

Louisiana State University

## LSU Scholarly Repository

---

LSU Doctoral Dissertations

Graduate School

---

7-5-2021

# Prüfung: A Deconstruction of Assessment Across Three Languages

Thomas Erich Benz

Follow this and additional works at: [https://repository.lsu.edu/gradschool\\_dissertations](https://repository.lsu.edu/gradschool_dissertations)



Part of the [Adult and Continuing Education Commons](#), [Critical and Cultural Studies Commons](#), [Discourse and Text Linguistics Commons](#), [Educational Assessment, Evaluation, and Research Commons](#), [Educational Sociology Commons](#), and the [Sociology of Culture Commons](#)

---

### Recommended Citation

Benz, Thomas Erich, "Prüfung: A Deconstruction of Assessment Across Three Languages" (2021). *LSU Doctoral Dissertations*. 5585.

[https://repository.lsu.edu/gradschool\\_dissertations/5585](https://repository.lsu.edu/gradschool_dissertations/5585)

This Dissertation is brought to you for free and open access by the Graduate School at LSU Scholarly Repository. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Scholarly Repository. For more information, please contact [gradetd@lsu.edu](mailto:gradetd@lsu.edu).

# ***PRÜFUNG: A DECONSTRUCTION OF ASSESSMENT ACROSS THREE LANGUAGES***

A Dissertation

Submitted to the Graduate Faculty of the  
Louisiana State University and  
Agricultural and Mechanical College  
in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

in

The School of Education

by

Thomas Erich Benz

B.A. Johannes Gutenberg Universität Mainz, 2014

M.A. Goethe Universität Frankfurt, 2017

August 2021

## Acknowledgements

With great pleasure I would like to express my profound gratitude to the facilitators, enablers, role models (of any form), educators, counselors, guides, trainers and friends, which have accompanied me in any form on this journey. To my parents who have made me understand long suffering and perseverance for as long as I can think and remember. *Ohne eure Vorbilder würde ich heute nicht verstehen, dass das Leben aus mehr besteht, als aus den Elfenbeintürme in denen ich mich so lange bewegte.*

I foremost would like to express my heartfelt thanks and outmost appreciation to my professor and mentor, Dr. Kim Skinner. She expressed compassion through allowing me to grow. Her support and encouragement motivated me to work harder and get to where I am today. I am indeed grateful for her meticulous and constructive feedback, attention, support, patience and guidance she bestowed upon me. I express my appreciation to Dr. Kenneth Varner who paved the way for me to commence my program at LSU. I appreciate him most for his wonderful ability to introduce opportunity and enact change.

I also extend my special thanks to Dr. Robyn Andermann. It has always been an honor to work with you and learn by example from you what it means to teach and to inspire students in the classroom. I thank her husband, both of them truly supported me, which mattered a lot as an international student during my program.

I appreciate Dr. Denise Egéa. *Merci à la personne qui représente la plus grande académique que j'ai jamais eu le plaisir de connaître, et qui m'a introduit aux théories de curriculum. Merci pour le dévouement avec qui vous avez accompagné mon projet de dissertation depuis le début.*

I honestly appreciate Dr. Roland Mitchell and Dr. Jacqueline Bach for serving on my committee and sharing their knowledge and experience with me. I appreciate the time they took to diligently go through my work and give me constructive feedback; it has meant a lot to me. Thank you for being available for me every time I needed you.

Furthermore, I would like to extend my thanks to all my professors who advised me throughout my program. I would also like to thank to my Dean's representative who has always (surprisingly) constructively contributed to my dissertation project. Thank you for your honest time, effort and the provision of new perspectives.

Finally, I want to excuse myself to my wife, who is married to a doctor, but unfortunately not to a medical one, or one of the sciences. Learning a lesson from a fellow historical citizen of my birthplace Trier, it appears we need to choose to either live a life of hardship and write a manifesto, or choose to live a happy life, but risk remaining invisible. With you and for our future, I would always risk the latter.

## Table of Contents

Acknowledgements .....	ii
List of Tables .....	vi
List of Figures .....	viii
Abstract .....	ix
Chapter One. Introduction .....	1
Statement of the Problem and its Significance .....	3
Research Questions and Significance: .....	8
Mode of Inquiry .....	10
Chapter Two. Theoretical Framework and Review of Literature .....	13
Intersections with Socioeconomic Status, Ethnicity, and Migration .....	13
Cultural Bias, Language and Diversity in Testing .....	21
Beneath the Educational Enterprise: Social and Cultural Capital .....	29
Competences as Social and Cultural Concepts .....	35
Chapter Three. Methods, Methodology, Sampling, Data and Data analysis .....	45
Methodological Considerations .....	45
Description of the Exam Data .....	56
Narrative Synthesis .....	63
Network Text Analysis .....	81
Chapter Four. Results .....	97
Comparison of Results: Narrative Synthesis .....	97
Results: Narrative Synthesis per Exam .....	111
Findings for the Abitur .....	114
Findings for the Baccalauréat .....	128
Findings for the SAT Test .....	141
Results Network Text Analysis .....	153
Findings: Representation of Choice .....	153
Explanations: Availability and Degree of Specialization .....	158
Explanations: Conductivity and Answer Choice .....	159
Findings: Representation of Content per Subject .....	161
Visualization of Semantic Networks .....	161
Complexity Values for Each Exam .....	171
Chapter Five. Discussion and Further Research .....	173
Abitur: Insights into Maturity and Choice .....	173
Baccalauréat: The Importance of Writing .....	175
SAT: Fractured Knowledge or Accessibility? .....	177
Different Approaches to Discovering Cultural Capital and Cultural Bias .....	179
Contributions Towards a Pedagogy of Assessment .....	183

SAT: the English Subject Test and Recent Developments .....	184
Insights and Limitations: 250 Years of Texts Across Three Languages .....	185
Insights and Limitations: Access and Authority .....	187
Insights and Limitations: Contradicting Axis of Choice? .....	190
Différance .....	192
Appendix A. Sample Table for Text included in the Narrative Synthesis.....	194
Appendix B. Coding Tables for Network Representation of Content. ....	201
Appendix C. Coding Tables with Raw Numbers for Network Representation of Choice .....	204
Appendix D. IRB Document.....	206
References .....	208
Vita.....	208

## List of Tables

1.1. Functions of Assessment Translated.....	5
2.1. Functional Domains of Assessment Translated.....	37
3.1. Between State Statistics for the Abitur.....	59
3.2. Impact Factor SAT Test.....	74
3.3. Impact Factor Baccalauréat.....	75
3.4. The Results of Sampling Exam Content.....	84
4.1. Summary of Narrative Synthesis: Production, Interpretation and Manipulation of Texts.....	99
4.2. Summary of Narrative Synthesis: Communicative Competence.....	101
4.3. Summary of Narrative synthesis: Modes of Reasoning.....	102
4.4. Summary of Narrative Synthesis: Themes with Partial Overlap.....	104
4.5. Summary of Narrative Synthesis: Culture and Morals.....	107
4.6. Summary of Narrative Synthesis: <i>Sitte, Reife</i> .....	108
4.7. Summary of Narrative Synthesis: Readiness.....	109
4.8. Summary of Narrative Synthesis: Social Factors for the SAT.....	110
4.9. Narrative Synthesis: Reference Table for Sampled Texts.....	113
4.10. Abitur: Communicative Competence.....	115
4.11. Abitur: Production, Interpretation and Manipulation of Text.....	118
4.12. Abitur: Mathematical, Scientific Reasoning.....	120
4.13. Abitur: Oral Mastery.....	122
4.14. Abitur: Classic Competence.....	123
4.15. Abitur: Mental Cultural Ideas.....	125

4.16. Abitur: <i>Reife</i> .....	127
4.17. BAC: Communicative Competence.....	129
4.18. BAC: Production, Interpretation and Manipulation of Text.....	131
4.19. BAC: Modes of Reasoning.....	133
4.20. BAC: Oral Mastery.....	135
4.21. BAC: Classic Competence.....	137
4.22. BAC: Culture and Morals.....	140
4.23. SAT: Communicative Competence.....	142
4.24. SAT: Production, Interpretation and Manipulation of Text.....	143
4.25. SAT: Modes of Reasoning.....	145
4.26. SAT: Readiness.....	147
4.27. SAT: Alignment with Authority and Degree of Social Assimilation.....	149
4.28. SAT: Social Factors of Learning .....	151
4.29. Complexity Values for Each Exam.....	172



## List of Figures

1.1. Visualization of Planned Inquiry.....	11
2.1. Assessment Landscape in France.....	19
3.1. Visualization of Planned Inquiry and Data Analysis.....	55
3.2. Generating a Population N.....	70
3.3. NTA, Application to Exam Data.....	82
3.4. Visualization of Exam Content.....	91
3.5. Visualization of Representation of Choice.....	94
4.1. Narrative Synthesis: Sample Size for Each Exam.....	112
4.2. Representation of Choice: Math.....	154
4.3. Representation of Choice: Native Language.....	155
4.4. Representation of Choice: History.....	156
4.5. Representation of Choice: Sciences.....	157
4.6. Representation of Choice: Foreign and Classic Languages.....	158
4.7. Representation of Content: Native Language Literacy.....	163
4.8. Representation of Content: History.....	166
4.9. Representation of Content: Mathematics.....	170

## **Abstract**

This research aims at deconstructing and questioning certainties about assessment as an educational institution on its most fundamental levels. To achieve that, I am utilizing theoretical frameworks inspired by concepts on the existence of cultural and social capital, by artifact mediated cognition, and by a recently proposed discipline on pedagogy of assessment. The research operates with the application of narrative synthesis and network text analysis of material, on which they have not previously been used. As such, I aim to contribute to a methodological application of both methods on exam data, understood as the totality of curricular documents which govern exams, of question sheets which make up the visible part of exams, and of endorsed exam responses which make up the concealed part of assessment. To accomplish the latter, I alternate between structural and post-structural paradigms, aiming for an integrality of both. In this work, I further provide a comparative analysis, comparing and synthesizing results from analyses on the French Baccalauréat, the German Abitur, and the US American SAT, using sources in French, German, and English. Neither the methodological mix, nor that specific comparison has been done before and I solve several issues, which arise through translation for example, through new applications of existing approaches of meta-description and categorization of textual content. The inclusion of three large scale exams from three different linguistic and cultural contexts allowed me to reflect on similarities and differences, which exist in the German, French and US-American societies, in terms of cultural or social competences students are expected to display when aiming for higher education. It also allowed me to reflect on issues arising through access, transparency and a lack thereof, which I discuss on the basis of sampling and procuring necessary texts and documents.

## Chapter One. Introduction

The German researcher Tinnefeld (2013) attempted to provide new lenses on already existing theoretical and practical approaches to assessment in education, in history and in the present time, ultimately aiming to lay foundations for a novel discipline, a *pedagogy of assessment* (p. viii). In his work he calls for critical reflections on testing and assessment that need to draw from philosophy, history, as well as from practice. Very few authors have actually taken up this exercise so far, among them, mentions are Shohamy (2016), Shohamy, Or and May (2017) who published and researched within context of language assessment in and on the English language. In Shohamy's (2016) *The Power of Tests: A Critical Perspective on the Uses of Language Tests* the author wrote:

Critical testing implies the need to develop critical strategies to examine the uses and consequences of tests, to monitor their power, minimize their detrimental force, reveal the minuses, and empower the test takers. (p. 131)

Her focus of research in that publication rests on critical analysis of language tests and the revelation of the social power structures in which these tests are embedded in. In Shohamy, Or and May (2017) *Language Testing and Assessment, Encyclopedia of Language and Education*, the authors point out the significance of language in such analyses:

Regardless of the purpose of an assessment – either to assess a test taker's language proficiency or their content knowledge in a particular area – these two constructs cannot be so easily disentangled. Any assessment of content will involve language, and any assessment of language that will be useful for making inferences about a test taker's ability to use language in a context outside the test itself will involve some content or topical knowledge (p. 4).

The complexity of *assessment* and the revelation of analytic lenses to understand *assessment* as not just test, exam, educational obstacle, occupied me throughout the following

---

<sup>1</sup> All German and French texts translated either directly or indirectly by the author. Direct translation immediately follows after the quotation, indirect translation in the consecutive paragraphs.

research. Socioeconomic factors played a role (Wells, 1999; Edwards, Coates & Friedman, 2012; Gök, 2010), but beyond that, I aimed to discover new analytical lenses through which assessment can be understood (Tinnefeld, 2013; Vygotsky et al., 1978). The above quote was informed by the authors' experience that proficiency and content testing commonly serve different purposes, but effectively entertain various entanglements with each other, just as much as Tinnefeld's hypothesis surrounding a new discipline, a *pedagogy of assessment*, is informed by an understanding of assessment which involves several paradigms and the combination of several analytical lenses. Before taking a deeper look at testing as an educational practice, I believe it is necessary to explain my own positionality:

I encountered a major problem right at the beginning of my review into different forms of testing. One example would be several types of university entrance exams. Despite a seemingly overwhelming importance for cohorts of aspiring students worldwide, there appears to be an equally overwhelming lack of critical research done outside of the US-based SAT or ACT examinations. This impression is also informed by my own journey through higher education and my experience with examinations of all different sorts that I gathered in my roughly nine years that I have been pursuing higher education degrees, certifications, and to some extent, tutoring and teaching others who do so. Exams and more specifically standardized exams related to access to higher education resources, have acted as gatekeepers to merit and achievement for myself and for many of my colleagues, especially when the students were not native to the country and academic environment in which the exam was administered. I have observed foreigners (without going deeper into definitions of *foreigners* here) struggling with language, content, and exam structures, to which they were not used in a country foreign to them, and I have, myself, been that foreigner several times during my educational journey up to this point.

Then again, coming from Europe, I have also experienced firsthand the value that is given to these foreign experiences in my and *their* countries of origin, a form of achievement that merits all the pain of adaptation and initial failures. This led me to a problem often overlooked in native discourses here in the United States. Where university entrance exams are concerned, the stakes are often high and the pressure resting on the shoulders of foreign or minority students is often significantly more than just academical in nature. These experiences, as well as the formal analyses of above authors (Edwards, Coates & Friedman, 2012; Gök, 2010; Shohamy, 2016; Shohamy, Or and May (2017); Tinnefeld, 2013; Wells, 1999) support my current interest into developing a more critical, deeper understanding of the different social, educational and cultural dimensions that are entangled in modern education, when testing is concerned.

### **Statement of the Problem and its Significance**

While the constructs underlying standardized testing and psychometrics appear to be topics widely discussed in scientific literature, my initial literature review into the topic assessment revealed an overwhelming absence of critical questions regarding assessment as an institution within education, as well as lacking a deeper analysis of the relationship between content, structures, and socioeconomic factors surrounding larger, *high stakes testing* situations, like encountered by students taking university entrance exams around the globe. To unpack these perceptions will be the aim of this first segment. A definition most appropriate to describe the notion of lack could be taken from Blaikie and Priest (2017):

Unlike classical positivists, who believed that the only things we can regard as being real are what we can observe with our senses, Critical Realists claim that, while we can observe regularities in the world around us and this is the starting point in any scientific investigation, we may not be able to observe directly the fundamental elements, the causal structures and mechanisms, that produce these regularities. Hence, the task of science is to discover these underlying elements, describe their nature and show how they produce observed regularities. (p. 169)

So, in other words, where I discovered quite an abundance in research tying the socioeconomic factors with the establishment of standardized examinations, a fundamental inquiry about the thoughts and ideas, which determine the shape, form, and content of exams is missing.

University entrance exams make up a subcategory of standardized examinations. With reference to common discourse within English speaking media there exist several denominators. One example would be *high stakes testing*, which relates to the gravity of the outcome of this form of examination (cf. Salaky, 2018; Partnership, 2014). A good summary of definitions for high-stakes testing is provided by Au (2007):

A test is high-stakes when its results are used to make important decisions that affect students, teachers, administrators, communities, schools, and districts (Madaus, 1988). In very specific terms, high-stakes tests are a part of a *policy design* (Schneider & Ingram, 1997) that links the score on one set of standardized tests to grade promotion, high school graduation and, in some cases, teacher and principal salaries and tenure decisions. (p. 258)

Au's (2007) account also serves to explain the categorical relationship between university entrance exams and their parent category standardized examinations. The high-stakes element is socioeconomic in its nature as it relates to policy decisions, promotion, and can even affect salaries and tenure. Beyond socioeconomic elements, assessments are considered institutions which carry a myriad of meanings and hold just as many stakes for different stakeholders. With the need to even begin to problematize these complex relationships and identify different stakes, I encountered the discourses surrounding the initially mentioned pedagogy of testing.

Furthermore, I encountered an intriguing taxonomy provided by German researchers Müller and Schmidt (2009). They provide a functional perspective on assessment based on the following three umbrella categories: function for recruitment, function didactics and function dominion and social assimilation (translated, p. 25). This taxonomy, together with the idea of a pedagogy of assessment provided the foundation and inspiration for better understanding and defining

*developmental and learning ideals* when assessment is concerned. Such understanding did not remain shallow or confined to one discipline or even paradigm of research. The following figure is taken from a German publication about *pedagogy of assessment* from Müller and Schmidt (2009, p. 25):

Table 1.1. Functions of Assessment Translated

<b>Function: Dominion and social assimilation (reproduction of the system)</b>
initiation
distribution of Status
legitimization
<b>Function: Recruitment (selection, election, allocation, standardization)</b>
cohort placement
selection within cohort
proof of qualification and competence
<b>Function: Didactics</b>
Test as timer and content structurer in a course
provision of orientation with regard to learning goals for lecturers and students
extrinsic motivation to learn
diagnostic tool
auditing over teaching and learning success of teachers and students

While going deeper into the sub-criteria listed under each umbrella term in the Table 1.1 above (adapted and translated from German) for the purpose of explaining developmental and learning ideas, I found the overarching categories helpful in identifying the stakes involved in high stakes testing, as well as coming to a conclusion about which form of assessment to consider excelling in most of these aspects. As a result of a preliminary review of the literature, I chose to focus on exams enabling access to higher education. Despite the overall lack of critical literature explaining the interaction of all the stakes involved in collective educational testing, assessment for higher education produced the richest results. In the following sections I discuss such university entrance exams while juxtaposing some of their aspects with the categories

above and shortly discussing each umbrella term. Assessments for higher education play a larger role in the socioeconomic fabric of different societies:

*High-stakes* tests in the form of university entrance exams are present within many different educational systems worldwide. Though abundant, it was hard to come upon a comprehensive, globally comparative study on these types of examinations. Edwards, Coates and Friedman's (2012) paper summarizes one of the rare attempts undertaken in that direction. Their paper was the result of an OECD funded research project and the authors describe the role of university entrance examinations as follows:

From the international study which inspired this paper, it appears that the main driver for the inclusion of such tests in selection processes is the belief that they can offer a more efficient, valid, or at least supplementary means for selecting the most appropriate candidates for university. In certain countries, tests are administered because of a lack of consistency in secondary school assessment processes, while others use tests in conjunction with various measures of achievement to strengthen selection methodology. (p. 4)

High stakes testing is the product of the demands that come with selective social pressure, and an economic situation of scarcity where the demand exceeds the offer. Müller and Schmidt (2009) established that assessment serves a need for *dominion and social assimilation* when it represents a rite of passage into certain social formations or into having access to limited resources. This also includes raising the social status of an individual work through granting prestige (p. 26). Edward, Coates and Friedman's (2012) perspective on university entrance exams comes from a comparative analysis of high stakes testing in more than 8 countries around the world. While they mention China, Turkey and South Korea as examples for state run, nationwide university entrance exams, most other countries around the world entertain a similar, semipublic or entirely private high stakes testing infrastructure (see Edward, Coates and Friedman, 2012, pp. 8-12).



Despite the social and economic stakes attached to these forms of tests, their intended purpose always goes beyond that. Regarding university entrance exams, Phelps (2009) discussed some of the expectations attached to these tests as follows:

The use of high-stakes tests originates in part from the desire to measure an important outcome—student learning—and to hold someone or something accountable for the result. (...) A great deal is riding on the promise of integrating cognitive psychological principles with educational measurement. (p. 237)

Due to my personal language constraints, for the purpose of this paper, I cannot delve deeper into a critical analysis of above grand examinations but have to restrain myself to university entrance exams done in German, English and French. Having access to the native language in which these tests are designed and admitted is imperative for this project, as an analysis must rely on describing the fundamental elements, as well as any causal structures between content and context (cf. Blaikie & Priest, 2017, p. 169).

The title of this work references deconstruction, a notion continuously pondered by Jacques Derrida and others (Derrida, 1990, 1993; Critchley et Al., 1996 for example). Though in the name, I cannot use deconstruction as a theoretical framework in the like others presented in the following chapters. The notion of deconstruction rather came to me at the end of my analyses and followed my mental struggles to better define what this work *is*. I will return with a conclusion at the end of this text, and instead focus for the moment on what this work actually *does*. In Derrida (2020) the authors note: “(...) the deconstructive process comes not from the reader/critic but from the text itself; it is already there, it is the tension between what (the text) manifestly means to say and what it is nonetheless constrained to mean” (p. 274). It is noted for example in Wood and Bernasconi (1988), that Derrida has tried to avert negative or violent notions of destructions, which arose in the context of the American school of poststructuralism (pp. 3-4). Though I aim to contribute to a critical analysis of high stakes exams with this

research, I will not be able to perform the critical resituating of my analyses. Given my limitations it has to suffice here to dissect and to reconstruct the high stakes exams in focus. Inspired by the following statement: “Rather than destroying it was also necessary to understand how an “ensemble” was constituted and to reconstruct it to this end” (Wood and Bernasconi, 1988, p. 3). The term deconstruction appears in the title as it permits me to interpret my own work. On a similar note, Derrida warns in Wood and Bernasconi (1988) against the appropriation of deconstruction either as a method or as a type of analysis:

All the same, and in spite of appearances, deconstruction is neither an analysis nor a critique and its translation would have to take that into consideration. It is not an analysis in particular because the dismantling of a structure is not a regression toward a simple element, toward an indissoluble origin. No more is it a critique, in a general sense or in Kantian sense. (...) I would say the same about method. Deconstruction is not a method and cannot be transformed into one. Especially if the technical and procedural significations of the word are stressed. It is true that in certain circles (university or cultural, especially in the United States) the technical and methodological "metaphor" that seems necessarily attached to the very word deconstruction has been able to seduce or lead astray. (p. 3)

Instead of being “seduced” or “lead astray” I was rather inspired by an element of Derrida’s theorem. The idea of *différance* expressed by Derrida (1973) inspired the methodological alterations I was forced to do to apply the *network text analysis* in combination with a dataset it has never been applied to before. As I will theorize later, it did not suffice to assume the existence of hidden structures (or a *hidden curriculum*), the reading of the non-present of exam data actually had to become an integral part of the analysis. It is the synthesis of all following elements, methodological and analytical, which inspired me to describe the following research as a deconstruction, in lack of a better interpretation.

### **Research Questions and Significance:**

In order to better understand the stakes and the purposes of modern collective assessment in education, I will be guided by the following questions:

1: Which social or cultural competencies are the SAT Test, the Abitur and the Baccalauréat intended to assess?

2: How are social or cultural capital embedded within and revealed in comparison across the Baccalauréat, Abitur and SAT Test?

Each of these research questions will be used to inform and justify the use of one of the three methodological approaches. The initial literature review revealed that assessment is neither socially equal, nor contextually sensitive and in many cases, with divides happening along cultural and language lines.

One purpose of the study is a multilingual contribution towards a critical pedagogy of testing, as proposed by Tinnemann (2002, 2013). A multilingual comparison of exam content and assessment in three different countries has never been done and can serve as a significant contribution towards a pedagogy of testing. While many studies have looked at internal coherences or socioeconomic factors of assessment, no study so far has tried to combine different disciplines and analytical lenses, to tie together content, context and discourse of, and surrounding university entrance exams in three different countries and languages. All this needs to be done here while juxtaposing the results forming one overarching research.

There truly appears to be a gap in the literature when searching for a critical, depth analysis of these exams. While in the German context, such attempts appear to be ongoing, my research will also hopefully provide an important contribution to introducing a pedagogy of testing into the English discourse and provide a contribution to the former. The methods I intend to use have never been combined and juxtaposed, nor have narrative literature reviews and network text analyses ever been done with exams. In general, a content analysis of the German and French exam samples, which I will settle on, has never been attempted, though both have

been around in different forms for more than 200 years. I also aim to provide new applications of the methodologies behind qualitative metasynthesis and network text analysis as per my unusual data sources, and finally through situating both methodologies in a theoretical framework inspired by Bourdieu (1977, 1986) and Lew Vygotsky (1978).

Lastly, not just the methodological mix is novel, but it is also rare to find publications doing large scale sampling and direct comparison of research conducted in three languages and three different social, educational, and arguably cultural environments. Not just the exams need to be translated, but sometimes also the scientific theory on which I intend to base the analysis. Adding to the appeal of the comparative language framework would be the fact that I intend to use first language sources to provide theoretical perspectives, from Adorno, Bourdieu and Marx to the provision of a major theoretical framework on didactics of assessment coming from German discourses.

### **Mode of Inquiry**

I am using two methods, as pictured in Figure 1.1, the first being qualitative meta-synthesis (specifically: narrative synthesis), the second being network text analysis. Both methods involve separate samples that are conceptually linked, but practically drawn from different sources. The sample for the qualitative meta-synthesis consists of publications by exam-stakeholders, or publications that are otherwise considered relevant. For this step, I use stratified purposeful sampling as outlined later in Chapter 3. The results of this method pertain to each exam individually at first and allow for a comparison in a subsequent analytical step. On the other hand, the second method, network text analysis, requires a cross-comparison among the exams to begin this analysis. The network text analysis operates with the exam content as

sample, and the analysis focuses on subject content, consisting of exam questions and responses, versus exam formalities. Figure 1.1 exemplifies both methodological approaches:

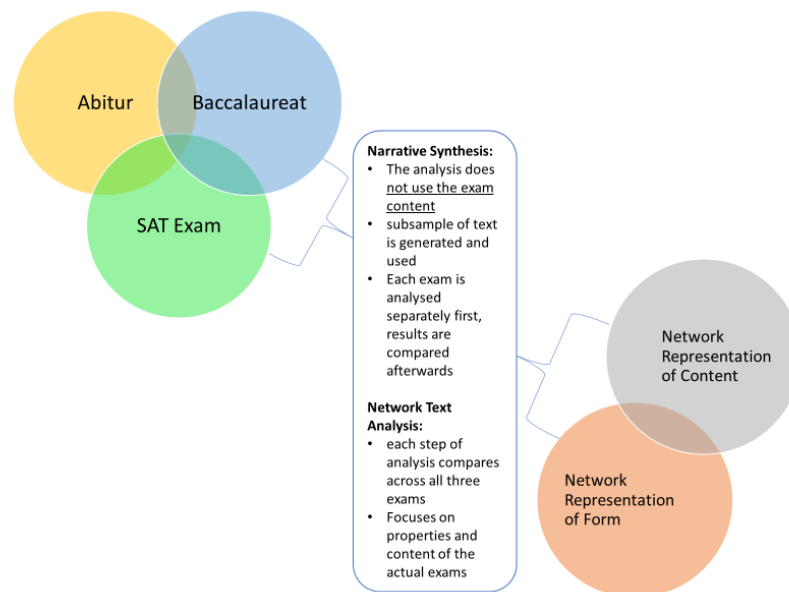


Figure 1.1. Visualization of Plan to Inquiry

In Chapter 3, I include discussions of methodological adaptations that were necessary to use both methods in the present scenario. These discussions include reliability and validity of my newly found adaptations of the methods supported by the theoretical framework. While I discuss the sampling for each of these three exams, a further layer of complexity presents through the progressive application of each method, as well as through the necessity of a comparison of results in Chapter 5. Despite a high theoretical complexity, with which I need to deal before, during, and after the application of each method, my proposed research is empirical, as it works with text as object, and with textual content as data in different forms.

In accordance with Blaikie and Priest (2017), the underlying paradigm of my research has many elements of critical realism, as it heavily relies on “retroductive logics” (p. 13), enabling a retracing of surrounding social, economic, racial and cultural realities to the patterns those

realities leave within cultural produce, objects of everyday life, and so forth. My research works with empirical observations of exams as cultured objects in many if not all contexts in which they appear. In addition, Blaikie and Priest (2017) summarized that notion: “(...) observed regularities are eventually explainable in terms of underlying 'real' causal structures and/or mechanisms” (p. 13). As my research design relies on three methodologies, each following a mix of different sampling techniques, so also does my research eventually include other paradigms.

## **Chapter Two. Theoretical Framework and Review of Literature**

The following chapter introduces the theoretical lenses, which I require to explain and justify my research design and the modifications I undertake to the methods and methodologies used in this work. It also provides a review of the relevant literature.

### **Intersections with Socioeconomic Status, Ethnicity, and Migration**

The constructs underlying standardized testing and psychometrics appear to be topics widely discussed in scientific literature (Abedi, 2012; Carver, 1974; Furr, 2017; Gordon & Terrell, 1981; Tinnefeld, 2013); nevertheless, I did not come upon a satisfying “depth-realist” (Blaikie & Priest, 2017) analysis of the socioeconomic structures linked to high-stakes testing. There is research available from all around the globe, exploring several societal phenomena linked to high stakes testing regarding recruitment. Taken from an older source, Takeuchi (1997) described the system of Japanese *juku*, so called “cram schools” (p. 184) and the necessity of cram schools in Japan through changed socioeconomic factors attributed to education.

In traditional society, status was ascribed, for example, by inheritance. Under such circumstances, education reproduced the existing social stratification and educational institutions were vehicles for the preservation and transmission of life styles. In modern industrialized society, with its emphasis on rationality and efficiency, educational credentials are used to assign people to occupations. As a result, educational attainment has come to have a greater direct impact than original social class on one’s socio-economic status attainment. (p. 184)

Takeuchi (1997) creates a direct link between feudal, hierarchical elements in modern Japanese society striving to preserve status in a modern society, and the need for high stakes testing. The Daigaku Nyūshi Sentā Shiken, which persists until today, serves as a means to redistribute socio- economic status and eventually separates the teenagers who are to be in high status, leading positions, from those who will make up the working class (pp. 183-187).

Looking at another example from the Turkish national university entrance exam, Gök (2010) pointed out how this high-stakes exam is rather serving to re-affirm (p. 127) already existing socio-economic differences in the population:

In such an educational system [the Turkish], namely in a system where progression between levels depends on selection exams, and in a society where education is not seen as a right, private preparation courses are the institutions that aggravate inequalities in education. (p. 126)

Gök (2010) perceived private education courses as key for success in the national Turkish college entrance exam. As only the already wealthy parts of the population can afford these private courses, she stated that they “(...) function as instruments normalizing the problem that those who cannot pay are doomed to fail and bound to be excluded from the formal educational system” (p. 127). With the help of Foucault, she continued to describe the categorization of the individual student through the institution of standardized testing: “education is determined by market dynamics and turned into goods and services, which can be sold and purchased” (p. 129). The following examples will be specifically questioning the *didactic purpose* of assessment.

The first of Müller and Schmidt’s (2009) functions is *didactics*, subcategorized into *didactic purpose*. That high stakes exams around the world entertain multiple intersections with categories such as culture, race and socioeconomic status became evident, but how about the dimension of *didactics*?

*Zum anderen erlauben Prüfungen aus diagnostischer Perspektive Rückschlüsse auf bis dato realisierte Lern- und Bildungsprozesse sowie mögliche künftige Verbesserungsmaßnahmen. Wie von verschiedensten Seiten betont wird, ist die hiermit verbundene Lernchance im Kontext von Prüfungen nicht nur auf die Prüflinge selbst zu beziehen, sondern vielmehr auch auf die prüfende Person sowie das prüfende System als gesamtes (...). (p. 27)*

On the other hand, does assessment have diagnostical potential in deducing on the quality of already occurred learning and development, as well as for informing the creation of future curriculum. As has been argued in several articles, the potential for development here is not to be seen solely with the examined, but to an even stronger degree with the examiner, if not also for the educational system as a whole.<sup>1</sup>



What is unique in the above approach is to postulate testing not just as a chance to gain information on the learning and development of the students, but rather to see it as a chance for the educational system into which the assessment is embedded, to reflect upon itself. The extent to which this could apply becomes clearer in the following mentions.

Returning to the US-American society, the role of race as a variable determining the outcomes of standardized testing appears to be extensively researched. Similar to the examples above, the US-American Scholastic Assessment Test (SAT) or American College Test (ACT) have been attributed with socio-economic achievement beyond their assessment of learner's aptitude (see Clark, Rothstein & Schanzenbach, 2009; Thompson, 2012; Wheeler 2013). The most impressive account for that relationship might be the role which, at that time, unexpectedly fell to the College Entrance Examination Board in the desegregation movement of schools during the 1960s. The CEEB (College Entrance Examination Board) was the entity that designed and administered early, standardized predecessors of today's SAT and ACT to evaluate the educational achievement of aspiring university students (Wheeler, 2013, p. 1). Wheeler examined how the SAT was initially designed by organs of the CEEB to "(...) keep black applicants out of their institutions" (p. 3). The reason these tests could effectively serve as gatekeepers, not letting segregated minorities like African Americans in, was that their design inherently catered to materials and the educational experience, in white schools (pp. 3-5). Though Wheeler (2013) provided a non-critical, historic account of the desegregation of SAT testing, his work serves as a rich source to further deduce the socioeconomic causes linked to high-stakes testing at the intersection of race with ethnical minority. Wheeler, through the words of a minority counselor in the 1970s, not long after segregation officially took hold, described these intersections as "(...) two main concerns – minority students' public access to SAT

information and their freedom from fear during testing” (p. 202). A lack of public access is a decisive factor, as the design of the SAT openly required additional preparation for students to succeed. Similar to the Turkish and Japanese examples above, university access depended not only on the developmental level of the individual student, but more so on access to social and economic resources, complementary to school education. The role that racial or ethnical background played here, should lead to an idea of similar difficulties foreign test takers must face.

As minorities within the US, it is questionable whether foreign applicants have the necessary social resources available to properly prepare for the SAT in their home countries. Unfortunately, any further or deeper analysis into this question is hindered by several roadblocks. Based on my literature review, there is no information available that takes a critical look at the socioeconomic factors involved in foreign students taking the SAT as a means to migrate for studies. A well-documented example for the intersections between nationality, language and high-stakes testing, which I encountered in the US, are the policies implemented by the El Paso school district between 2004 and 2010.

Researcher and journalist Kamenetz (2016) revealed how these policies formed a campaign of conscious suppression of Latin American immigrant children who migrated at an age close to their 10th grade (pp. 22-26). “Credits were deleted from transcripts or grades changed to move students forward or back a grade in order to keep them out of the tenth-grade test” (p. 24). The tenth-grade test she refers to, was a high-stakes, standardized test similar to the SAT in its socioeconomic dimensions. Money, structural subsidies, and political reputation elevated the results of this test far beyond just the sphere of educational assessment. Migrants were systemically removed from the test or put into a lower grade as continued practice has

shown that they underperformed due to their different level of acculturation and English as their second language. In reality language and cultural background, played a larger role here than the educational maturation of these immigrant children, and until these policies were uncovered in 2010, the district's performance was outstanding. This example hints towards a larger underlying negative correlation between the performance in that 10th grade standardized exam, and second language, difference in culture, and migration background.

While the US College Board (2019a) appears today to be striving to provide an account of academic and developmental maturation through the SAT and ACT, which is accurate for each student independent of race, ethnic, or cultural background, elsewhere high-stakes tests are used under a completely oppositional paradigm. University entrance exams in Germany and France are most often used because there is the general assumption that foreign university applicants do not have the necessary skills or educational ability to begin studies at a higher education institution in these countries. For as far as the German *Feststellungsprüfung* is concerned (Deutscher Akademischer Austauschdienst, 2019), it serves as a means to decide about the equivalency of non-German high school credentials with those of German students. Here appears an even more pronounced lack of information about the underlying causes, that led the German education system to recognize some countries' high school diplomas as equivalent to the German one, while rejecting others. The only resource for that matter is a registry provided by the German central office for foreign education, *Kultusministerkonferenz – Anabin*, which lists countries and high school diplomas, as well as their recognition for direct access to higher education in Germany or their lack thereof (Kultusministerkonferenz, 2019). There is a lack of any rationale that would explain the non-recognition of a certain credential. Through *Anabin* future students may be able to discern, whether they require to take the German high stakes

exam, *Feststellungsprüfung*, but not why. I found one document describing the *Feststellungsprüfung*, its goals and structure in more depth. From a private German research institute (ITB Consulting, 2019) that majorly constituted in the design of this high-stakes exam, the following passage is taken:

*Eine wichtige Fragestellung ist daher, inwieweit die TestAS-Ergebnisse mit den Schulleistungen von Oberstufenschülerinnen und -schülern vergleichbar sind und inwieweit der TestAS eine geeignete und angemessene Schwierigkeit aufweist. (p. 2)*

A major determinant for the test's validity therefore consists of the correlation of the test takers' test results in terms of performance, with the performance of German high school students, which allows for conclusions to be drawn toward the appropriateness of the exam's difficulty.

Apart from being highly German centric, this passage reveals that the major rationale for the creation of the German *Feststellungsprüfung* is a direct comparison between foreigner's educational achievement with that of German graduates of the *Abitur*. The *Abitur* is a comprehensive examination which happens at the end of high school in Germany, after 12 years of schooling. The document reveals a shocking vacuum of theoretical reflection on the socioeconomic circumstances of foreigners, as well as on the cultural differences or rationales of comparison not related to testing. The non-proclaimed truth that reads within this document is that foreign students are expected to assimilate to the extent of equal sociocultural development with German high school graduates. Page 5 of the report caters to this fact to some extent by declaring that the presence of sociocultural difference is in fact a confounding variable that the test cannot account for. After this initial research into the German high-stakes exam, it appears that here, we are even further away from an in-depth realist description of this phenomenon than in the US-American context.

In France the current testing landscape appears more complex. The following diagram is taken from a report funded by the French ministry of education and the OECD, authored by Dos Santos, Rakocevic, Lefresne, and Trosseille (2012):

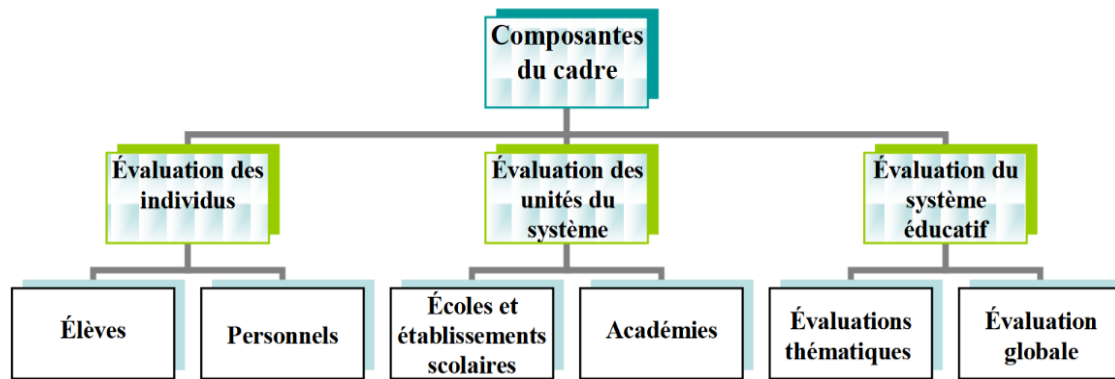


Figure 2.1. Assessment landscape in France.

Source. Dos Santos, Rakocevic, Lefresne, and Trosseille, 2012, p. 29.

Standardized exams are common practice, and through this diagram the authors reveal three stakeholders, who base their decisions on the administrations of different forms thereof. The left box refers to examinations taken by individuals, by students and by administrators, the middle box refers to evaluations of educational bodies of different status, and the right box refers to evaluation of the productivity of the educational system in terms of learning outcomes. For the latter, Dos Santos et al. (2012) refer to the *PISA* study among other things (p. 29). What is most interesting in this French landscape for my research to be conducted, would be the fact that the country has a de facto national, collective standardized entrance exam, the *Baccalauréat* (*BAC*).

The testing curriculum for this annual end of high school exam is composed on a national level and does not distinguish between different high schools. Depending on the *BAC* results, the students have access to different formations at the French university level (Dos Santos et al., 2012, pp. 77-91). Though the colonial age is past, France still has several overseas departments

(Guadeloupe, Martinique in the Caribbean and Mayotte and Réunion off the coast of Africa), which French law and administration treats as an integral part of the French Republic.

The French researchers Delcroix, Forissier and Anciaux (2013) conducted a study towards the interactions between culture and language of native students in these overseas departments and their educational success within the French education system. Unsurprisingly they came to the result that the centralized model of the French education system interferes with local education and learning in the high schools, which they included in their observation (pp. 180-185). Though their study focused on sociocultural interactions between local learning and the French curriculum, it is implied that the BAC serves to disadvantage (pp. 159-160) these learners in many ways due to their cultural and social distance to the French mainland. These findings are supported by recent research showing weak overall performance of the overseas departments in terms of literacy (Jeantheau et al., 2016) and maths (Arneton, Bocéran, & Flieller, 2013).

Finally, exams are often stated to have lost their purpose (Reynolds & Brown, 1984; Manning & Jackson, 1984) as forms of assessment of educational development and have taken up a merely social function as gatekeepers to societal achievement and educational status. Widely unquestioned and unreflected upon as institutions of assessment, their societal functions and failures take center stage in most of the current English publications that I could find. Similar to the above tendencies of cultural bias in the French overseas assessments, cultural bias was already discussed during a period in the 1980s, in the wake of a debate on mental testing. Most prominent here are Reynolds and Brown (1984), as well as Manning and Jackson (1984). Manning and Jackson (1984) advocate for a departure from mental testing and question the purpose of assessment in the then already widely used ACT / SAT college entrance exam:

Many colleges ask applicants to take the College Board Achievement Tests as well as the SAT, and these achievement tests are designed, insofar as possible, to test an understanding of certain core concepts and bodies of knowledge common to high school courses in particular disciplines. But the SAT is much more widely used by colleges to obtain common assessments of students' levels of development in certain abilities (such as reading comprehension) that are fundamental to further academic work. (p. 195)

Manning and Jackson (1984) continue to argue that the SAT is often used as a general intelligence test, or a further obstacle, not designed to test content knowledge, but rather out of the pure necessity to select who has access to resources and who does not. They also take up more recent concerns about equity and achievement of non-native speakers, to the debate surrounding cultural bias in mental testing.

All research mentioned here pointed towards an urgent need of a relocation of testing within interdisciplinary approaches, as exams entertain conceptually rich relations with many different spheres of social interaction. There lies incredible value in more wholistic description of assessment, as each of the above research shows, as well as in a reinterpretation or even a deconstruction of what we know about testing.

### **Cultural Bias, Language and Diversity in Testing**

The debate surrounding cultural bias in testing (Kruse, 2016; Manning and Jackson, 1984; Misra, Sahoo and Puan, 1997; Reynolds and Brown, 1984) is born from research on international English language testing and has sprung up from analyses of English language and later other language exams. The intersections between language and test results have been a topic in the previous chapter and second language interference in education already has been of concern. Beyond this, there is overwhelming evidence in literature, that the mere use of language can encode differences in thought (Coates and Friedman, 2012; Edwards, 2014; Ehrhardt, 2014; Kruse, 2016). As a researcher, I am here faced with the task to continuously translate source material from French and German into English, which can include up to three mental procedures,

on which I am unable to reflect on at each given moment. The first procedure, understanding, is followed by description, which, given my methods, is just a precursor of meta-analysis. Though however impractical my given research design is, reflection on each of the three steps would hold the potential for a research of its own, as language and the understanding of lemmata, words, and sentences are socially and culturally determined. The first word of the title of this research serves as a good example. The German “Prüfung” has several English translations, test, trial, assessment, exam. It can hold technical connotations (auditing), as well as it can appear in religious contexts (god given trial, fated). The vast semantic relativity of this word stands in example with the role language and culture have with assessment as an educational institution.

Language and its use can provide insight into human thought and social discourse and therefore needs to be taken into account as another category of performative distinction in comparing all of the above high-stakes exams with one another and explaining the different situations that minority test takers are facing. The most prominent example would be the so-called Sapir-Whorf hypothesis. twentieth century linguists and language philosophers Sapir and Lee-Whorf stand at the beginning of a debate stretching across the humanities and spanning more than 30 years about whether, and to what extent, language structures thinking (for a summary: Ahearn, 2011). In a less extreme version, the structure of language has a significant structural impact on thought - debate surrounding the Sapir-Whorf hypothesis of the mid-20th century (Ahearn, 2011, pp. 69 – 73).

More recent discussions on the impact of linguistic categories on social and cultural organization are led by American, but also European authors until today. The first example comes from German author Blum-Barth (2016), who critically looks at diversity under the premise of culture and more specifically at processes which she describes as “transcultural” (p.



116). In the process of a critical literature review, Blum-Barth concludes that cultural diversity cannot be explained through assumptions of “trans-culturality” (*Trans-kulturalität*, p. 118). She argues that *transculturality* operates on underlying models of cultures as uniform and enclosed entities, which derive from antiquated ideas of nation states and seclusionism. For her, cultural demarcations are never clear enough to warrant for transient processes that are happening in our globalizing world. Instead culture is constantly shaped by its practical manifestations.

Hence, she claims that the idea of transculturality is inappropriate to serve as an analytical category or to understand diversity. Instead, she introduces the idea of linguistic hybridity (pp. 124-125). “Linguistic hybridity” (*Hybridisierung*), according to the author, is marked through the meeting and intersections of vocabularies from different languages. Such meetings are realized through amalgamations of word-elements from different words or the same word, in different languages. The most important function she distills from her analysis of linguistic hybrids is the revelation of hidden meaning, which ultimately points at diverse social practices (cf. pp. 124-126). The study of linguistic hybridity can lead to an analytical lens, through which researchers might be able to get a better grasp on the actual contexts and contents of modern “diversity.” Her approach establishes the possibility to discover social realities, as well as cultural differences through an inter-language comparative linguistic analysis; a framework, which underlies my own endeavor.

Another insight towards the relationship between language and social organization comes from the Luxemburgish researcher Ehrhart (2014). In her article *Le développement plurilingue et interculturel en milieu éducatif ouvert à la diversité* she looks at linguistic diversity through the lens of educators. Following years of research executed across country and language borders with educators and education administrators from France, Luxembourg, Russia

and Germany as research participants (pp. 90-91), she is convinced that language cannot exist without the underlying social realities to which it references, therefore exists in a close relationship with the latter.

After an in depth findings sections on the Luxembourgish research project LACETS (pp. 99-102), she succeeded in complementing her ideas concerning the inter-connectedness of linguistic development and cultural competence with research data: the successful development of linguistic competence was directly correlated with the degree of immersion of the students in the corresponding cultural networks, e.g., scholarly, familial and societal. Moving toward a concrete research design to answer the questions I posed during the introduction, I will keep borrowing from the assumptions of strong linguistic influence on thought. Language serves as a gateway to explore social reality and language encodes meaning, but language can provide ordinal data as well, which can be measured in an empirical fashion. From the onset of generative grammar through Chomsky (Bußmann, 2002; Chomsky, 1956, 1995) to the examples above, linguistic diversity has always presented a challenge and a source of information for research. Furthermore, the discussion about cultural bias in mental testing hints towards the deterministic nature of assessment based on where and in which culture it is executed.

While talking about linguistic differences, this segment asks for a reflection on cultural differences as well. Culture has already been defined in many different ways (for a debate: Clifford, 1988; Geertz, 2008; Lentz, 2011; Vertovec, 2001). The cultural anthropologist and historian Clifford (1988) reviewed it as follows: “‘Cultural’ difference is no longer a stable, exotic otherness; self-other relations are matters of power and rhetoric rather than of essence. A whole structure of expectations about authenticity in culture and in art is thrown in doubt” (pp. 13-14).

As a consequence, the ethnographic modernist of that time searched, in Clifford's (1988), words for the "universal in the local, the whole in the part" (pp. 3-5). The local was the observed information which the travelling ethnographer often collected through his or her records. Clifford referenced and challenged a certain type of belief here, that cultural identity could be understood through fixed boundaries while in reality, these boundaries began to dissolve and give way to ever floating constructs of cultural identity in the space of an urban metropolis:

The "exotic" is uncannily close. (...) An older topography and experience of travel is exploded. (...) Difference is encountered in the adjoining neighborhood, the familiar turns up at the ends of the earth. (...) "Cultural" difference is no longer a stable, exotic otherness; self-other relations are matters of power and rhetoric rather than of essence. A whole structure of expectations about authenticity in culture and in art is thrown in doubt. (Clifford, 1988, pp. 13-14)

Given the complicated and often hidden interactions of culture, native language, and assessment that I already discussed above and in the last subchapter, no essential definition of culture can inform this study. Any essential power discourse cannot suffice to explain the complicated entanglements of exams which are taken by internationals from around the world.

The study of globalization is an ongoing endeavor in the social sciences until today and there are several perspectives on how the global and the local interact in urban spaces. Two prominent authors in the field are Huntington (2000) and Barber (2014). "The planet is falling precipitantly apart AND (*sic*) coming reluctantly together at the very same moment" (Barber, 2014, p. 32). Barber describes in his article *Jihad vs. McWorld* a world in which the fabric of democracy is constantly torn between ideologically contrasting motions. He uses the terms "Jihad" and the "Lebanonization of the World" (p. 36) to describe the battle of local forces which seek fragmentation and follow tribalistic logics, against a world spanning homogenization of economics, culture, and societies. While these terms, and here particularly the term *Jihad*, relate to certain mindsets, rather than easily observable and measurable physical phenomena, the

author identifies their entanglements as reaching far into the political realities of global and local phenomena (p. 39). Above discussed intersections between collective high stakes exams and migration patterns offer a great example for his notions. Democracy is challenged by “meritocratic elitism” on one side and the local cultural resistance against the former, by “retribalization” on the other side of the spectrum. It is here fitting, that a critical research of high stakes exams, as already established, needs to go beyond the direct manifestations of the products, the exams, as their *institutional capital* reaches far beyond the direct implications of the exam’s execution worldwide. On the several accounts presented in the first section of this chapter, exams, especially if administered on a global scale, serve as gatekeepers to social and economic achievement, either in the society of origin of test takers or in reference to a society, into which test takers want to migrate.

Baker and LeTendre (2005) examine the interactions between global culture and national schooling in their study *National differences, global similarities: world culture and the future of schooling*. Their contribution is ultimately a critical analysis of education narratives. They claim that the narrative of “an educated citizen” as a “national product” to carry forth the nation as an “entity” (p. 1) is fundamentally flawed, because curricula all over the world today, are pervaded and shaped by “global ideas”: “The current situation in schooling across nations is wholly unpredictable from the image of unique national models of schooling” (p. 2). The authors believe in identifying general patterns of education in every classroom around the world, irrespective of the local culture, and further support their argument with the success of “state sponsored mass-schooling” as a specific, an industrial form of education export. The recipients of this mass-schooling model are public schools all over the globe.

Ultimately, Baker and LeTendre (2005) fail to address the actual problem which underlies their discoveries, namely the consequence that exporting Western education as “a pervasive cultural product (some would say even a hegemonic product)” (p. 12) has for the diversity of world views, cultures, and traditions of increasingly diverse classrooms. The conflicts between the local and the global in the sphere of education, the spread of global ideas and the role that high stakes exams play in here requires another theoretical lens, one that comes through a differentiated understanding of diversity, which can be applied to classrooms, but also one which ultimately could expand Bourdieu’s cultural capital and generate new categories of analysis.

Clifford (1988) gave an account of a demise of stable cultural otherness, one which needs to be taken into consideration for a framework based on the idea of cultural capital. Where “essence” is removed, there is a demand for more diverse explanations of culture. The term diversity can be considered as one of the least coherently defined terms in the English language, especially when it is applied in the context of human interactions. Looking at definitions underlying diversity, there are a myriad of discrepancies between different schools within the field of *diversity studies*. As the British researcher Vertovec (2015) described it:

Within businesses and corporations, state agencies, universities and a variety of public institutions, diversity is an essential reference point within mission statements, strategy frameworks and staffing structures. (...) Consequently, real challenges arise when academics attempt to probe such a term critically and analytically while public bodies are using it normatively or instrumentally. Confusion and miscommunication are bound to arise. (p. 1)

Vertovec (2015) expressed a dire need for different academic paradigm shifts regarding the topic following this quote. A drastic mismatch between the actual deployment of terms related to diversity in social spheres and the academic research. While the former draws from context, the latter struggles to operate with content. As a consequence of the contextual deployment of

*diversity*, Vertovec (2015) observed that there are nearly as many underlying definitions of diversity in discourses, as there are social contexts of its application: “In public discourse and policy, then, diversity has no clear content or overall aim” (p. 3). His theoretical assumptions are based on a UNESCO and University of Oxford co-founded research project into better understanding the demographics and regulatory demands of London’s diverse population.

Thinking about the inclusion of categories that describe such a diverse population, Vertovec (2015) devised a list of categories which typically are used under the umbrella of “diversity” in many examples of the public and corporate sector:

(...) understandings of diversity include categories of difference such as: race, gender, ethnicity, culture, social class, religious belief, sexual orientation, mental ability, physical ability, psychological ability, veteran or military status, marital status, place of residence, nationality, perspectives, insights, background, experience, age, education level, cultural and personal perspectives, viewpoints and opinions. (p. 2)

Following the exploration of these statistical categories, he (2007) devises the idea of super diversity, which positions itself in stark contrast to the uniformity of cultural essence, that Clifford (1988) already condemned to failure:

(...) there is deep suspicion about mapping cultures onto places, because multiple cultures and identities inevitably inhabit a single place and a single cultural identity is often situated in multiple, interconnected spaces.” [and giving the example of London:] “A city or neighborhood may have small numbers of new migrants but relatively high indices of diversity. (p. 1042)

Spaces become interconnected, as cultural identities spread over all delineations.

Relating this realization back to the task of comparing exam content across different languages, the assumption presents that such is possible if the threads that structure their diverse networks of production can be revealed. The mention of migration in the above quote is an example for one of such traditional indices which are used to demarcate a specific space from another. But spatial demarcations no longer suffice to grasp the reality of cultural production. What does that mean? Super diversity dissolves the idea of space rather than using it as an

instrument for comparison. In its stead, a multitude (or multiplication) of significant variables steps in and forms something Vertovec (2007) pictures as *multifaceted networks*, opposed to stable entities:

(...) the existence of social dynamics (is) marked by ‘an articulate core and a disarticulated plurality of peripheries’ differentiated by variable conglomerations of race, immigration status, gender, economic activity and neighbourhood. (p. 1044)

In a space where several variables interact to form conglomerations, the category of space itself dissolves and is replaced by a plurality of features which correlate among one another to form these multifaceted networks. Such correlations are the key to discovering new instruments and one of such could be formed through an understanding of *cultural capital* as will be the concern in the following chapter. The high stakes exams under observation, when looked at through the lens of super diversity, become vehicles of thought and tools of globalization at the same time.

With the help of Vygotsky’s and Bourdieu’s theoretical backdrop, I will let myself be inspired by ideas of multifaceted networks when using network text analysis.

### **Beneath the Educational Enterprise: Social and Cultural Capital**

The first dimension of the Müller and Schmidt (2016) taxonomy was “dominion” in the form of social assimilation. The discussion of cultural bias in testing provides an important contribution to better understanding how assessment is related to social hierarchies within societies, but how about cultural dominion? Bourdieu (1977) in his much-accomplished work *A theory of practice*, provided a critical theory about the culturally structured order of everyday life.

When initially researching about high stakes examinations, I perceived a lack of critical analysis of the examinations currently employed and identified one way to address this lack through a critical realist lens, which has social equity and the formation of human beings in society as a focus. (Blaikie and Priest, 2017, pp. 171-173) In *A Theory of Practice*, Bourdieu

developed different frameworks that can be used to shed light on the relationship between the living individual, member of a society, learner, test taker and the encultured and socially encoded object that is a high stakes exam. His theoretical insight of *cultural capital* might provide a framework through which I could access informal education.

Cultural capital can exist in three forms: in the *embodied* state, i.e., in the form of long-lasting dispositions of the mind and body; in the *objectified* state, in the form of cultural goods (pictures, books, dictionaries, instruments, machines, etc.), which are the trace or realization of theories or critiques of these theories, problematics, etc.; and in the *institutionalized* state, a form of objectification which must be set apart because, as will be seen in the case of educational qualifications, it confers entirely original properties on the cultural capital which it is presumed to guarantee. (Bourdieu, 1986, p. 247)

Here Bourdieu's (1986) theory touches two crucial dimensions that need to be explored if I could assume that the dead objects, which are exams carry in themselves a sociocultural encoding that can serve as artefact of learning and/or development. The objectified state of the cultural capital encoded in high stake exams could hence be determined through a proper analysis of their content in relation with social discourse.

Cultural capital, in the objectified state, has a number of properties which are defined only in the relationship with cultural capital in its embodied form. The cultural capital objectified in material objects and media, such as writings, paintings, monuments, instruments, etc., is transmissible in its materiality. A collection of paintings, for example, can be transmitted as well as economic capital (if not better, because the capital transfer is more disguised). But what is transmissible is legal ownership and not (or not necessarily) what constitutes the precondition for specific appropriation, namely, the possession of the means of 'consuming' a painting or using a machine, which, being nothing other than embodied capital, are subject to the same laws of transmission. (Bourdieu, 1986, p. 250)

Cultural goods must be appropriated in their material, as well as in their symbolic nature (economic and cultural capital of objects). The symbolic nature of the material can be understood through the embodiment of the knowledge of its creation, the science, math, or the ideological understanding, but also through proxy. This, according to Bourdieu (1986), generates the basis



for executives and engineers, who have no possession of the material, but embodied cultural capital and its temporary use thereof (pp. 250-253).

The same principles could be applied to an assumed perfect achiever in a high-stakes exam. How would they need to act, which embodied cultural capital would they need to inform their content knowledge? Academic credentials, for example, neutralize to some extent and for a certain period of time the biological limitations of their bearers. As opposed to that, autodidacts can be called into question of their skill at any time. Further, institutionalized cultural capital can be used to shed light on another eminent property of high stakes exams, namely on the actual *stakes* that they bear for their audiences and the trust that is confined in them through the institutions and societies which endorse them. Here, a potential analysis needs to depart from the content of the actual exams and go beyond a narrative literature review to the conception of these exams.

The origin of Bourdieu's (1977) theory supports that assumption, as it sprang from the careful ethnographic observation and description of social, cultural, political, and economic transactions with a group of Berber people whom he observed as a participant over long periods of time. Reapplied from his original *field* unto so called Western societies, the identification of social classes led to the formulation of cultural capital when Bourdieu (1986) tried to explain differences in scholastic achievement among children originating from different classes (pp. 250-253). Embodied cultural capital is further described as an accumulative process departing from a substantial understanding of the human individual, a "singularity" (p. 49). Education allows the human individual to accumulate culture through familial and institutional education. Embodied cultural capital further breaks with certain distinctions of inheritance and acquisition of capital,

as it represents acquired inherent properties: „It thus manages to combine the prestige of innate property with the merits of acquisition” (Bourdieu, 1986, p. 250).

Bourdieu (1986) further identified two other determining factors to the realization of embodied capital, which are time and scarcity (p. 250). The former relates to one of the reasons working class children are more likely to maintain the status of their parents as the time invested in their education often does not exceed the minimum necessary to work in the labor market. In fact, he moves in the spheres of economy and allusions to such abound. *Institutionalized cultural capital* becomes an economic factor as it links embodied capital and the individual to the mechanisms of free market:

By conferring institutional recognition on the cultural capital possessed by any given agent, the academic qualification also makes it possible to compare qualification holders and even to exchange them. ... Furthermore it makes it possible to establish conversion rates between cultural capital and economic capital by guaranteeing the monetary value of a given academic capital. (Bourdieu, 1986, p. 51)

On Bourdieu’s theories in application to high-stakes or university entrance exams, research appears to be sparse. Hong and Youngs (2008) serve as the only example of a research directly linking the concept of institutionalized cultural capital with high stakes exams. The authors assume institutional cultural capital as an actually measurable variable and draw conclusions about the effect of high-stakes testing on ethnical minority groups and low-income social groups in the US population. Their analysis follows an empirical paradigm which I would like to abandon through my own application of Bourdieu’s cultural capital.

Though not related to high stakes testing, a research conducted by Erel (2010) caught my attention. He researches into the cultural capital of migrants within the British and German society. As my problem statement was informed by the absence of meaningful research into the relationship between foreign test takers and university entrance exams, so would also a content analysis of high stakes exams need to be resituated back into discourses of migration and

diversity. At this point, Erel's (2010) research proves one thing, i.e. a theory of cultural capital is not sufficient if there is a theory of culture or even diversity lacking to inform this inquiry. With the words of Erel (2010) the problem is as follows:

(...) cultural practices within a migrant group are differentially validated in gendered, classed and ethnic ways. Cultural practices acquire different meanings and validations according to the local, national and transnational context. Rucksack approaches do not adequately take account of the struggles over particular cultural practices and the differential ways in which cultural practices articulate with forms of femininity, ethnicity and class to create complex hierarchies of distinction. In order to understand how cultural capital signifies distinction and produces recognizable social identities and positionalities it is important to consider the meanings the actors give to cultural practices. (Erel, 2010, p. 656)

The *Rucksack approaches*, to which Erel refers here are his idea of a direct application of the framework cultural capital as a static one-sided reading of culture and socio-economic differences. For my research, I will instead use the knowledge of cultural capital to inform and refine already existing frameworks, like those postulated through the pedagogy of testing discourse.

As initially mentioned in Chapter 1, I am using two methods which rely on text as data for an analysis. Exams have textual components (e.g. the question sheets) but given that any exam's outcome is not determined by the question, but by the corresponding responses, I am faced with a significant obstacle, which I first need to overcome conceptually here before I can later tackle it through methodological adaption. I choose to complement the assumption of cultural and social capital expressed above with a theory of perception, formulated by Vygotsky (Van der Veer and Valsiner, 1994; Vygotsky et al., 1978, Vygotsky, 1980). In the context of research into educational development in children, Vygotsky (1978) develops a theory of *artefact mediated cognition*, where a degree of socially and culturally formed perception determines the degree of developmental acquisition of *tools*, with which children and in general, learners, can influence and participate in the world around them. The concept of tool that

Vygotsky (1978) employs here has its origin in material dialectics and Engel's (2019) idea of labor and tool use transforming nature and human nature in the process (p. 7). In Marxist material dialectics the tool is the means through which we can accumulate production and therefore economic capital: "The specialization of the hand-this implies the tool, and the tool implies specific human activity, the transforming reaction of man on nature" (Vygotsky et al., 1978, p. 8). In allegory to the tool being used to transform human nature through labor, Vygotsky established language, symbols, and signs as the tools with which we mediate our social relationships, and with which we transform our nature in allegory of Engel's (2019) dialectic method of capitalistic transformation (see also Wygotski, 1930, p. 309). In Vygotsky's (1978) words "(...) the basic analogy between sign and tool rests on the mediating function that characterizes each of them" (p. 54). The transformative aspect, which carries on from the above dialectic method, hints towards a two-sided form of action of the psychological tools available to us. "The mastering of nature and the mastering of behavior are mutually linked ... We call the internal reconstruction of an external operation *internalization*" (Vygotsky et al., 1978, pp. 55-56).

For the response to research question 1, which asks about *competencies*, I will lay the here presented as a foundation for an altered definition of the term. The *mediated artefact* is represented by the non-expressed, officially endorsed, and considered correct, response to an exam question. Hence though not realized in the question, the response turns into an indispensable part of the exam data and of my analysis of such.

## Competences as Social and Cultural Concepts

An important element of the research question are *Social and Cultural Competencies*, as connected to the ideal score. In light of the discussion so far, it is important at this point to better frame these terms. An essential definition will be obtainable after the analysis of the results of my research. In a short mention during the introduction above, I talked about Tinnefeld's (2013) treaty for the creation of a subdiscipline of testing. Within his chapter 2.1 where he attempted to define the field, he took note of a lack of wholistic research on assessment, one combining different disciplines. He noted that there is a lot of compartmentalized research, but too little has been done to locate assessment within a multidisciplinary framework (pp. 81-84). He calls to bring forward the societal structures that organize and give meaning to assessment, juxtapose the assessed with the role of the tester (p. 83).

In general, he called for an approach to assessment, which has long since been deemed necessary for the larger subject of education, for example through *Curriculum Theory* (Huber 1981; Pinar 1989, 2004, 2006; Lincoln 1992). During the introduction and the first chapter of *What is Curriculum Theory*, Pinar (2006) reflected on several changes affecting education in schools and in academia, during the second part of the last century, such as several waves of “reconceptualization” of the curriculum (p. 39; p. 155). Pinar (1989) stated:

The curriculum (...) is a holistic life experience, the journey of becoming a self-aware subject capable of shaping his or her life path. As a perpetual struggle, the curriculum ... is never a finished product that can be finally mastered and passed along to an awaiting new generation. (p. 130)

Of special importance to this text is what Pinar (2006) later termed the “out-of-school experience” (p. 27). He mentioned the authors Aoki (1988) and Van Manen (2006), where the former created guidelines for phenomenological research and together, they laid the groundwork for phenomenological curriculum research at the University of Alberta (pp. 408-409). In its

method, such research looks at “(...) densely textured moments which point beyond the immediacy of the context in which they occur” (Pinar, 2006, p. 407) and then aims to create a form of “wide-awakeness,” which “(...) attunes and pulls the investigator and the student more deeply into the world” (ibid, p. 408). Through the assumption that curriculum is inherently linked to the process, to social interactions between the individual and institutions, Pinar (2004) continues to illustrate the assumption of hidden curricula.

Hidden curricula are defined by “(...) all of the experiences planned and unplanned that occur under the auspices of the school. (...) The hidden curriculum is the ideological and subliminal message presented within the overt curriculum” (Pinar, 2004, p. 27). He described it as a social construction which governs our perception of history, and therefore of the present. As an example, he used the construction of history around memory and forgetting on examples of accounts of the First Nations. Such untold narratives of the vanquished First Nations eventually justified a physical and ideological suppression, which enabled the emergence of the American nation state. To expose such devolving structures should be the goal for an inquiry into *hidden curriculum* (Pinar, 2004, pp. 39–40). For this research, the Hidden Curriculum serves as an inspiration to deconstruct exams in terms of their content, context, in terms of what is said about them, and in terms of what is hidden beneath structures of authority and regulatory axioms. Regulatory axioms determine in what manner, shape and form each student can take an exam. All exams, and especially the Abitur, Baccalauréat, and SAT, follow their own canon of rules pertaining to, for example, response format, subject combinations, intensity of content being tested etc. As the network text analysis reveals, these structures can be understood as semantic networks, connecting the exam with social, cultural concepts of assessment and spheres beyond education. I will continue this thought in Chapter 3, with the methodological discussion of

network text analysis. By including concepts from Said (2003), Pinar (2004) theorizes the dissolution of the human subject, the dissolution of text as linear knowledge, and a replacement of these through decentered networks of knowledge grouped around “nomadic centers” and “provisional structures” (p. 139).

I have already assumed the existence of a form of these provisional structures, when I described artefact mediated cognition in the section above. In order to better link a research of assessment with a similarly wholistic stance as had been taken by the curriculum reconceptualists above, I continue orienting myself Müller and Schmidt’s (2009) functional framework, which was mentioned in the introduction. With the aim to understand and reveal the different stakes of assessment, Müller and Schmidt (2009) created the following table based on what they call “*die wohl differenzierteste Systematisierung von Prüfungsfunktionen.*” (p. 25) [The most differentiated categorization of assessment], in direct reference to a 1970s German education theoretical publication by Fuchsig (1976):

Table 2.1. Functional Domains of Assessment Translated.

<b>Function: Dominion and social assimilation (reproduction of the system)</b>
initiation
distribution of Status
legitimization
<b>Function: Recruitment (selection, election, allocation, standardization)</b>
cohort placement
selection within cohort
proof of qualification and competence
<b>Function: Didactics</b>
Test as timer and content structurer in a course
provision of orientation with regard to learning goals for lecturers and students
extrinsic motivation to learn
diagnostic tool
auditing over teaching and learning success of teachers and students

As previously outlined, socioeconomic status is directly connected to successful exam results in many contexts and societies, including the ones from which my research data would stem. The studies to which I refer in the context of this review, proved that the highest mark will probably also involve a degree of social and cultural assimilation. Table 2.1 posits such influences into a functional descriptive framework involving *initiation*, *distribution of status* and *legitimization*. Exams tended to often be nothing more than obstacles to overcome, but if we take a look at the above presented taxonomy, then this is not a flaw, rather a feature.

Müller and Schmidt's (2009) second function identified for assessment is *Recruitment*. This category focuses on the property of exams to positively or negatively influence career opportunities of individuals who stand in direct competition for the same career paths with other individuals (p. 26). That category theorizes properties of standardized assessments, as all previously mentioned exams to serve as tools of selection and placement. But another mechanism that falls into this category is the decision-making process for the distribution of funding within the US that I discussed in section 2 of this chapter through the example of an El Paso, Texas school district.

While the third subcategory *legitimization* might apply to schools struggling to legitimize their existence within an educational system founded on rigorous accountability, like the one that made the El Paso incident possible, it might also relate to the legitimization of degrees. Such links widely to the discussion of cultural capital, specifically in its objectified and institutionalized form. Bourdieu's (1977) concept of cultural capital becomes very important here as it can contribute to also better refining the category *Dominion and social assimilation*. Müller and Schmidt (2009) elaborated on the initiation and legitimization aspects of said



category, in that both would serve mainly for reproduction and constant reaffirmation of certain social status quo through the society (p. 26-27). However Bourdieu (1977) goes much deeper:

Cultural goods must be appropriated in their material, as well as in their symbolic nature (economic and cultural capital of objects). For the example of NGOs that want to deliver so called developmental aid to so called third world countries, need to foster cultural capital as well as economic remittance. The symbolic nature of the material can be understood through the embodiment of the knowledge of its creation, the science, math or the ideological understanding, but also through proxy. This generates according to Bourdieu the basis for executives and engineers, who have no possession of the material but embodied cultural capital and temporary use thereof. (p. 250)

Academic credentials neutralize to some extend and for a certain period of time the biological limitations of their bearers. Autodidacts can be called into question of their skill at any time.

What does that tell about us who have tried for years to climb to the highest academic ladder available (pp. 250-251). Institutionalized cultural capital works similarly, but instead of focusing on the individual, it is defined through the mechanics of the free market approach:

By conferring institutional recognition on the cultural capital possessed by any given agent, the academic qualification also makes it possible to compare qualification holders and even to exchange them. ... Furthermore, it makes it possible to establish conversion rates between cultural capital and economic capital by guaranteeing the monetary value of a given academic capital. (Bourdieu, 1977, p. 51)

Such a shift of focus from the individual towards groups, towards the workings of institutionalized cultural capital within exam environments is necessary to fill above categories beyond what the authors provided.

The last function of the taxonomy above, *Didactics*, is concerned with the exam situation as an educational, rather than going with the surrounding contexts. Müller and Schmidt (2009) explain the educational situation through a functional analysis of the teacher, as well as through a description of the students' roles (pp. 27 – 29). They conclude that a proper analysis of the function within each category might have the following effects:

*Auch auf systemischer Ebene könnte im Anschluss an Prüfungen ein Ausbau von Betreuungsangeboten für Studierende ins Auge gefasst werden. Zusätzlich wären auch*

*inhaltliche oder strukturelle Modifikationen bestehender Veranstaltungsreihen oder des Modulaufbaus denkbar, wenn etwa festgestellt wird, dass in einem späteren Semester Wissens- oder Kompetenzlücken in bestimmten Domänen bestehen.* (Müller and Schmidt, 2009, p. 28)

Judging from a systemic stance, the lessons learned from the act of assessment might lead to adaptation of tutorship and support for future test takers. In addition, exams might reveal shortcomings of content and structure of already existing curricula and timetables. Such could be deduced from domain specific analysis of the exam performance and lead to adaptations of already existing lectures, lessons in a later semester.

Though the authors remained less specific about the domains to which they referred, this provides an opportunity to adapt the above presented ideas and formulate my own theoretical basis for learning and developmental ideals, which in turn might lead to a structured analysis for my methodology chapter. Vygotsky's (1978) theories of learning and development could provide valuable theoretical foundation here.

Vygotsky (1978) provides some insights into the processes of learning and development, and developed *sign*, *artefact* and *semiotic mediation* as primary venues of cultural and social development and learning. He articulated these concepts in *Mind in Society*, where, through the words of Koffka, he states the following about development: "(...) development is based on two inherently different but related processes, each of which influences the other. On the one hand is maturation, which depends directly on the development of the nervous system; on the other hand, is learning, which itself is also a developmental process" (p. 81). He established learning as contextual, and as social, and describes signs as tools for the mediation between the learner and society.

From the material dialectics perspective, Vygotsky (1978) borrowed several formative elements of his concepts on psychological tools. In an allegory to the tool being used to transform human nature through labor, Vygotsky established language, symbols, and signs as the tools with which we mediate our social relationships, and through which we transform our nature

in allegory of Engel's (2016) dialectic method of capitalistic transformation. The transformative aspect, which carries on from the above dialectic method, hints towards a two-sided form of action of the psychological tools available to us. "The mastering of nature and the mastering of behavior are mutually linked (...) We call the internal reconstruction of an external operation *internalization*" (Vygotsky, 1978, pp. 55-56). One of the aims of my envisioned wholistic analysis of high stakes exams is to discover and reconstruct the knowledge that encoded in these exams as archetypes of learning. While earlier I posited the idea of unrealized responses to exam questions as inherently necessary for an empirical analysis of exam data, archetypes of learning might be structured by the relationship between sign and artefact as defined here. In the context of my research, *Artefact* describes textual artefacts of learning and development encoded in the content of, and necessary for the success in high stakes exams. However, an application of sign in this context might prove more challenging.

Following the assumption of constructedness, even of social and cultural interactions, as well as their symbols, Vygotsky (1978) further explains thereinto: "To study something historically means to study it in the process of change; that is the dialectical method's basic demand" (pp. 64-65); and further Vygotsky mentioned that "Behavior can only be understood as the history of behavior" (p. 65). In Marxian material dialectics: "It is not the individual perception that limits the individual, it is how they are socially perceived which ultimately limits their action." (Marx, 1975, p. 9) (*Es ist nicht das Bewusstsein der Menschen, dass ihr Sein, sondern umgekehrt ihr gesellschaftliches Sein, das ihr Bewusstsein bestimmt*).

A similar dialectic nature is displayed in sign systems, which are produced by collectives of individuals or societies, but ultimately are internalized in the development of the individual through their repeated application. As such, the dialectic method is to be understood as an

attempt to enable a proper description and analysis of the fleeting and ever-changing process of learning and development. It also opens up new horizons of analysis in conjunction with the envisioned research and the preceding historical, narrative literature review into the several university entrance exams, understanding the content of these exams as not just constructed, but also positioned in time within a social and cultural discourse that deployed the signs used to generate the visible artefacts therein.

Just as in Marxism, the social cell permits an observer to read in it the fabric of the society. For Vygotsky, it is the single psychological cell that gives out the anatomy of psychology. The dialectic method discussed earlier is one specific example of an expression of these elements within his corpus. Adding to that, Vygotsky's (1978) particular historical, sociocultural approach is further summarized as follows:

In the development of higher functions – that is, in the internalization of the processes of knowing – the particulars of human social existence are reflected in human cognition: an individual has the capacity to externalize and share with other members of her social group her understanding of their shared experience. (p. 132)

In another source, the capacity to internalize social existence is further described as a two-way process along a literal interpretation of dialectics. While the social system exists on the outside of the individual's psyche, the latter shapes around the structures given by the former, while permitting to add their personal expression onto it through a process which could also be named appropriative (see Wygotski, 1980).

Vygotsky's (1978) dialectic approach was taken up and developed into studies of different nature. For example, the field of second language acquisition knows many applications of cultural mediation as framed above. Wertsch (1998) used this framework to locate cultural mediation within the interaction of child and adult, hence positing the adult as the conveyor of sign and as the enabling element to the child's language acquisition and the internalization of

sign, which results from such processes. Staying within the field of second language acquisition and early childhood development, James Lantolf (Lantolf, 2000; Lantolf and Thorne, 2006; Lantolf and Beckett, 2009) developed sociocultural theories of language acquisition on the foundation of Vygotskian dialectics. An example for an application of cultural mediation in a different field could come from Hennig and Kirova (2012) who identified toys and play as artefacts in the intercultural learning processes of young children. The application of Vygotsky's theory in my research, though, marks a major departure from its common application in the field of educational intersubjective practice and towards an application with the analysis and ordering of content in forms of educational archetypes. For that purpose, signs and artifacts will play an elevated role, which I employ on the foundation of Vygotskian theory. By taking away the intersubjectivity, the explanation of culture and signs will lack an important dimension of analysis, which requires to be replaced by another theory to support such analysis.

This chapter made clear that on one hand, wholistic approaches to assessment are rare as research is often compartmentalized, but they are necessary, because assessment, as an institution of education, does have far reaching socioeconomic implications for the lives of exam takers of certain types of exams. I assume in the following that the stakes involved with high-stakes testing are neither revealed through any existing curricular description of exams, nor through empirical measurements of exam outcomes. A curriculum theory approach to assessment with its interdisciplinary requirements comes with several implications, of which I discussed the most important ones pertaining to my research in this chapter as well. Exams as material objects, or texts have several hidden layers of cultural and social significance, which are encoded within controlled responses and within language used to describe these exams. A comparison across three languages holds value in itself, as socio-cultural determinism is often

expressed through the language we use. In order to respond to the research questions posed in chapter 1, I use theoretical frameworks developed by Bourdieu and Vygotsky, while using some assumptions made in the context of a German discourse on pedagogy of testing to operationalize these frameworks and adapt the methods I use here for the application to my data sources.

## Chapter Three. Methods, Methodology, Sampling, Data and Data analysis

The following chapter focuses on the methods used in their application, but also contains a necessary reflection of methodological change, which each method needs to undergo. The methodological change is necessary, as I use the methods in a unique setting.

### Methodological Considerations

Though the method section in a dissertation and/or many other academic publications is often an axiomatic instance, often not justified, it is just there. Though the here presented research does not allow me such oversight. Due to the high complexity, it is imperative to reflect the tools, which I am employing out of the rich *academic toolbox*. That notion of an academic toolbox was not invented by me. I rather am inspired in my thinking by Fairclough (2001) who used this term while attempting to describe a methodological application of the then much discussed *critical discourse analysis*:

I should declare at once that I have certain reservations about the concept of ‘method’. It can too easily be taken as a sort of ‘transferable skill’ if one understands a ‘method’ to be a technique, a tool in a box of tools, which can be resorted to when needed and then returned to the box. CDA is in my view as much theory as method – or rather, a theoretical perspective on language and more generally semiosis (including ‘visual language’, ‘body language’, and so on) as one element or ‘moment’ of the material social process (Williams, 1977), which gives rise to ways of analyzing language or semiosis within broader analyses of the social process (...). (p. 121)

I do share his concerns about the unreflected separation of method and theory, which informs this subchapter. I believe that, without a methodological discussion, there cannot be any application of the following methods, especially as I am use each method in a context it has never been used before. The following methodological discussion will serve to dissect each method and will help me identify the prospect of each of their application within my research paradigms, as opposed to just using each as a “tool in a box of tools” (Fairclough, 2001, p. 121).

The first method on which I need to reflect is *qualitative meta-synthesis* in some of its different forms. The primary goal is to produce knowledge as rich as possible on the purposes of assessment. Though I intend to be guided by Bennett, Driver and Trent's (2017) narrative literature review approach, I also intend to alter it, giving it a quasi-historical dimension.

Once turned capital, culture turns into a currency, in a capital market paradigm, such can be exchanged for status and a myriad of social resources. But once cultural capital is either objectified or institutionalized, it inevitably turns finite, difficult to access. The premise of Bourdieu's (1977) original work was provided through the challenges he faced as an ethnographer understanding kinship functions within a group of Berber people.

Reading about approaches to observation led me to the idea of retracing the axiomatic properties we can observe today when assessment is concerned, surrounding all that is observed and true in relation with the exam data from the BAC, Abitur and SAT. Another support for such an approach is coming from Bhabha (1997), who discussed the locatedness of culture in his treaty on postcolonialism, *The location of culture*. The idea here is to first situate university entrance exams within their specific cultural contexts, which will lead towards a framework to later identify learning and assessment situated in diverse networks of "difference" and "sameness" (Bhabha, 1997). A revelation of the concrete research that led to the inception of high stakes exams, might provide a deeper reflection on the institutionalized capital surrounding these high stakes exams today. The *qualitative meta-synthesis* results in a description of scholarly culture, which is shared among the stakeholders, executives and eventually all exam takers with a perfect grade. The results will also help in providing flesh to the theoretical construct of scholarly culture, which I first employed in the previous chapter.



Following the theory behind cultural capital I elaborated on in Chapter 2 section 3, Bourdieu (1986) uncovers the existence of such in 3 states, *embodied*, *objectified* and *institutionalized*. As a reminder, he theorized that cultural goods must be appropriated in their material, as well as in their symbolic nature (*economic* and *cultural capital* of objects). The symbolic nature of the material can be understood through the embodiment of the knowledge of its creation, the science, math or the ideological understanding, but also through proxy. According to Bourdieu, this generates the basis for executives and engineers, who have no possession of the material but embodied cultural capital and temporary use thereof (Bourdieu, 1986, pp. 250-253). In adapting the method of narrative synthesis, I draw from the implications above posited for my research, namely defining *scholarly culture* through assortments of capital shared by the creators and decisionmakers of the Abitur, Baccalauréat, and SAT.

Doing a narrative literature review, I am borrowing methodology from Bennett, Driver and Trent's (2017) theoretical approach to the method of narrative literature reviews:

(...) a narrative literature review refers to a systematic review that synthesizes the individual studies (...) systematically extracting, checking, and narratively summarizing information on their methods and results (...) Narrative reviews are beneficial in providing conclusions for researchers who examine topics that do not have one optimal way of measuring outcomes. (p. 896)

Similar to Bennett, Driver and Trent (2017), the narrative literature review will allow me to organize the, probably very different, research conducted for the conception of these modern university entrance exams in different countries and languages along certain emerging categories. Eventually, topics and themes will arise in terms of test objectives and findings. I will use this emergent approach to enrich my following frameworks of analysis. Another argument for the use of a narrative *meta-analysis* as Petticrew and Roberts (2008) called it, would be the hope that it helps me to go deeper and deconstruct certainties in the perception of high stakes exams. Similar to how Bourdieu, after careful theory backed ethnographical inquiry, broke apart

and narratively refigured kinship into the different forms of cultural capital. The authors write: “Systematic reviews also flag up areas where spurious certainty abounds. These are areas where we think we know more than we do, but where in reality there is little convincing evidence to support our beliefs” (p. 2). When responding to research question 1, the certainty about the Abitur, Baccalauréat and SAT lies not so much in the abundant reflection of their role as social and cultural institutions, but rather it stems from a long operational tradition of *doing things that way*. Especially with regard to uncovering social and cultural competencies, the meta-synthesis will prove useful in discovering certainties surrounding these exams, which need to be reconstructed.

Patticrew and Roberts (2008) discussed different terminologies used to relate to similar approaches which ultimately can be subsumed under the umbrella of a narrative literature review as Bennett et al. (2017) point out. Examples used in literature for a resembling set of methodologies are *systematic reviews in the social sciences* (Jones, 2004; Baumeister 2003) and *qualitative metasynthesis*, very prominent in the field of medicine (Sandelowski et. al., 1997). Patticrew and Roberts (2008) discussed systematic reviews in the social sciences, which have more relevance to my research and provide the following guidelines for the processes of inclusion and exclusion of certain studies (sampling):

If one wants to know “what works,” the review may include randomized controlled trials (RCTs), if there are any. If one wants to investigate “what matters,” qualitative data are going to be important. Many reviews do not have to address either one or other of these types of questions alone. (p. 57)

While sampling for the purpose of a narrative literature review will always include studies conducted by other researchers, a first distinction is considered above between *what matters* and *what works*. Judging by the nature of my research as guided by my postulated research questions, I will have to know which research stood at the foundation of the selected exam samples, at the

time of their conception and selection of content, which will help reveal the locatedness of the three exams within their societal and cultural foundations.

Patticrew and Roberts (2008) call this procedure of decision making “typologies of evidence” (p. 59). Generating a typology for the selection or non-selection of available evidence demands for knowledge of the study design. Following Patticrew and Roberts (2008), there are a few available, but the major criterion of distinction is the *degree of control* involved. Following the narrative literature review design as applied by Bennett et al. (2017), I will have no means of control for reliability, therefore sampling and data collection will have to rely on an open-ended exploration of connections, rather than on a measurement to the degree to which studies influenced the generation of high-stakes exams in their current forms. Bennett et al. (2017) did not discuss this issue, but presented a rationale based on their research question through which they conducted their review of literature (p. 896). In allusion to above, the idea is *what mattered*, therefore I will leave the definition of social and cultural competencies open and undefined, or rather to be answered and provided by the literature. Selected literature needs to postulate learning and developmental goals in any form and definition, that are either measured by the student’s success in these high-stakes exams or supposedly required to reach a perfect score in these exams. This statement will turn into my typology and eventually into the rationale for sampling, for the selection and non-selection of studies to review.

An important departure from the above approaches commonly termed narrative literature reviews, is that the nature of my text required me to deviate from some of the postulates presented by Bennett et al. (2017). All texts chosen are either qualitative in nature or used a wide range of mixed methods to qualitative methods. For that reason, I here focus completely on what Leamy et al. (2011) outlined as a *narrative synthesis*. The latter can be considered a rather recent

adaptation of qualitative meta-synthesis to work with qualitative research, as opposed to empirical, often quantitative one. Leamy et al. (2011) went even a step further and developed a framework to combine and cross-tabulate findings from empirical, as well as from non-empirical, theoretical, research. As the texts I will include for my meta-synthesis are all qualitative in nature, often do not adhere to modern scientific standards, but are nevertheless highly relevant for either one of the exams, I will continue referring to my method as *narrative synthesis*.

As another departure from Bennett et al.'s (2017) approach, I will establish a rationale for the inclusion of literature, which is a degree of relatedness. When I just mentioned *highly relevant* above, the rationales for inclusion I employ serve to establish this relevance. Though I continue being inspired by Bennett et al.'s (2017) concept of *degree of control* as introduced above, I cannot dogmatically follow their approach, but I need to reinterpret it somewhat. While Bennett et al. included only literature published in peer reviewed journals, presenting some type of firsthand research (as opposed to a theoretical discussion), they assessed the authority and validity of their data collection. As I focus on already established, widely employed exams, the validity of my data collection will not necessarily come from the data being scientific, but rather from the data being *endorsed*. I define endorsement of research, studies, and literature through its retraceable relation to the exams in question. This relation could be established through direct appearance in reference lists. More to that follows in the section on sampling in this chapter.

### **Network text analysis**

Network text analysis (NTA) is a method represented by Diesner and Carley (2005) in one of their more recent applications as following:

(...) in all cases, networks of relations among concepts are used to reveal the structure of the text, meaning, and the views of the authors. Further, these networks are windows into

the structure of the groups, organizations and societies discussed in these texts. This structure is implicit in the connections among people, groups, organizations, resources, knowledge tasks, events, and places. NTA is a specific text analysis method that encodes the links between words in a text and constructs a network of the linked words. The method is based on the assumption that language and knowledge can be modeled as networks of words and the relations between them. (p. 83)

Diesner and Carley's (2005) specific iteration of the method (short: NTA) works with 3 layers of complexity, the first one being a computer assisted quantitative linguistic analysis of the concepts represented in a set of texts and the frequency of their appearance within texts, which the authors selected according to their research design (see pp. 91-92, *Concept Identification*). "A concept is an ideational kernel – a single idea" (Carley, 1997, p. 81). In the earlier iteration of the methodology from which this quote is taken, Carley (1997) established network text analysis as a methodology focused on the content of texts, as determined by concepts.

Such concepts are emergent: "In cognitive networks (both at the individual and societal, or social-conceptual levels), concepts are devoid of meaning except as they relate to other concepts" (Carley, 1997, p. 81). Meaning is hence created by going beyond the inherent structure of text and language, through which the text is encoded. In order to distill a method from this, Carley worked with several assumptions, through which meaning emerges beyond the inherent structure of textual objects (p. 83). Important for the purpose of adapting her methodology and applying it to exams as objects are the following: Conceptual networks have no beginning and no end, therefore are radically relational; all data produced is relational, specifically in relation to a "focal concept" (pp. 84-85); such focal concepts are determined by the researcher and relate to the research interest of the project. The location of *focal concepts* can be measured through the dimensions "density, conductivity, and intensity" (p. 87), as well as "imageability" and "evokability" (p. 92). The following graphic stands as the final product of these measurements:

While density, conductivity and intensity are measuring the prominence of the focus term within the local network (the social discourse), the arrows indicate the direction and the strength of imageability and invokability of closely related terms. Arrows pointing in both directions are imagined and invoked properties of “hacker.” Though complex at first, this methodology has been adapted and continues to be adapted, for example through the current work of Heather Rackin at the LSU department of sociology (Rackin, 2020).

As posited later by Disner and Carley (2005), it allows to reveal the implicit structure of texts and within them “(...) connections among people, groups, organizations, resources, knowledge tasks, events, and places” (p. 83). After having revealed through the narrative literature review, what research is linked to the creation of my exam objects, this method will potentially produce a more informal dataset on the current *entanglements* (Shohany, 2016) of assessment and society. One aspect still requires specific attention.

In a figure on page 96 called *Local Network for “Hacker”*, Carley (1997) presented a local network of her focus concept “Hacker”. Such local networks are opposed to the larger but unrelated networks within language and the totality of its use. In order to establish a local network, for example for the purpose of a network text analysis of exams, it becomes necessary to limit the discourse and determine the focus. Such has to be done through a so-called *thesaurus*. A thesaurus helps frame the interpretation of text level concepts and in turn generates local networks.

Disner and Carley (2005) described *Thesaurus Creation* as follows: “(...) a generalization thesaurus translates text-level concepts into higher-level concepts. A single higher-level concept typically has multiple text-level entries associated with it in a thesaurus” (p. 92). At the same time, the authors have to admit that there is no predetermined formula available

yet to generate higher-level concepts from the in-text observations. In the following method sections and for my own research design, I will have to draw from my own theoretical lenses, outlined in Chapter 2. The methodological challenges which arise in applying network text analysis to exam data are created first by the fact that question sheets as exam sheets cannot be assumed to have the same properties as texts. As the sampling for this research reveals, exams are not purely text, there is little to be learned by measuring the word frequencies within and across the questions in exam sheets, as the heart of the exam lies in the response. I cannot look at relationships between lemmas, words, or phrases within an exam document. In order to overcome this problem, I will treat the exam text as both revealed and concealed.

The visible part consists of the questions, while the hidden or concealed part are the responses to these questions, as well as formal axioms. Formalities, which determine the nature of a test and an exam, arise from formal limitations of subject choice (Abitur, Baccalauréat and SAT do not have the same subjects to choose from), combinability of subjects, which eventually leads to each student having a different assortment of exam papers for their own individual *track* to completion of the respective exam. The idea to include concealed properties of exams aligns with the wholistic approach of my analysis as expressed in Chapter 2, for example through hidden curricula, through all presented intersections of exams with social and cultural determinants outside the act of assessment or even the educational enterprise. As I posited in my theoretical framework, these elements are not extrinsic to assessment as understood in this research, they are rather an integral part of assessment, beyond exam sheets and exam results. Despite these extraordinary circumstances, network text analysis presents an ideal candidate of a method, as the fundamental assumptions of the research conducted by Carley and Diesner (1997) and Popping (2000, 2003) naturally assume the existence of hidden structures, which need to be

recognized and analyzed. The minutia of how I adapted the method accordingly follow after a closer examination of the data with which both here presented methods will operate.

### ***Challenges of a Comparative Data Analysis***

Both methodologies presented here operate with separate samples that are conceptually linked through the concept that a deep analysis of assessment needs to go beyond the questions that make up exam sheets. The results of the narrative synthesis pertain to each exam individually at first and allow for a comparison in a subsequent analytical step. While the narrative literature review did not require the exam data, it was necessary to collect textual data linked to the three exams, but content-wise unrelated. The sampling has to happen in different stages, being tweaked and expanded for the application of each of the three methods. While doing so, translation and the question of comparability of foreign language content will pose another significant threat to the validity of the sampling. On the other hand, the second method, network text analysis, requires a cross-comparison between the exams to begin its analysis. Unlike established in Diesner (1997) and Popping (2000, 2003) the network text analysis operates with the exam content as sample and the analysis focuses on subject content, consisting of exam questions and responses, versus exam formalities. At the end of this section, the logics behind data collection and behind the resulting analysis deserve some more attention.

All of the presented methods need to include a procedure for the purpose of overcoming inherent linguistic bias through the creation of meta-structures for all content throughout all subjects of these exams. Unifying vocabulary will come from further definition and postulation of developmental and learning ideals, through the rich theoretical background provided by Bourdieu's (1977, 1986) cultural capital, through Vygotsky's (1978) artefact, and through the functional categories developed by Müller and Schmidt (2009), within the discussion of a



pedagogy of testing (Tinnefeld 2013). I tried to exemplify this logical challenge in Figure 1.1, which I presented in Chapter 1, but repeated here for a better visualization:

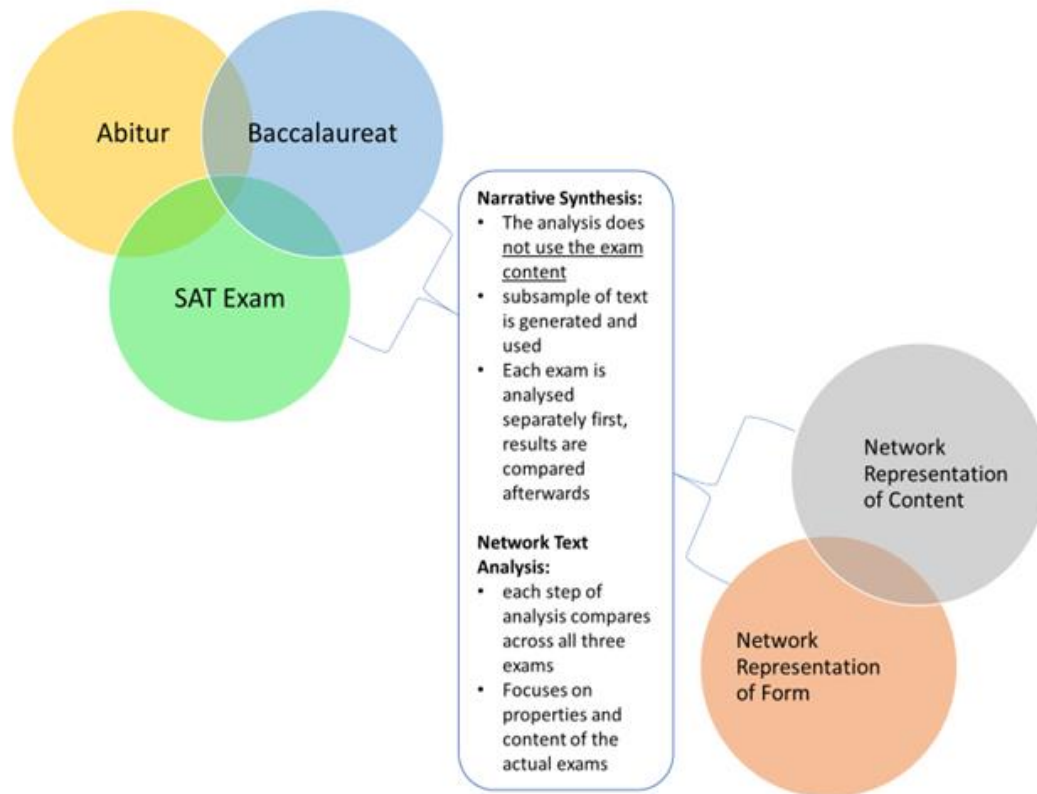


Figure 3.1. Visualization of Planned Inquiry and Data Analysis

The narrative synthesis will produce separate results depending on the French, German and English discourses and patterns revealed after coding. Before continuing with the network text analysis, the results of method 1 requires a preliminary comparison and reflection. The so generated results might, together with theoretical lenses from Chapter 2, even inform consecutive analytical steps, as each, to a higher degree, requires a form of subjective interpretation. The results of both analyses together will be included the discussion in Chapter 5 of this research.

## Description of the Exam Data

The French and German comprehensive exams, Baccalauréat and Abitur, lead to the graduation of high school and therefore provide the most comprehensive exam content when aiming to respond to Research Question 1. In this subchapter, I will reflect on the selection and sampling procedures, while providing information about the exams I believe most important to focus on. For the German exam data, there are two possibilities, the Abitur and the TestAS, which I mentioned in Chapter 2. As the TestAS, was created specifically for testing foreigners, there are many more implications involved, that would require a deeper analysis on its own, one that would not be as fruitful in comparison with the other two exams. While the SAT does have a significant number of foreign test takers, the exams purpose is to provide a general measure of higher education preparedness. Beyond that, the French Baccalauréat (short BAC) is not designed to be taken by foreigners at all (despite creating cultural conflict in France's overseas departments, see Chapter 2). Hence, in this section I will be concerned with a description of the Baccalauréat, Abitur and the SAT. I also aim to reflect on their comparability by reflecting on the exams' contents. A framework provided by the UNESCO Institute of Statistics on the comparability of certification around the globe will help in this endeavor.

### *Baccalauréat*

The sample which I use for this research is taken from *Baccalauréat general 2019* (*L* – literary focus, *ES* – economics, social sciences, *S* - STEM). The 2019 iteration of the *terminale* [last year of secondary school], the written final at the end of French secondary school. The 2019 iteration of the BAC is unique in two aspects: For once, it represents the last BAC examination pre-COVID pandemic, and second it stands as the last BAC before the implementation of new laws pertaining to the Baccalauréat (see Ministère de l'Education Nationale et de la Jeunesse,

2020a). However, after careful consideration, I found that the changes are completely structural in nature and seemingly do not touch the content of the exam itself.

The French Baccalauréat is a collectively held examination at the end of France's secondary education, for a graduation diploma after a total of 12 or rarely 13 years of primary and secondary education. At the same time, the BAC serves as an exam providing the entrance into higher education at a French public university. Moreover, the results of the exam determine which subjects the students have access to and which ones are difficult or impossible for them to access. Such is done through cut-off marks within the application procedures to French universities. In general, the French BAC is targeted at higher education and until this year, offered three general streams of specialization:

The general academic stream culminates in the *diplôme du baccalauréat général*. The technology-based stream earns students the *diplôme du baccalauréat technologique* and the vocational stream leads to the *baccalauréat professionnel*. (Getahun, 2018, n.p.)

While all three streams allow for a different *mix* from a general pool of subject-based assessments, the student's specialization determines in which subjects the students receive longer, more in-depth exams. For this year (2020) the French ministry of education (*Ministère de l'éducation nationale et de la jeunesse*, 2020) communicated the following:

*Le baccalauréat général vise la poursuite d'études supérieures longues. En 2020, l'examen du baccalauréat général reste organisé en séries économique et sociale (ES), littéraire (L) et scientifique (S) pour la dernière année.* (website)

The BAC has for a goal to enable the pursuit of normal university higher education. In 2020 for the last year, the exam baccalauréat général remains divided into specializations. These specializations are the socio-economic focus (ES), the literary focus (L) and STEM (S). [note: but this division will not persist after 2020]

In favor of further streamlining in the following years, several subcategories are being planned to be removed, where the Baccalauréat général will replace to some degree the

vocational (*professionnel*), as well as the technological (*technologique*) streams. Since for my research, I needed to sample a representative set of exam data from the BAC, I needed to decide between including solely the specialized questions for each subject area versus composing a sample of solely the non-specialization exam parts. (As the BAC S does include ES and L questions with lower difficulty and so on).

The former would then lead to a BAC sample with exam content of the highest difficulty, the latter option would lead to a representation of BAC exam content of moderate difficulty. I eventually decided to rely on a sample including only exam sheets which represent the highest difficulty for the given specialization, as it might lead to a more complete representation of the overall competencies expected from different sets of exam takers. The same rationale applies to the following exam, the German Abitur.

### ***Abitur***

#### *Exam Sample: Abitur Niedersachsen 2020 (Leistungskurse)*

Kramer, Neumann and Trautwein (2016) provided a historically descriptive account of the German Abitur in *Abitur und Matura im Wandel: Historische Entwicklungslinien, aktuelle Reformen und ihre Effekte*. Apart from that, much information is fragmented as it appears that the Abitur does not require a general description or definition for most of the German public. This is not surprising as the practice of a common exam to evaluate teenagers' ability to study has been around since at least the 18th century, before the establishment of the German nation. What also is not surprising on this background, is that each federal state in Germany still has a high degree of sovereignty in executing the examinations.

On the nation level there are three important variables that provide some uniformity within this examination: First, the Abitur has to be annually administered by every public high school within Germany to allow students to graduate from high school after 13 (or 12, depending on the state) years of schooling; second, the Abitur follows federal curricular guidelines which determine to a high degree the content of all state iterations of the exam offered; and third, the results of the Abitur directly determine which university subjects are permitted to be studied by the prospective university applicant (for many students the dream of studying medicine ends before it even begins). The governing administrative body, *Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland* (2020), provides the following statistics for each year of Abitur as shown here in Table 3.1:

Table 3.1. Between state statistics for the Abitur. *Source.* Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2020.

Sekretariat der Ständigen Konferenz  
der Kultusminister der Länder  
in der Bundesrepublik Deutschland

#### Abiturnoten im Ländervergleich 2019

Land	BW	BY	BE	BB	HB	HH	HE	MV	NI	NW	RP	SL	SN	ST	SH	TH
<b>Zahl der Prüfungen</b>	<b>51.152</b>	<b>43.067</b>	<b>13.013</b>	<b>9.803</b>	<b>2.722</b>	<b>9.479</b>	<b>23.108</b>	<b>5.420</b>	<b>33.900</b>	<b>84.465</b>	<b>17.591</b>	<b>3.336</b>	<b>13.069</b>	<b>5.981</b>	<b>14.044</b>	<b>6.962</b>
- bestanden	49.339	41.392	12.417	9.332	2.583	9.015	22.232	5.024	32.041	81.162	17.001	3.211	12.660	5.675	13.296	6.799
- nicht bestanden (abs.)	1.813	1.675	596	471	139	464	876	396	1.859	3.303	590	125	409	306	748	163
- nicht bestanden (%)	3,5	3,9	4,6	4,8	5,1	4,9	3,8	7,3	5,5	3,9	3,4	3,7	3,1	5,1	5,3	2,3
<b>Notenmittel</b>	<b>2,44</b>	<b>2,32</b>	<b>2,38</b>	<b>2,27</b>	<b>2,40</b>	<b>2,42</b>	<b>2,37</b>	<b>2,34</b>	<b>2,56</b>	<b>2,44</b>	<b>2,49</b>	<b>2,37</b>	<b>2,24</b>	<b>2,32</b>	<b>2,55</b>	<b>2,18</b>
<b>Note</b>	<b>Häufigkeit</b>															
<b>1,0</b>	726	847	260	230	55	141	549	109	369	1.587	277	68	308	144	117	200
<b>1,1</b>	476	546	154	159	28	103	312	55	224	964	162	41	210	65	91	122

The first row features an abbreviation for each German state, the second row features the total number of exams held in that state. This total is further divided into passed (*bestanden*) and failed (*nicht bestanden*). Continuing downwards, the following rows list an absolute number for each grade achieved (*Note*). Going forward with sampling, these basic descriptive statistics could serve as categories of stratification to apply to the other two exams as well.

### ***SAT (Scholastic Aptitude Test)***

*Exam Sample: SAT Subject Tests 2013-2020 (The Official Study Guide for All SAT Subject Tests)*

The American Scholastic Aptitude Test (SAT) is administered to determine access and funding within the US. Unlike the previous two exams, the SAT is managed and administered by a private organization. The best source of statistical information appears to be the website of the umbrella organization, The College Board (2019 b). In 2019 around 8 million university applicants from around the world took the SAT test. Though I found no number on how many of these were non-US test takers, the US National Center for Education Statistics, NCES (2019) published the number of high school graduates for 2018 to be 3.2 million, while 2.2 million freshmen enrolled in college in 2018.

The Board [College Entrance Examination Board] voted in 1925 to administer the Scholastic Aptitude Test (SAT) for the first time the following year during the June week of examinations. More than 8,000 candidates turned up to take the test, almost all of them applicants to Harvard, Yale, Princeton, Pennsylvania, Smith, Vassar, or Wellesley. (Valentine, 1987, p. 35)

At this point, I can only estimate the importance of the SAT for US university and college applications to be high, especially as the College Board (2019) published data not just on the numbers of SAT exams, but also on the importance of their test for college funding and grant distribution. Going forward with this research design, I will find more reliable statistical information about the SAT exam to ascertain comparability with data available for BAC and Abitur. What can be said, based on the current statistics, is that the SAT serves as a tool for selection of university access in the US similar to the previous two exams. Furthermore, due to a lack of another homogenous high school examination across the US, the SAT exam as centrally organized is the closest thing in terms of structure and national reach to the French and German exams.

### *General Considerations for Sampling*

International Standard Classification for Education (ISCED) comparability criteria will inform the categorization of all three exams. With the here following ISCED classification scheme, I justify categorizing each exam as a de facto university entrance exam, even though the Abitur, as well as the BAC are actually high school degrees. The ISCED according to van der Pol and UNESCO Institute for Statistics (2012):

The International Standard Classification of Education (ISCED) belongs to the United Nations International Family of Economic and Social Classifications, which are applied in statistics worldwide with the purpose of assembling, compiling and analysing cross-nationally comparable data. ISCED is the reference classification for organizing education programmes and related qualifications by education levels and fields. ISCED is a product of international agreement and adopted formally by the General Conference of UNESCO Member States. (p. 6)

The SAT is a choice since it is the closest comparable test to the German and French exams, given their function to determine access to higher education. The SAT is collectively administered and centrally organized. Though in the US, high school certificates do have an impact on university admission, the SAT / ICT exams are often equally or even more important to determine access to college and university degree programs.

Furthermore, unlike in Germany and in France, the US or its states do not administer homogenous examinations in all schools, devised by the state or even on the federal level. On the contrary, the SAT / ICT examinations are homogenous, centrally organized and can be located within ISCED level 3 if located within the UNESCO framework based on their function. While the SAT and the French BAC are uniform for each round / year of assessment, the German Abitur has different contents depending on the state in which it is administered. This poses a few questions concerning sampling, the most important would be which state's exam to focus on and compare to the uniform SAT and BAC.

I used a mix of purposeful and stratified sampling. Purposeful sampling at first will be guided by the Research Question: which social or cultural competencies are these exams intended to assess? For that, the steepest possible representation of exam content that the students can encounter in any iteration of the exams is used. As the SAT is uniform for each of its iterations, but in its form linked to several test dates per year, time is an important variable. First of all, all samples need to be collected from roughly the same point in time.

Concerning the German Abitur, more factors need to be taken into consideration. Unlike BAC and SAT, the Abitur is not uniform, but contains different questions for each federal state of Germany. Stratifying the sampling might pose a solution to ensure better compatibility and comparability with the former: Statistical measures are already provided by the German administration above in Figure 3.1, which includes the total numbers for tests taken per state, the total and average of passages and failures per state, as well as a total number of the grade categories into which results fall per state. At the end I needed to settle for one state's Abitur to serve as sample for this research and to be compared with the French and American data.

### ***Ethical Considerations***

The format of my research requires that I work with, translate, and incorporate German and French source material into English, and seamlessly incorporate it into an English research text. This research design can be supported by my personal profile as a researcher and translator: I am a native speaker of German with French being my second and English my third languages and I have working experience as a translator. I do as well hold a German business license for the purpose of offering freelance translations into German from English and French. I will be able to avoid some of the risks of direct translation, as discussed in Chapter 2, because these will be rarely required. I will instead rely either on indirect translations as conceptualized by Behr



and Shishido (2016), who called the procedure “adaptive translation,” or on codes which use an English metalanguage system to describe textual content (Library of Congress, 2021). Excerpts in French, which I cite from the source material and which have to be directly translated (as opposed to coded), will be cross-checked by a native speaker of French in order to minimise any threats to validity that might be caused by translator bias, as discussed in the corresponding section in Chapter 2.

Both of the methods I use, have not been exhaustively used before. On the one hand, that invites me to adapt and experiment with the methodology in application to my own data, on the other hand, the reliability of each analysis has to be considered lower than it is with already more established procedures. While the narrative literature review relies heavily on translation for the comparison of data between the three languages, the consecutive methods, network text analysis and critical discourse analysis did not face a similar threat to validity. As I am going to outline, both methods rely on a meta synthesis of information contained within the exam data; hence if done correctly, they should result in codes that can be compared between all three languages. In the light of these considerations, I explain on how I adapted each method to my specific research design.

### **Narrative Synthesis**

This method treats its sample as exam data, though neither of it is physically linked to the exams an assumption, which follows a critical pedagogy of testing (see Müller & Schmidt, 2009) and my reasoning to treat hidden structures pertaining to each exam not as exherent, but inherent to analysis. Language holds power, as described in section 2 of Chapter 2, therefore it appears natural to me to analyze what *is said* about all three exams. In the following, I discuss sampling

strategies and analyses with an application of narrative synthesis upon curricular and other closely related exam documents.

Following the wholistic lens that I imposed on myself in making sense of the exam data, a narrative synthesis helped to identify those causalities which stand at the creation of the exam's content and institutional frameworks. As laid out in the methodology, the method required practical adaptations concerning data collection and analysis, especially as I focused on research specifically linked to the creation of each of the exams in question. Bennett, Driver and Trent (2017) introduced the following as rationale for the inclusion of articles to their query: "The first and second authors completed simultaneous electronic and ancestral searches for peer-reviewed articles using the online database *PsycINFO* and seven databases from The Elton B. Stephens Co (...)" (p. 896).

I deviated from the approach of a synthesis of all published research, as my goal here was not to give an overview of the most current published research for university entrance exams, but to give a perspective on whatever research is linked to the creation of the actual content. There is the factor time involved, which will require to go back and uncover research linked to the creation of my selected high-stakes exams, that is no longer openly associated, nor mentioned in relation to these entrance exams in current publications. Guided by the outcome of the narrative literature review, I generated a rationale for comparison, before collecting current versions of university entrance exams (SAT, Abitur, BAC)

### ***Sampling***

The sampling procedure for the narrative literature review involved a synthesis of published research on all three exams and therefore, it worked with a different set of data than initially collected in the form of the exam objects. The relationship between the exam and

published research had to be established, to replace query guided sampling as laid out in the methodology section. A narrative literature review originally (Bennett et al., 2017) measured a scientific discourse in the present or without giving much consideration to the variable that was the date of publication. Inspired by the idea that cultural and social capital can only be fortified through their continued reaffirmation over time, I was required to include older documents. Though that imposed some added difficulty for my sampling, I remained open to the inclusion of texts published on the SAT, the Abitur, and the Baccalauréat throughout the time of their existence.

My sample eventually came to include historical texts dating back to the 18<sup>th</sup> century. They had a significant impact on the design of the exam at the time of their publication. There is no erasure trace or stamp of removal. In fact, an initial review of available texts for all three exams suggested, that it was entirely possible that ideas had been transplanted from generation to generation, to the point of their persistence in the present. All these texts can be considered impactful as they often represent the sole publication about any of the three exams by the exam stakeholders. Unlike sampling documents based on rationales like scientific impact, as is often done by narrative meta-analyses, I needed to redefine *impact*. *Impact* here meant carrying social and political weight towards the exams. The major stakeholders for the three exams remained more or less the same, from their earliest mention to today. While the *College Board* as a major stakeholder of the SAT underwent a slight name change from *College Entrance Examination Board*, the German and the French mentions of their respective exams date further back and include publications done by government bodies belonging to empires in place of today's modern democracies. While the French and German political systems underwent major changes

over the last 300 years, the authority over the Baccalauréat and the Abitur has always been singularly positioned with the governments and its respective bodies of education.

After a thorough review of the available literature, I selected reference chronicles for each exam to locate and evaluate historical publications. I selected Bölling's (2010) *Kleine Geschichte des Abiturs*, and particularly his timeline of the most influential legal documents pertaining to the Abitur (cf. p. 197). For the Baccalauréat, I used Meuriot's (1919) *Le baccalauréat. Son évolution historique et statistique des origines (1808) à nos jours*, and a resource provided by the French Ministry of Education (*Ministère de l'Education Nationale et de la Jeunesse*, 2020b), which provided official key dates to the Baccalauréat, as well as official publications. The historical texts that I eventually sampled, I found in various digital archives. The exact source for each included historical publication I listed in Appendix 1.

Another sampling strategy included looking for texts which are directly published on the respective websites of each exam's governing entity. These entities are: The *Ministère de l'Education* of France, the *Kultusministerkonferenz* for Germany and the *College Board* for the US-based SAT exam. This strategy exclusively produced modern publications to each of the exams, which could be best described as being curricular in nature.

Finally, I also sampled texts involving calculation of impact factor: In the world of journal publishing, one important constant is the impact factor of a journal. This factor represents a comparative quantitative tool for measuring the number of citations. McKiernan et al. (2019) summarize the statistical operations underlying Impact Factor Calculation. The journal impact factor is calculated over a given time period, as well as through the total and relative amount of citations. Both measurements are available for and could be transferred to single publications. IN

that case, the total citations in a journal here could be simply replaced with the total citations containing keywords related to one or more of the selected exams.

I was able to use Google Scholar's *citation count* to search documents based on a rationale similar to the impact factor. As I could not directly establish the relationship between citation count and impact factor, I added the search criterium *publisher* to the query. As publisher, I used the national stakeholders for each exam listed above. After a first trial, I realized that all three entities have published peer reviewed articles, which fall into the search scheme of Google Scholar. The query hence looked as following: exam name [title] AND governing entity [publisher].

I have used different ways to describe the conceptual proximity, which all texts considered entertain with the exams. For the sake of simplification, I will continue to reference these texts as *curricular documents*, though the attribute is at least questionable.

### ***Validity and reliability***

In the light of the publication base of this method, an obvious threat to reliability with my altered design imposes through the fact that measures of proximity have not been defined as such before. Then again, the method appears to still be in an experimental state, as elaborations on its general reliability fall short, and were rarely mentioned during my literature review therein. In that light, there is an obvious threat to validity of my proposed design: My methods for data collection are slightly altered and add these above statistical procedures to produce a sample that better fits my research design, but also in order to improve of some shortcomings in terms of querying that I encountered. For example, Bennett and Driver's (2017) and Petticrew and Roberts' (2006) studies both appear to lack a description of why and how query terms were included or rejected. I included the linguistic statistics above to remedy a shortcoming I

perceived when I decided which patterns are worth mentioning within the text data, and which patterns need to be rejected.

The method tries to make away with the shortcoming of expert interpretation and with a degree of subjectivity, a widely accepted shortcoming of qualitative research. By using transparent algorithms to guide the inclusion of texts and the creation of linguistic patterns, the narrative literature review can replace subjective decision making to some degree. Popping (2000) described such procedures as *logical formalism*: “Logical formalism is found in expert systems. An expert system is a computer application that performs as if it were a human expert. It uses a rule base, a knowledge base and an inference system (p. 4).” Popping also advocates the use of linguistic statistics to realize logical formalism.

### ***Systematic Sample Reduction***

I conducted a review of all available texts pertaining to the SAT, Abitur and Baccalauréat to arrive at an idea of the population size  $N$  from which a sample can be drawn. There are three key dates, one for each exam, which set the commonly assumed onset of each exam, and the beginning for inclusion of texts directly related to each exam towards data extraction. Pertaining to the Abitur, I used two historical publications (Bölling, 2010; Bundeszentrale Politische Bildung, 2016) as frame of reference. There appears a general consensus within German discourses, that the first referable iteration of the modern Abitur was established per decree 1788 in the kingdom of Prussia. Hence the possible population segment of texts and publications linked to the Abitur, in theory, can include everything published between 1788 and today. This posed several challenges: First, it predated the founding of the German nation by roughly 100 years. I needed to find rationales to link texts about the Abitur that were published by monarchs

and by an emperor with those that were and still are published surrounding the Abitur as the modern democratic educational institution it is today.

Similar problems did not arise with the French Baccalauréat and the SAT test. There is some consensus (Ministère de l'Éducation, 2020b; Marchand, 2009) that the former was created under Napoleon I., i.e. the year 1808, when France already existed as a nation, and when the country's form of government, the republic, already resembled that of today. Nevertheless, the data extraction for the Baccalauréat also posed several challenges, which I discuss in detail in the following subchapter. The SAT has been comparably well documented, from its inception in the 1920s (Valentine, 1984) to its modern iterations today. Then again, all tests yielded significantly different results along the three procedures of stratified sampling. I specifically focused on describing these rationales of data extraction. Each procedure is informed by notions of structural proximity or dominion, which I expressed in the previous sections of this chapter. As a result of all sampling techniques, minus a differentiation provided through text evaluation by cite count, I generated a population  $N$  containing 2268 texts. Figure 3.2 below represents all steps necessary to generate this population.

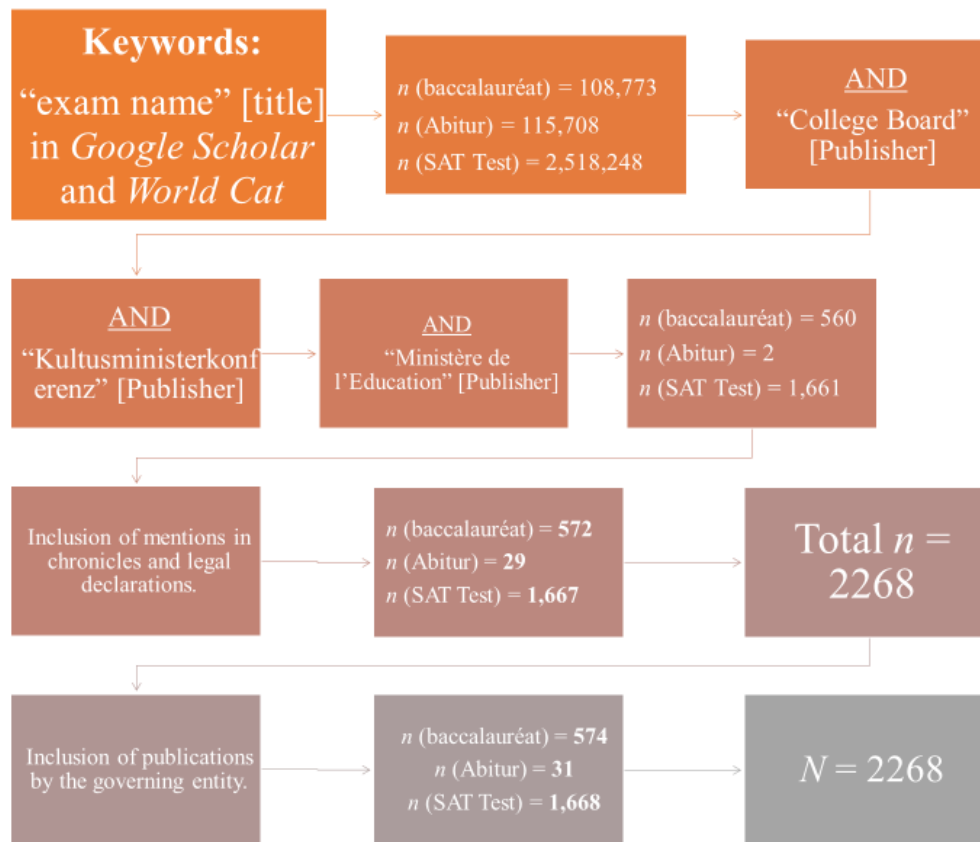


Figure 3.2. Generating a Population  $N$

Like Leamy et al. (2011), I reduced the population size either by hand or through the cite count rationale. Given the large number of texts included without further query attributes, I systematically had to reduce the sample size. This became necessary as the final sample needed to reflect the relevancy of several types of publications therein included, for all three exams. Aspects or relevancy are a limited selection of properties, relations ascribed to each publication in view of its impact and *control* over each exam. *Control* is a word, which I use as outlined by Müller & Schmidt (2009). Their framework of research enables the following procedures to become valid for data extraction in the first place. Especially the subdomain of their category control which is *legitimization*, plays a larger role here. On page 25 ff., the Müller and Schmidt (2009) describe an essential function of assessment as legitimizing different forms of educational



achievement and mental properties embodied by the test takers, hence giving social substance to ethereal and naturally invisible dispositions. At first, this theorem plays a large role to give significance to the different publications which govern each exam in its national context. More precisely the following task of sample reduction becomes feasible when selectively focusing on relations of control and governance. Each of those concepts informs one of three systematic procedures to reduce the size of the sample.

### ***Publications by the governing entity***

For purposes of quality assessment and further reliability of the texts included, I initially excluded all texts which were not directly published by the main governing body, being in charge of each exam. The texts which fall under this category have been limited exclusively to recent publications, no version is older than five years, the format of the texts included here is best described as educational standards.

As shown in Figure 3.2 above, the so added texts were no more than five in total. The low number also relates to what I perceived as an urge for streamlining and standardization with all three entities. An example would be the document *Vereinbarung zur Gestaltung der gymnasialen Oberstufe und der Abiturprüfung* with its omitted version from 1972, and its only published version dated 2016. While being subject to more than 40 years of input and change concerning the Abitur, the only version of the document published is the one representing the current iteration of the exam since 2016. The same is true for the Baccalauréat standards document provided by the French Ministry of Education and the standards document from the CEEB.

### ***Impact Factor Search***

The search in Google Scholar was not limited by dates the only criteria were, as stated in the figure 4-1, “publisher” AND “topic”. The initial query resulted in 935 results, which I then had to narrow down. In lack of additional reviewers and coders, I used the rationale *impact factor* (Chapter 3, section 4), here presented through Google Scholar’s proprietary citation count. The table below includes all articles which have more than 50 citations to their name. This query revealed significant structural differences among all three governing entities. First, the query for the Baccalauréat revealed 560 publications, which seems quite impressive! A search on the official website for publications (Ministère, de l’Education Nationale et de la Jeunesse, 2020b) with the same query term resulted in merely 480 publications. Not only did the systematic query in Google Scholar provided 80 results more, it also included Google’s proprietary citation count, which is the key rationale for establishing the impact factor. Though numerous, only very few citations were marked as being cited more than 50 times; They all are listed in table 3.2 below. Similarly, the query of publications pertaining to the SAT test yielded 1661, a large number of results (see. Figure 3.2 above). It remains to say that the College Entrance Examination board appears to value scientific, peer-reviewed, outlets even more than the French ministry. The fact that much of the 1661 publications are published in competitive CEEB owned journals speaks to that as well. Given this structure, it also seemed to me that the citation count was the most efficient indicator of impact among all three exams. Table 3.2 below lists all documents which had a citation count larger than 50.

On the other hand, the query for the Abitur provided merely two results pertaining to the query criteria. Similarly, the official website of the exam’s stakeholder, the *Kultusministerkonferenz* (2020b) has very few publications. All texts published there fall into

either one of the following categories: official statement, or standards documents. Official scientific publications like the ones I found with the French Ministry, appeared to be all lacking the information I am looking for here (curricular input, reflections on social or cultural competencies required for the exams). Given these observations the reliability of my here applied sampling rationale appeared low, and I did not include the two results.

In a final step of the sampling procedure, I looked at the abstracts of each of the texts listed in Table 3.2 and Table 3.3. With the overarching Research Question 1 in mind, I marked the most promising candidates red for inclusion, or greyed out those, of which after a closer reading, I believe carry not much significance towards the research objective. This resulted in a final sample of five texts, of which all are from the College Board, and none from the selection of the Ministère de l'Éducation.

Table 3.2. Impact Factor SAT Test

Publication by College Board	Google Scholar Citation Count > 50	Google Scholar settings: Search term "College Board" in category: publisher; Search term "SAT" in category: topic (10.18.2020)	
Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). Validity of the SAT® for Predicting First-Year College Grade Point Average. Research Report No. 2008-5. College Board.	483		
Mattern, K. D., Patterson, B. F., Shaw, E. J., Kobrin, J. L., & Barbuti, S. M. (2008). Differential Validity and Prediction of the SAT®. Research Report No. 2008-4. College Board.	155		
Burton, N. W., & Ramist, L. (2001). Predicting Success in College: SAT® Studies of Classes Graduating since 1980. Research Report No. 2001-2. College Entrance Examination Board.	373		
Camara, W. J., & Echternacht, G. (2000). The SAT [R] I and High School Grades: Utility in Predicting Success in College. Research Notes.	383		
Chall, J. S. (1977). An analysis of textbooks in relation to declining SAT scores.	149		
Kobrin, J. L., Sathy, V., & Shaw, E. J. (2007). A Historical View of Subgroup Performance Differences on the SAT Reasoning Test™. Research Report No. 2006-5. College Board.	91		
Wyatt, J., Kobrin, J., Wiley, A., Camara, W. J., & Proestler, N. (2011). SAT Benchmarks: Development of a College Readiness Benchmark and Its Relationship to Secondary and Postsecondary School Performance. Research Report 2011-5. College Board.	71		
Everson, H. T., & Millsap, R. E. (2005). Everyone Gains: Extracurricular Activities in High School and Higher SAT® Scores. Research Report No. 2005-2. College Board.	64		
Camara, W. J., Copeland, T., & Rothschild, B. (2005). Effects of Extended Time on the SAT® I: Reasoning Test Score Growth for Students with Learning Disabilities. Research Report No. 1998-7. College Board.	68		
Mattern, K. D., & Patterson, B. F. (2009). Is Performance on the SAT® Related to College Retention? Research Report No. 2009-7. College Board.	61		
Patterson, B. F., & Mattern, K. D. (2009). Validity of the SAT® for Predicting First-Year Grades: 2009 SAT Validity Sample. Statistical Report No. 2012-2. College Board.	54		
Stanley, J. C. (1986). SAT-M Scores of Highly Selected Students in Shanghai Tested When Less than 13 Years Old. College Board Review.	98		
Angoff, W. H. (1974). Criterion-Referencing, Norm-Referencing, and the SAT. College Board Review.	64		
Stanley, J. C. (1977). The predictive value of the SAT for brilliant seventh-and eighth-graders. College Board Review, 106, 31-37.	60		
Schneider, D., & Dorans, N. (1999). Concordance between SAT® I and ACT™ Scores for Individual Students. Research Notes. RN-07. College Entrance Examination Board.	55		

Table 3.3. Impact Factor Baccalauréat

Publication by Ministère de l'Education	Google Scholar Citation Count > 50	Google Scholar settings: Search term "Ministère de l'Education" in category: publisher; Search term "baccalauréat" in category: topic (10.27.2020)	
Vallet, L. A., & Caille, J. P. (1996). Les élèves étrangers ou issus de l'immigration dans l'école et le collège français: une étude d'ensemble. Ministère de l'Education nationale, de l'enseignement supérieur et de la Recherche, Direction de l'Evaluation et de la Prospective.	253		
Moisan, C., & Simon, J. (1997). Les déterminants de la réussite scolaire en zone d'éducation prioritaire. Rapport pour le Ministère de l'Education nationale.	138		
Caille, J. P. (2001). Scolarisation à 2 ans et réussite de la carrière scolaire au début de l'école élémentaire. Éducation et formations, 7-18.	95		
Nakhili, N. (2004). Impact du contexte scolaire dans l'élaboration des choix d'études supérieures des élèves de terminale. Éducation et formations, 72, 155.	63		

The decision to exclude some articles (greyed out above) was made after reading the abstracts and some of the content of the articles presented with RQ1 in mind. For that matter, none of the texts listed in Table 3.3 for the Baccalauréat included useful statements about the exam, as they all discussed effects and results of the exam, not touching the assessment itself. The two articles relating to Abitur, failed to meet the impact factor threshold of 50 and, judging from the abstract, both appeared irrelevant for inclusion given the above rationale.

### ***Re-tracing within chronicles***

As described, with this third approach, I focused on extracting key dates and decisions surrounding each exam, made public in form of documents at any given time. Figure 3.1 in the previous section gives the numbers of texts added to the sample. Considering the SAT exam, I followed a historical examination by Valentine (1987) published as a monography by the College Entrance Examination Board (or *CEEB*). This helped me get references for eight publications, all of them from between 1919 to 1934 ((inception phase of the SAT; all listed in Appendix 1). Getting access to all eight texts proved to be simple, as all were available at

Louisiana State University's library. A point of concern here would be the fact, that there is a gap in time of publication in the SAT sample, as it does not include texts published between 1934 and the 1970s. The reason for that being on one hand, that it was impossible for me to still locate curricular documents from that timespan, though some were mentioned in chronicles. On the other hand, it can be assumed that the College Board makes an effort to have older curricular documents available.

Considering the Abitur sample, this data extraction procedure proved to be yielding the majority of texts. I used a chronicle published as a monograph by an established German professor of education and history, Rainer Bölling (2010). In his work, he included the exact information I was looking for, decrees and official communications from the governing body of the exam, in form of a timeline from 1788 *Erstes Abiturreglement in Preußen* to the most recent publications by the *Kultusministerkonferenz* (for explanations and summaries see appendix 1). In total, I so identified 26 references, with contextualization on the significance and on the social and political context of each. After further reading into the nature and content of each of these 26 publications, I opted to include seven into the Abitur sample. Unlike with the SAT, the date range of these 7 publications is rather evenly distributed among the 222 years of the existence of this exam, minus the time of the Nazi reign.

For the baccalauréat sample, I relied on three different sources of historical information and contextualization, the first being a historical document published by the *Ministère de l'Education Nationale et de la Jeunesse* (2020a), which describes the first modern iteration of the baccalauréat as being intrinsically linked to the reign of Napoleon I, with significant publications starting from 1808. The second text is a chronicle by Marchand (2009) published in a peer reviewed French journal, and the third being a chronicle by Meuriot (1919) published at the

beginning of the last century in *Journal de la société statistique de Paris*. All chronicles combined helped me to add 12 references to the baccalauréat sample.

### ***Analysis (Coding)***

As outlined previously, the coding of my sample of texts from all three languages happened in two steps, one being inductive, and the other deductive coding. Both steps require a coding rationale which is provided through research question number one, the question of social and cultural competencies as connected to the perfect score in the three selected exams. For the data analysis of the so collected textual data, I believed *deductive codes* as proposed by Bennett, Driver & Trent (2017) presented a good fit, as the definitions of cultural and social competencies are open ended, and I did not have to include any form of prior definitions. Deductive codes are also a convenient choice, as at first glance, they represent a perfect extension of the logical formalism discussed above. Bennett, Driver & Trent's (2017) deductive codes were based on the definitions of certain key concepts he drew from literature on critical race theory.

Denzin & Lincoln (2002) mention the possibility of using so called *keyword in context* (KWIC) as linguistic tools used in the analysis of textual content. Such KWIC rely on correlations between keyword and context, which are defined through several structural relations. KWIC will rely on collocations as well and I will here continue to operate with collocations as defined during the sampling subchapter. The initial determination of KWIC will inform a search for deductive codes. Bennett, Driver and Trent (2017) describe the search for deductive codes as following:

Based on multiple readings of the selected 26 articles, deductive codes were determined insufficient in capturing the full expanse of results. Therefore, inductive codes were created to strengthen analysis. For example, the deductive code “social construction of Whiteness” was considered too broad and inductively changed to both “Whiteness as property” and “racial positioning” based on a further reading of the articles. (p. 897)

I followed the same analytic design here, but instead of solely relying on my reading to reveal *inductive codes* from the text sample, I also used an additional layer of logical formalism to guide my reading. This logical formalism was enabled through the use of KWIC and linguistic statistics. *Keyness* can be described as the relative frequency of terms appearing in a specific corpus in relation to a generalized reference corpus for a specific language. *Keyness* refers to the higher/lower frequency of particular words (termed keywords) in one corpus when compared to their frequency in another corpus (Scott 1999). The significance level of this frequency difference is statistically calculated via log-likelihood tests. Keyword analysis was used to determine words which were significantly more frequent in broadsheets and tabloids (pp. 24-25). The math behind *keyness* is often associated with problems early corpus linguists faced when applying traditional statistics (for example parametric tests) to textual data. Ted Dunning (1993), a linguistic statistician describes the problems as following:

Statistics based on the assumption of normal distribution are invalid in most cases of statistical text analysis unless either enormous corpora are used, or the analysis is restricted to only the very most common words (that is, the ones least likely to be of interest (Dunning, 1993, p. 11). He (1993) goes on to contrast the use of two methods of statistical tests used to provide what came to be called *keyness* in more modern publications: the first method relied on Pearson's Chi-Square statistic, while the second analysis happened through the log-likelihood correlation ( $-2 \log A$ ) (Dunning, 1993, pp. 68-70). He looked at the significance of occurring *word bigrams* or *multigrams*. His results showed that the log-likelihood generated word bigrams coincided more neatly with semantic significance on closer inspection. Beyond the scopes of his paper, Dunning's experiment sheds light on the nature of *keyness* and its relation to *collocations*. *Collocations* are generated around the core parcel of query terms



combined with their most significant neighbors where the occurrence is not random, or due to chance. In other words, collocations are correlations among words, where the occurrence of these pairs or triplets is predicted to be higher than chance.

Unlike Baker (2007) or Koteyka (2010), I used the tool *SketchEngine* (Kilgariff, 2004) from the website [www.sketchengine.eu](http://www.sketchengine.eu). SketchEngine is a fairly advanced statistical toolbox for corpus linguists, which offers functionality similar to *WordSmith* and beyond that of *AntConc* used by Koteyka. Furthermore, SketchEngine offered accessibility and transparency. The concrete statistic procedure which underlies SketchEngine's keyness scores is slightly altered from the statistics described above, but still relies on the same comparative procedures. At the end, a feasible sample for analysis must be reduced to the amount that the resulting text can be read multiple times by one person, the researcher, myself, as reading will be key to the following data analysis. Hence I was able to decide on size of the sample myself, depending on the time constraints and my personal ability to read.

The previous results in a tabulation of qualitative data as being most relevant to the research question. That means a criterion for inclusion would be that the publication under consideration makes succinct statements about *social competence*, as well as *cultural competence* aligned with the three functions of assessment outlined in Chapter 2. First function: Dominion and Social assimilation (reproduction of the system), second function: Recruitment (selection, election, allocation, standardization) and the third function: Didactics (see Chapter 2).

As outlined by Petticrew and Roberts (2005, p. 191), a narrative synthesis of the content of qualitative text can be facilitated by some useful online software. I chose SketchEngine (Kilgariff, 2004), which is a fairly advanced statistical toolbox for corpus linguists, that can be used for narrative synthesis. Furthermore, SketchEngine offered accessibility and transparency.

Following Leamy et. Al.'s (2011) conceptual approach of "inductive, open coding techniques" (p. 446), I created three text corpora including all texts for each language sample.

For each language reference corpora are provided through the team of the British / Czech Lexical Computing Research company founded through and continuously developed by Kilgarriff et. al. (2014) collected from a wide range of digital sources from the year 2007 forward. Within the reference corpus for English 15,703,895,409 words are contained from published and written English from a wide range of sources (e.g. newspaper, blogs,). The same holds true for the German reference corpus, *DeTenTen13*, which contains 16,526,335,416 words. Both reference corpora were accessible through the SketchEngine interface through which I could query them for single terms or combinations of terms (collocates). Another important mechanism provided through the tool was determining the word frequencies and keyness scores of all the different words contained between different collocates from the reference corpora (or any consequently generated corpus for that matter).

### ***Coding***

The coding of term combinations with the highest keyness across all sampled documents resulted in the expression of certain themes with each theme's significance being underlined through the sum of keyness of all the term combinations it included. These themes served as an expression of *cultural and social competency*, and represented a new application of Bourdieu's theorems.

Cultural capital can exist in three forms, two of which are most relevant here: in the *embodied* state, i.e., in the form of long-lasting dispositions of the mind and body, in the *objectified* state, in the form of cultural goods (pictures, books, dictionaries, instruments, machines, etc.), which are the trace or realization of theories or critiques of these theories,

problematics, etc.; Bourdieu (1986) states that cultural capital in its objectified state has properties acquired only through understanding cultural capital in its embodied state. In consequence though, objects can transmit the material realization of these embodied cultural properties. For the exams, that means I need to search and identify traces of all and however formed, embodied properties the students need to possess in order to score a perfect mark on each exam (cf. Bourdieu, 1986, p. 249 – 251). In relation with Research Question 1, I grouped the lists of terms looking for hints of realization of these properties. I searched for the expression of dispositions of the mind and educational development as pertaining to Research Question 1. With the words of Bourdieu I looked for traces of student “habitus” (cf. Bourdieu, 1977) in the expressions of the terms. The results are presented in Chapter 4.

### **Network Text Analysis**

The second method I employ specifically served to respond to Research Question 2: How is social or cultural capital embedded within and revealed in comparison across the Baccalauréat, Abitur and SAT Test? As previously mentioned, my application of network text analysis will deviate from previously mentioned research, through the fact that I did not look at relationships among lemmas, words or phrases within a document, but rather at both physically visible and invisible information. It treated the exam text as both revealed and concealed. The visible part consisted of the questions, while the invisible part were the responses to these questions. As such, the application of the method itself, and the degree to which it succeeded to produce significant results, can be seen my take on responding to the problem of revealing social and cultural capital, which is embedded in textual and outside of text structures. The notion that this is feasible stems from the holistic approach to assessment, which I presented in Chapter 2 section 1.

Furthermore, the application of network text analysis (abbreviated *NTA*), is inspired by both Vygotsky (1978) and Bourdieu (1986), hence the expected results were analyzed accordingly. This has never been done before, exam data has never been sampled for that purpose and the approach to analysis in connection with the theoretical framework is entirely new. As NTA aims to reveal socially determined, already existing structures, it is not an entirely hermeneutical exercise. Upon careful examination I decided to rely on commonly and internationally used classification schemes, more specifically, the Library of Congress (2021) classification, to determine type – token relationship inspired by Popping (2003). I also relied on the existence of semantic networks as per previous definition, which are represented in artefact form through the structure and content of exams. The determinants for my semantic networks are produced through the relationship between questions and answers, both integral parts of exam data. The following Figure serves to illustrate what I describe above:

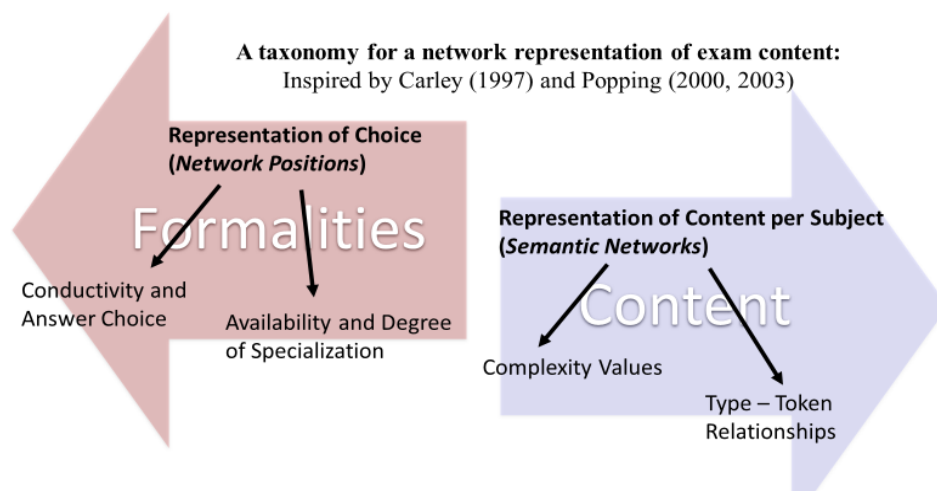


Figure 3.3. NTA, Application to Exam Data

The two sides of Figure 3.3 both signify sources of data for this method. The content analysis focuses on exam questions and responses (provided through answer keys). In another important step of analysis, I looked at the determinants of exams, which go beyond exam content

in the form of question and response. I termed the latter in lack of another good descriptor “Formalities”, though as I came to understand, they much rather coincide with networks in the sense of semantic networks, as presented by Popping (2000, 2003).

### ***Sampling***

Table 3.4 below illustrates the extent of my exam sample. I was able to collect exam data in the form of question sheets and answer keys from all three exams. The Abitur exam data is from 2020, which I received from the German state of Lower-Saxony. Though the state publishes the entirety of its 2019 Abitur exam questions, many subjects are simply not included in the state’s current iteration of the Abitur, as they lack a representation in schools. Similar issues of access arose with the SAT and Baccalauréat samples. The owner of the former (The College Board) does not publish exam question sheets in its entirety. I received access to both, question sheets and answer keys, through a publication from 2013, which is commonly mentioned as preparatory standard work (references) for the SAT subject tests until today. Though the French government does provide exam sheets in their entirety, they do lack publications of official or endorsed answer sheets. I had to rely on semi-official answer keys for my analysis, which I thematize in Table 3.4. The sample I collected is from the *Baccalauréat général (E, ES, L)* of the year 2019. All field specifically marked with an N in the table below pertain to the diverse and colorful catalogue of several West-German states, whose exam questions are sadly not freely accessible to the German public.

Table 3.4. The Results of Sampling Exam Content

Subject	Abitur	SAT Test	Baccal auréat	Subject	Abitur	SAT Test	Baccal auréat	Subject	Abitur	SAT Test	Baccal auréat
Math	Y	Y	Y	Modern Greek	N	X	Y	Nutritional Science	Y	X	X
Literacy Native Language	Y	Y	Y	Italian	N	Y	Y	Pedagogics	Y	X	X
History	Y	Y	Y	Ancient Greek	N	X	Y	Hebrew	N	Y	X
French	Y	Y		Geography	Y	X	Y	Japanese	N	Y	X
English	Y		Y	Music	Y	X	Y	Portugese	N	X	X
Biology	Y	Y	Y	Agriculture	N	X	Y	Islamic Religion	N	X	X
Spanish	Y	Y	Y	Cinema	X	X	Y	Psychology	Y	X	X
Physics	Y	Y	Y	Circus Arts	X	X	Y	Law	N	X	X
Latin	Y	Y	Y	Dance	X	X	Y	Turkish	N	X	X
Chemistry	Y	Y	Y	Theater	X	X	Y	Chinese	N	Y	Y
German		Y	Y	Dutch	N	X	X	Korean	X	Y	X
Art History	X	X	Y	Physical Education	Y	X	X	Political Science	Y	X	X
Social Studies	Y	X	Y	Jewish Religion	N	X	X				
Economy	Y	X	Y	Linguistics	N	X	X				
Philosophy	N	X	Y	Christian Religion	N	X	X				
Geography	Y	X	Y	Ethics	Y	X	X				
Creative Arts	Y	X	Y	IT	Y	X	X				
Engineering	N	X	Y	Russian	N	X	X				

The exams are represented above by columns, where the corresponding subjects occupy rows. All cells marked with green are part of the analysis as they belong to subjects, which are represented across all three exams, the SAT, the Baccalauréat and the Abitur. Cells which are not colored and marked with an X indicate subjects that are not available in the exam. Access was a major challenge to the sampling procedure, as none of the stakeholders did sufficiently provide exam data, the questions and answer keys, in their entirety. Where I was absolutely not able to receive access to the exam data is marked with an N in Table 3.4, which, unlike X, does not indicate that the subject is not tested in the corresponding exam. Y indicates that the exam data was available, both in form of questions and responses. An important adaptation I had to make to the methodology of network text analysis, is the subsummation of answers under exam data. That necessity arose from analytical standpoints and is better explained in the following

subchapter on Analysis. For the sake of my later analysis, I assumed that an exam question is not complete without the officially recognized response, or the perfect response as perceived by graders. For all three exams, graders rely on official grading guidelines, so the question cannot be perceived as exemplifying exam data without the corresponding response. Questions and responses share the same theme / topic, but only through the response, the mental operations required to succeed in the exercise become understandable.

Referring to the responses, ideally meant having access to an officially endorsed answer key/ teacher sheet. As I mentioned before, I succeeded in sampling the latter for the Abitur, the SAT, but have to rely on semi-official answer keys for the Baccalauréat. Official endorsement is missing here, as I had to rely on anonymous uploads on an educational files sharing platform. Webpedago (n.d.) aims to collectivize knowledge and provide access. At the same time I was not sure who the author of these answer keys was. I had to use a careful assessment of the answer keys provided based on my own knowledge of the subject matter. As a result of that, I included only subjects of which I was confident to have appropriate content knowledge. The content analysis in Chapter 4 is limited to Native Language Literacy (for each respective exam), History and Mathematics. As such I did not include the other subjects (see Table 3.4) Physics, Chemistry, Spanish and Latin in my comparative analysis.

### ***Thesaurus creation informed by artifact and cultural capital***

As discussed for the methodology, the creation of a thesaurus is critical for any consecutive content analysis as it informs local networks, but the thesaurus did also to some degree influence the structured representation during the sampling. As it plays a role for both steps, and marks the clearly interpretative aspect of this method, I decided to subsume the following information on thesaurus creation under data analysis. Carley (1997) established the

pre-existence of ideas about the content of the text as a prerequisite for the creation of concepts. Such concepts then lead to the creation of a thesaurus. For the specific purpose of my data analysis, I can choose from several ideas for the creation of a thesaurus, but most importantly, I needed to be guided by the elements discussed in Chapter 2 and specifically on Chapter 2, developmental and learning ideals, as these were determined per my research design, what I am looking for within the data of my exam objects. Pinar's (2004) take on hidden curriculum in this context especially seems to perfectly fit the requirements for this thesaurus. Pinar (2004) theorizes the dissolution of the human subject, the dissolution of text as linear knowledge and a replacement of those through decentered networks of knowledge grouped around "nomadic centers" and "provisional structures" (p. 139).

Above all, Vygotsky's theory on the relation between learning, symbol and artefact, as laid out in Chapter 2, presents a valuable framework. Vygotsky (1978) acknowledges that learning happens through symbols, as well as their mediation through a child's social environment, but on the other hand, a symbol as well as its context, would represent the foundation of a proper and individual inner structure created in the child's psyche. With the aim in mind to formulate a thesaurus for the NTA, Vygotsky's concept of artefact mediated cognition (pp. 54-55) can serve as higher-level abstraction for the previously identified concepts, where the textual concepts enter the role of signs, per definition given for sign above, which control, whether or not, the exam takers have the necessary tools to get to the correct response. Another theoretical mention to inform the creation of a thesaurus here would be the idea of hidden curriculum as outlined in Chapter 2.

As the aim of this analysis is to reveal semantic networks underlying assessment, I chose to use a preexisting classification system of content. Classification schemes and consecutive



tracking of subjects plays a large role in the Abitur and the Baccalauréat, as the analysis in chapter 4 proves, but subjects and rationales for grouping certain subjects together often remains institutionally undefined.

Consensus about limitations of subjects towards other subjects is often absent. Hence the best way to establish subject as a coding rationale was to include an exterior system of classification. I chose the Library of Congress (LOC) classification of subjects, as it appears one of the most widely used formats internationally (Batley, 2014; Library of Congress, 2021).

### ***Validity and reliability***

The validity of my application of NTA relies on a sound implementation of the statistical procedures underlying semantic networks. Unlike above with the narrative literature review, this procedure will primarily inform the consecutive data analysis. While I was not able to completely eradicate subjectivity and the process of subject interpretation, the semantic network approach allowed to generate an analysis of the exam's content spanning all subject fields represented within each exam, through clear categorical analysis and representation of content. For the purpose of analysis, Carley (1997) calls for the creation of a thesaurus, which is a subjective practice and needs to be customized to my own theoretical framework and research design. For this to not devolve into an exercise in hermeneutics, I employed an already existing framework of representation of subject content, the Library of Congress System of classification of content in the following analyses. Further threads to validity include the fact that I have to use different years of publication of the exam data. This became necessary through the difficulties I faced sourcing official answers for each exam, especially the SAT.

## *Analysis*

While the already collected exams served as data, the sampling process for this method included the gathering of focus terms, as well as the creation of a local network. More specifically I relied here on Popping's (2000, 2003) "semantic network" approach to create, what he called, a "structured representation." A *structured representation* of the content of texts is here opposed to "logical formalism", which according to Popping (2003) relates to the corpus linguistic analysis I used for the narrative literature review above (pp. 92 – 94). "The goal of a semantic network is to make clear the structure of knowledge, so it aims at a structured representation" (Popping, 2003, p. 92). When looking at exam content, I required a *structured representation* of both, the properties of exam questions, as well as those relating to the endorsed, correct response to the question.

Unlike texts, such content cannot be read directly without an additional layer of abstraction in place. A structured representation of such content helps to unify the way I then talked about the data. That does not mean that a *semantic network* can rely on my interpretation alone (as opposed to hermeneutics). Following the network theory of Popping (2003), I had to structure my own network according to the relationships between *tokens* and *types*. According to Popping's (2003) theory:

A token denotes an individual that can be chosen from a universe given by the discourse. as an example, "Pluto" is a token, and "dog" is a type. The choice of the individual might be restricted, and the restriction follows from the relations attached to the token. The relationship between token and type is denoted by ALI (alike). There are seven relations between types: PAR, CAU, AKO, ORD (ordering), ASS (symmetric association), EQU (equal, symmetric) and DIS (distinct). (p. 96)

Whether the token "Pluto" above refers to the planet or the Walt Disney character determines its "type." Here the author chose the latter, and therefore token and the specific type enter an ALI connection. Then the author goes further into semantic connections between types,

which span linguistic categories from syntax, semantics, as well as pragmatics. Without further going into the seven relations listed, I will shortly present a network example for CAU, a causal relationship, presented within the same publication.

Popping (2003) on page 103 gives a rough idea on the shape of my envisioned sampling through the figure depicted on that page. Popping provided a structured representation of a discussion of the labor market within a doctoral dissertation. While the concepts needed to be defined through the research intent of the paper he analyzed, the operators allow to mold information within different texts into a uniform structure, which in turn will allow for comparison among different sources of text and even across the borders imposed by language and translation. As the structure of the so produced sample is roughly described here, another issue remains. In the methodology section, I stated the need for a thesaurus; this need became apparent here again and I realized that Popping's structure cannot work without. The need for a thesaurus arises from the expectation to reveal already existing networks which the text entertains with the society in which it is produced in. In the section above, I mentioned the Library of Congress classification as a solution.

Assuming that an exam text consists of the question, but even more so of the correct, endorsed response to the question, without which an exam taker cannot succeed, I operated with a duality of question and response. In a representation of content, I assumed that the exam question is the token, while the response is the type. Together they are subsumed under *representation of choice*, which reflects a semantic network as previously defined, and which required of me to code the exams and answer keys based on categories I found within the LOC guide (Library of Congress, 2021). This method works through references to existing classification systems, a fundamental assumption of the relationship between types and tokens. I

make use of the Library of Congress classification systems (LCC), among which several exist. Batley (2014) summarizes the difference between extant classifications based on a continuum between “enumerative and faceted” (p. 6) types. Enumerative classification schemes such as the Library of Congress classification (LCC), attempt to project the totality of existing subject matter. Batley (2014) listed as advantages, that these schemes have been used around the world and for the longest time, that their notation systems are fairly short, and that they attempt to produce a larger picture. The disadvantages of enumerative schemes are their rigidity, overbearing size, and the fact that they often fail to project newer developments of research within the respective academic subjects. Faceted schemes on the other hand, are more flexible due to a composite approach to the projection of scientific developments. Faceted schemes build up concepts through arecombination of existing ones (Batley, 2014, pp. 9-11).

***Representation of Content Consists of:***

Among others, Batley (2014), Harris (2013) and Kumbhar (2011) established the Library of Congress classification system as one of the most widely used content classification systems in libraries around the world. Besides its prominence, I also found that the LCC currently has the most versatile and inclusive classification structure of Western and Non-Western disciplines and research foci. Questions and responses share the same theme / topic, but only through the response, do mental operations required to succeed in the exercise become understandable. When referring to the response, that ideally means having access to an officially endorsed answer key/ teacher sheet.

The idea itself to create a semantic representation of the exams is inspired Vygotsky’s theoretical approach (Van der Veer and Valsiner, 1994; Vygotsky et al, 1978), which I discussed in section 3 and 4 of Chapter 2. I specifically related what Vygotsky (1978, pp. 54 - 55) termed *artefact mediated cognition*. In brief this means that the correct response to any exam question is

a mechanism of cultural control, which controls the acquisition by the examinees of a sufficient toolset to get to the correct response. The question in itself is then an *artefact* of what the exam's stakeholders, or the society, perceives as the most relevant learning and developmental achievements. Such artefacts are mediated through topics, while responses require certain mental operations to get to the perfect score. Through my analysis of the exam questions and answer keys, I match the semantic representation (*type – token*) with the above theoretical approach and used the LCC system as notational framework (*thesaurus*).

For a better overview I visualized the generated information in a sankey diagram, which also helped me to visualize the semantic network, as demonstrated in the following example. To create this visualization, I used the free software SankeyMATIC (2021):

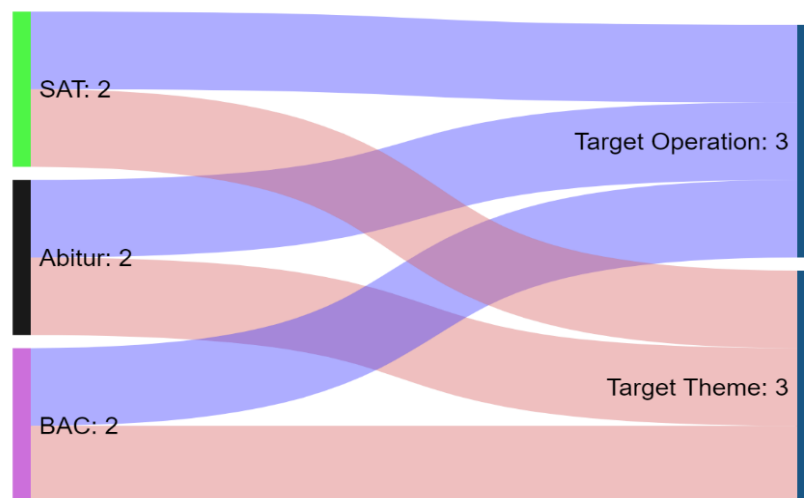


Figure 3.4. Visualization of Exam Content

The diagram in Figure 3.4 should serve to visualize the duality between type and token, the former represented through topics in red, the latter represented through mental operations, in blue. The number indicates the total amount the topic or the operation appeared in either the question sheet or the answer key. This number allows several other inferences, that needed to be better understood. One inference arises from the difference in ratios in comparison across the

three exams for each subject, for which I introduce the *focal concept* of complexity (for “focal concept” see: Carley and Diesner, 1997).

### ***Complexity***

Complexity is a focal concept born from the necessity to express inherent and exherent properties of the exams, for which the previously described analyses could not account. It looks at the ratio between question content (exam sheets) and required answer steps, or *mental operations*. Mental operations are here inspired by Vygotsky’s assumptions on artefact mediated cognition, as priorly laid out. I introduce the ratio here as measurement to shed light on the reasons for the on first glance apparent discrepancy between the lower number of questions contained in the Abitur exam, as opposed to the large quantity of questions contained within the SAT test. All this, while students receive about three times as much time to complete the Abitur exam. It also accounts for some of the limitations of multiple-choice questions, which the SAT exam purely relies on, while the Abitur contains none, exclusively free response questions. Hence the ratio might shed some light on the structural differences in expectations, where a higher number potentially indicates a higher amount of operational complexity.

The complexity I perceive here is not expressed in the questions alone, but created through a relationship between visible and concealed properties of each exam question. Repositioned into the theorem of artifact mediated cognition, a higher complexity score then reveals something about social control. Like an iceberg, much more is hidden under the ocean’s surface, when we have high complexity values. These hidden parts are neither realized nor expressed in the question but require the students to nonetheless apply the correct arsenal of tools in order to acquire the perfect grade.

The next dimension of my analysis aimed at presenting representations of *choice*, where choice relates to significant decisions each exam taker needs to make. Such decisions include the choice of subject, the choice of specialty, or to elect certain subject as being treated differently in the exam.

### ***Representation of choice***

Befitting my wholistic approach to assessment, I am naturally assuming that structure and content play hand in hand and are both an integral part of the exam. So especially for the representation of choice, I operated within the exam text and matched information thereof with information that I deduced from regulatory measures and restrictions pertaining to each exam. For this analysis I looked at structures inherent in the exam, but not directly expressed through a superficial reading of content.

As an inspiration for each dimension of the analysis, I took Carley's (1997) formal properties of *density*, *conductivity*, and *intensity*. In allusion to the formalist approach employed by Carley, I propose *Availability / Degree Specialization*, *Conductivity* and *Answer Choice* as categories for the analysis of exam data. All of the following are ways through which I assume, social capital in its institutionalized form is revealed. As a resource for all the information I require for this analysis, serve the official pages of the French Ministry of Education, (Ministère de l'Éducation Nationale et de la Jeunesse, 2020a), the Ministry of Education of the German state of Lower-Saxony (Niedersächsisches Landesinstitut für schulische Qualitätsentwicklung, 2020) and the official pages of the US American College Board (College Board, 2021).

I came to term this section *choice* with the assumption that all dimensions reflected here have a direct effect on, or root within the idea that students must make choices in order to succeed in the exams. What I term representation of choice is inspired by Bourdieu's (1977, 1986) conceptions of the objectified state of cultural capital. Following a preliminary sighting of

regulatory documents for the Abitur, the Baccalauréat and the SAT, I understood choice as the best descriptor of the information revealed here. I chose to visualize these properties as in the following example:

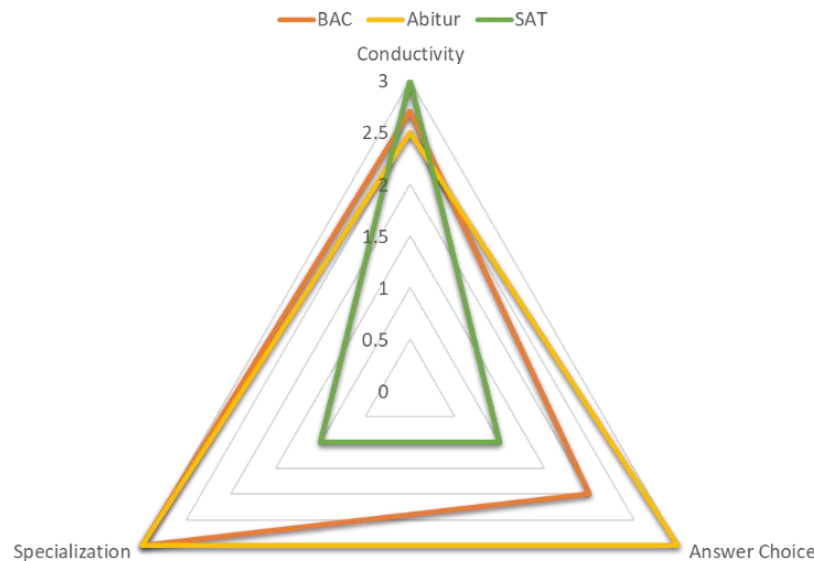


Figure 3.5. Visualization of Representation of Choice

The diagram in Figure 3.5 visualizes the most important dimensions that I found useful to describe “choice” in application to my dataset, per exam subject. The numbers on which each dimension is ranked are categorical numbers. As all three dimensions produced different ordinal values, I categorized the results into sets of three based on the respective range from the smallest to the largest value. This step became necessary to be able to visualize and compare all conceptually linked dimensions of this analysis with one another. I included the tables with the actual ordinal numbers for each subject in the Appendix 3.

Where Carley (1997) uses *Density and Intensity* to quantify the presence of certain concepts in a text, I use *availability and degree of specialization* (or short *specialization* in the diagrams) to describe the presence of a subject in an exam and the degree to which a subject is present in the exam. The availability of subjects alone does not begin to express the difference between the SAT and the Abitur, or the Baccalauréat, as the latter have a complex system of



specialization in place for each subject, which directly reflects through the exam content. Furthermore, do all three exams differentiate in allowing students to leave out certain subjects in favor of others. The dimension represents facets of *choice* pertaining to the intensity of the study of one subject over another. Between the two radicals, leaving a subject out versus including it, the Abitur and the Baccalauréat also allow students to take one subject with a higher or a lower profile, hence to choose early on, what to specialize into. This information is coded through a ranking, which I explain in the following chapter.

For Carley's *Conductivity* I use *Conductivity* and *Answer Choice*. The basic underlying question being, what goes with what? Both *Conductivity* and *Answer choice* are conceptually related, as both determine which exam content is compatible / incompatible (through question grouping or through the *tracking* of subjects). Answer choice is mostly present in the Abitur and Baccalauréat and conceptually groups together questions and the underlying competencies to respond to these questions. Answer choice can also serve to let students further specialize in their approach to completing the exam, as it provides the exam takers with the choice to respond in-depth to one set of questions, while disregarding the other.

For an explanation of *conductivity*, it is especially important to understand the German and French tracked systems, where subjects are categorized into established fields (humanities, science and IT, etc.). For the SAT, the conductivity score is fairly easy to calculate (and its results are not of high significance), as the SAT subject test does not restrict students in terms of subject combinations. It follows that each subject can be combined with every other subject on the menu, which means that each subject of the SAT has the same conductivity score as another. The situation changes with the Abitur and BAC, wherein subjects are categorized according to overarching themes, as for example math belongs into the group of science. If the

plan foresees a required science subject, the student can choose between biology, chemistry, physics or math. As a consequence, a subject which stands alone has a higher value of combinability than a subject that is aggregated in a group of subjects under an overarching theme. I am here referring as *tracking* to the categorization of subjects, and therefore the structural limitation of possible subject combinations.

I use the following formula to calculate the total number of possible combinations:  $nCr = \frac{n!}{r! * (n - r)!}$ , where  $n$  represents the total number of items, and  $r$  represents the number of items being chosen at a time, this means that for each subject aggregate, the subject's value is reduced by 1. As expressed earlier in this section, the corresponding scores in the diagrams are ordinal ranging from 1 to 3. The analysis of conductivity does not fit into the ordinal system, so for the sake of comparison, I needed to categorize the results of conductivity into 3 groups, based on the range of numbers between the smallest and the highest value.

## **Chapter Four. Results**

The following results are presented in an anachronous order, in that the final results precede the minutia of analysis. I chose this design as it appeared catchy and provides an oversight of the research to the impatient reader, as well as the depth of analysis to the so inclined reader. The following sections will describe results from my narrative synthesis across the exams, before going deeper within each exam. The narrative synthesis helped me generate themes, which are theoretically grounded in Bourdieu's (1977, 1986) descriptions of social and cultural capital. When coding, I looked for social and cultural competencies demanded of students by the curricular documents in my sample relating to each exam, the Baccalauréat, the Abitur and the SAT. In coding, while grouping terms with a sufficiently high significance, I was further inspired by Bourdieu's (1977) explanations towards "habitus" (p. 72 ff.), and its being a realization of embodied cultural and social competence.

The latter sections of this chapter are dedicated to the results of the network text analysis, which focused on structures revealed and concealed within the exams, and where the results helped me to respond to Research Question 2, to respond to ways through which cultural capital is embedded within and across the exams and structures relating to the exams.

### **Comparison of Results: Narrative Synthesis**

Narrative synthesis is an analysis of what is said and how it is said, as opposed to network text analysis. An analysis of the total language that is used to describe assessment. – objectified cultural capital (Bourdieu, 1986) in curricular documents. As these documents naturally carry or carried a high degree of authority over the selected exams, I assumed not just their content, but as objects their agreeing to the definitions, which Bourdieu perceived as the foundation of cultural or social capital becoming object.

Though all of the listed themes are included in all literature pertaining to all three exams, the expression and the extent, of the theme's overall significance, differentiate. The tables, which I use in the following sections should serve to make these differences apparent and accessible for further discussion. The score column refers to the total sum of keyness of all term combinations included under a specific theme. Each theme description is taken and referenced from the in depth analyses and descriptions which follow in the sections after the larger comparison, with section titles for each exam. The sum of keyness for each theme helps compare the significance a theme holds in comparison across all three exams. As previously expressed, my analysis operates under the assumption that there exists fundamental social capital (Bourdieu, 1977; 1986), which conditions the production and continued reproduction of certain developmental ideals, that are incorporated into curricula, syllabi and other formative documents pertaining to the three exams. Through my analysis, I aim to discover and spell out some aspects, which make up this capital and the following definitions of synthesized themes.

Coding these themes was a statistically supported, interpretive endeavor into the language of the sampled texts. I was aided in my coding decisions by some of the theoretical frameworks which I presented earlier, specifically the functional categories established by Müller & Schmidt (2009). These categories were especially helpful in locating some of the differences among all three exams.

#### ***Social / Cultural capital: Overlap Among all Three Examinations***

Each table includes a theme that developed across the comparison of texts for each exam, with a simplified, summative description of the individual expression of each theme for each individual exam. First, this format should help provide a better overview, second it should serve to introduce the consecutive sections where each theme is given in-depth focus. The themes unite

the most relevant results of coding the sampled text on cultural competencies, as well as social competencies required of exam takers.

The theme ***Production, Interpretation and Manipulation of Text*** has the highest representation within texts linked to the Baccalauréat, with a sum of scores more than double of those of the other two exams. It is followed by the Abitur and SAT uniting scores with a sum of less than half of the former.

Table 4.1. Summary of narrative synthesis: Production, Interpretation and Manipulation of Texts

<i><b>Exam</b></i>	<i><b>Production, Interpretation and Manipulation of Text</b></i>	<i><b>Score</b></i>
BAC	<p><b>A literary thesis imitates great classics</b></p> <p>Rhetoric mastery is at the heart of literary theses, and is honed through in depth knowledge of great classics, as well as through their imitation.</p> <p>Through a dictate, students reveal their mastery of the French language, as well as insight into the ingenuity of the text.</p>	1143
Abitur	<p><b>Aesthetic insight into form is related to the mastery of the German language</b></p> <p>The eloquence of the use of the German language is of importance, as well as philosophically and philologically precious prose of global format with a special focus on Greek classics.</p> <p>Understanding the Aesthetic as a unique and special form of perception, creation and awareness.</p>	598
SAT	<p><b>The mastery of text is related to social and political fluency of the students</b></p> <p>Students express their mastery over text through the application of skills relating to rhetorics and argument structure.</p> <p>Mastery of text implies a close study and in which reflection of political and social events in the society the students live.</p>	419

Historical and contemporary texts pertaining to the Baccalauréat majorly stressed the importance of a student's ability to write a literary thesis. Such was often seen as the expression of mastery of subject content, but as well mastery of the French language. Older publications often stressed a need for students to imitate literary classics of Greek, Latin and European Renaissance origin to express such mastery. Historical and contemporary texts pertaining to the Abitur also focused on the eloquent application of the German language, but different in what they perceived as measure of mastery. Understanding and analysis of aesthetics often found an expression in the curricular documents, when the mastery of textual production and manipulation was concerned. Such aesthetic understanding was historically (in older texts of the sample) closely linked to the understanding and appreciation of ancient Greek drama and philosophy and to the understanding of text as basis for dramatic performance.

Historical and contemporary texts pertaining to the SAT Test took quite a different stance on the theme. What I found could best be described as a mentalist approach with focus on technicalities, where students are often expected to construct an argument based on an understanding of the author's positioning and rhetoric devices used. The latter was often embedded in definitions of critical thought on contemporary political developments.

The following theme, which I call *Communicative Competence*, is again vastly represented in documents related to the Baccalauréat with a summative score more than double that of each of the other two exams.

Table 4.2. Summary of Narrative Synthesis: Communicative Competence

<i>Exam</i>	<i>Communicative Competence</i>	<i>Score</i>
BAC	<p><b>Mastery of foreign languages and mastery of French</b></p> <p>Modern foreign languages carry a importance similar to that of classic languages. Thorough and age appropriate foreign language study is in the interest of a vital citizenship.</p> <p>The importance of modern foreign languages, held an importance at least equal to the mastery of the French and ancient classic language.</p>	1610
Abitur	<p><b>Mastery of foreign languages, stressing romance languages and cultural fluency</b></p> <p>The ability to orally translate from German into Greek, whereas special attention is given to etymology, proper syntactical sequence, and correct grammar overall.</p> <p>An exceptional position that the romance languages have always had in the German Abitur.</p> <p>Communicative competence as a means to bridge intercultural difference. Communication in a foreign language needs to be embedded in a sociocultural understanding of the culture of speakers of that language.</p>	544
SAT	<p><b>Communication in English is a strategic tool for success</b></p> <p>Word choice and the expression of meaning needs to be accompanied by an understanding of the speaker's situation and proper tone.</p> <p>The written language inspires and helps develop mastery of the spoken language. Standard written English.</p> <p>Mastery of communication strategies is crucial for academic and career readiness</p>	417

While curricular documents relating to both the Baccalauréat and the Abitur expressed mastery of foreign languages as crucial for achieving Communicative Competence, the SAT documents focused on the strategic use of the English language.

While documents relating to the Baccalauréat stressed the acquisition of a diverse selection of modern and classic foreign languages, as being similarly important for communicative competence, documents relating to the Abitur stressed the importance of classic philologies and the romance languages. Some passages in documents relating to the Baccalauréat stressed building communicative competence in younger generations, as being directly related to building a strong French citizenship. On the other hand, documents relating to the Abitur often laid out the importance of acquiring cultural understanding in tandem with language skill, as important for communicative competence.

Documents relating to the SAT stressed the importance of the mastery of Standard English for communicative competence, with expressions on the mastery of communication being directly linked to the academic and professional success of students.

The following theme, which I call ***Modes of Reasoning***, is vastly represented in documents related to the SAT with a summative score more than double that of each of the other two exams.

Table 4.3. Summary of narrative synthesis: Modes of Reasoning

<i><b>Exam</b></i>	<i><b>Modes of Reasoning</b></i>	<i><b>Score</b></i>
BAC	<p><b>Understanding lived realities leads to good choices</b></p> <p>Students' final examinations coherently reflect their individual educational track and subject choice. Their methods display refinement in their respective specializations.</p> <p>Programs are designed to form students intellect and integrity, which in turn helps them to best understand their lived realities.</p>	482

(table cont'd.)



<i>Exam</i>	<i>Modes of Reasoning</i>	<i>Score</i>
Abitur	<p><b>Mathematical reasoning is at the forefront</b></p> <p>Mathematical reasoning is of utmost importance, especially when comprehending the natural properties of the world that surrounds us. Aside analytical and applied geometry, a proper historico- developmental framework for math within scientific endeavors is also of the essence.</p>	424
SAT	<p><b>Critical reading and overview matter</b></p> <p>An indicator for development of reasoning is the student coping with rigors of reading in college. Critical reading of texts.</p> <p>Reading comprehension and mathematical reasoning are two important poles, which often contradict each other.</p> <p>Reasoning in science entails knowing the relationships, through which ideas in different fields are linked with one another.</p>	1701

In terms of the ideas included in the theme *Modes of Reasoning*, statements in curricular documents among all three exams vary to some extent. Documents relating to the Baccalauréat often stress the ability to make informed decisions, which are anchored in a good perception of the lived realities in the French society. Intellectual development comes hand in hand with the development of active citizenship, through which the students are able to know themselves and their place in society.

Curricular documents relating to the Abitur significantly stress Math reasoning and the development of mathematical skills as being important for the students in subjects beyond Mathematics. On the other hand, documents relating to the SAT stress reading comprehension of college level texts and science literacy as most important for this theme. There are some passages, which also set mathematical reasoning and literacy development as opposites on a scale. The following section will focus on themes, for which I only discovered sufficient

evidence among documents relating to the Baccalauréat and the Abitur, but not in the documents relating to the SAT.

***Social / Cultural capital: partial overlap between two exams***

Similar to the previous section, the following table indicates the name of the theme with a descriptive summary of statements I found for the theme across documents relating to either the Abitur or the Baccalauréat. As the table is shorter, I decided to include both themes which fall under this section in one table and discuss each below the table.

Table 4.4. Summary of Narrative Synthesis: Themes with Partial Overlap

<i><b>Exam</b></i>	<i><b>Oral Mastery</b></i>	<i><b>Score</b></i>
BAC	<p><b>Oral mastery of subject matter is a sign of professionalism and specialization</b></p> <p>An oral exam is an integral part of the students' performance in the subjects in which they specialize.  The oral mastery of subject matter proves sound reflection and in depth understanding of the subject.  Being fluent in a subject indicates on the part of the students' professionalism.</p>	732
Abitur	<p><b>Oral mastery expresses in-depth knowledge and is rooted in a written component</b></p> <p>The oral exam is developed as a defense of the written performance of students.  The oral exam needs to focus on the student's most accomplished subject matter, therefore it is the result of continuous, in-depth specialization.</p>	253

(table cont'd.)

<i><b>Exam</b></i>	<i><b>Classic Competence</b></i>	<i><b>Score</b></i>
BAC	<p><b>The study of old languages reflects and enriches the modern</b></p> <p>Classic Studies are based in a humanist conception of education; they help develop a combination of memory, imagination, spirit, and soul.</p> <p>Classic competence has to be developed through the application of old languages to modern topics.</p> <p>The study of old languages is a means to come to a deeper understanding of linguistic structures, which underlie those languages, modern French and other European languages.</p>	1143
<i><b>Exam</b></i>	<i><b>Classic Competence</b></i>	<i><b>Score</b></i>
Abitur	<p><b>Classic competence is linked to theology and the European Renaissance.</b></p> <p>Classic competence is traditionally the mastery of ancient Greek and Hebrew, as it enables a higher education in Christian theology.</p> <p>Classic competence is developed by the students over lengthy studies of ancient Greek and Renaissance writing.</p>	304

I identified two themes in this category. One theme being ***Oral Mastery*** as linked to the concept that any exam needs to include an oral component, which has until today been the case for the Abitur and the Baccalauréat. Reading and coding the documents relating to both exams, I understood that the reasoning for the inclusion of an oral element is very similar for the stakeholders of both exams. Documents relating to both exams express the idea that an oral exam supplements the written exam in a subject, whereby the students are required to perform exceptionally well in one of the subjects in which they specialized. Therefore, the idea of mastery of subject matter appears to be directly linked to both a written and a structured oral performance about the content knowledge.

The second theme in this category is *Classic Competence*, which relates to the knowledge in and about ancient languages, more specifically ancient Greek, Latin and Hebrew. The expression of the theme is slightly different between the documents of both exams. French curricular documents often express the concept that the acquisition of ancient language helps students to a deeper understanding of linguistic structures of modern European languages (and of their native language, French). Therefore the acquisition of both, modern and ancient languages, is considered important to for the overall linguistic development of students. The approach to ancient languages is set in a comparative framework, which one source termed as inspired by the humanist belief in a wholistic development of students, as old language studies provide perspective on contemporary languages.

On the other hand, documents relating to the Abitur often expressed the importance of the study of ancient languages for the reading and translation of theological works. *Classic competence* often appears to be embedded in a framework of religious studies (Christian, Jewish) which are subjects offered uniquely by the Abitur among the three exams. Beyond that, the study of old Greek is often linked to an assumed importance in the documents, of the European Renaissance and its writings.

### ***Social / Cultural Capital: Themes Unique to Each Exam***

The following section discusses themes, which appear solely in one exam, but not the others. I treat each of these themes separately for each exam and unlike before, it was necessary to provide some reflection on the translation of each of these themes. Some themes cannot be directly translated, which requires that I should define them through the descriptions I provide.

I found one unique theme for the Baccalauréat, which I termed *Culture and Morals*, and which unites several assumptions on behavioral objectives of successful exam takers.

Table 4.5. Summary of Narrative Synthesis: Culture and Morals

<i><b>BAC</b></i>	<i><b>Culture and Morals</b></i>	<i><b>Score</b></i>
	<p>Scholarly culture consists of historically evolved fundamental concepts.</p> <p>Scholarly communities have the authority to decide upon their fundamentals, and it is the aim of the student to appropriate these fundamentals through intellectual growth.</p> <p>Through this growth, the students refine their intellect, moral decision making, and build social integrity. Scholarly culture takes into account scientific reasoning, as well as humanism.</p>	649

Table 4.5 gives a concise description of some of the concepts, which I identified belonging to this category. Many documents relating to the Baccalauréat stated assumptions of fundamental values, which the students require for intellectual growth. One concept would be the existence of a scholarly culture. Successful exam takers would have acquired the necessary acculturation, to become working members of a scholarly community in subsequent higher education. This process of acculturation reminded me of what Bourdieu (1986) described as habitus, since such acculturation measures the extent to which the students can operate and relate to the French society.

I found two unique themes with documents relating to the Abitur, both similarly significant and both describing behavioral objectives, similar to the Baccalauréat theme above.

Table 4.6. Summary of Narrative Synthesis: *Sitte, Reife*

<i>Abitur</i>	<i>Sitte, Mental and Cultural Ideas</i>	<i>Score</i>
	<p>Moral leadership is valued as much as academic excellence and needs to be accessed before any student can be admitted into higher education.</p> <p>Work attitude is as important as the knowledge of proper academic methods to succeed and graduate from upper secondary school.</p> <p>Aside the ability to conduct higher studies, students receive a propaedeutic education, especially in sight of the applicability of their knowledge within practical scientific contexts.</p> <p>Propaedeutic is the understanding of philosophical and anthropological questions of essence and existence. Ethics here stem from Greek classics, as well as from Christian doctrine.</p>	475
<i>Abitur</i>	<i>Reife</i>	<i>Score</i>
	<p>Academic mastery springs from the continued long-term endeavor to acquire knowledge. A certificate of maturation (historical name for Abitur, wording still present in Austria) is not given to anyone displaying a motly assortment acquired in forced and hasty studies.</p> <p>Higher education maturity should be distinguished from a social, cultural maturation of the individual within society. Previously used nearly synonymously, the relationship these terms share is marked through commonalities, as well as through differences. Future debates will have to succeed in that distinction.</p>	559

The first theme as stated in Table 4.6 is *Sitte*, which I left in German in lack for a better translation and described it instead as a mental and cultural concept, native to the German language and society. The documents linked to the exam, which expressed this term are mostly historical, with mentions until the 1970s, and the term is not very common in modern German anymore. What made me include it despite its being seldomly used, is the fact that its descriptions persist in modern curricular documents, but with different descriptors. *Sitte* often relates to a type of moral leadership, which is deemed valuable for the success in the Abitur and in subsequent higher education. Further descriptions in the Abitur documents entail mentions

about the students' ability to reflect on philosophical questions of essence and existence as related to schools of thought of Greek antiquity and Christian doctrine.

The second theme I left untranslated for similar reasons, is named ***Reife***. To some extent *Reife* means “maturity”, but the concepts which make up a better definition of *Reife* are not included in the English definition of the word. Curricular documents mentioned the term often in relationship to an extended period of in-depth study and careful consideration of the subject matter and of the subject in relation to society. Other passages often expressed the idea that exam readiness for the Abitur (or *Maturat*, as it is still called in Austria until today), is not the result of inherent academic talent for the subject or excellent scholarly achievements, but is the result of an academic “maturity” meaning the outcome of long term academic study and maturation of students on a personal level.

Following are three themes, which I found uniquely represented in the SAT. The last two are closely related, and although separate, they unite high amounts of keyness and many term combinations, which made me consider to keep them separate during the coding.

The first theme is ***Readiness***, which has the highest keyness score among all unique themes in this section.

Table 4.7. Summary of Narrative Synthesis: Readiness

<b><i>SAT</i></b>	<b><i>Readiness</i></b>	<b><i>Score</i></b>
	<p>College Readiness can be examined through the SAT score. Determinants for a high score are four years of English and three years each of mathematics, science, and social science/history.</p> <p>The SAT is designed to collect evidence from student performance in support of a broad claim about what students know and can do.</p> <p>At its core, the exam is a postsecondary admission and diagnostic exam, and as such is designed to be a strong predictor of postsecondary success.</p>	997

Table 4.7 provides a summary of some of the descriptions I found in the curricular documents on Readiness. The theme describes a functional property that is accorded to the SAT by its stakeholders. If devised correctly, the exam is assumed to have the ability to predict college, higher education success through the exam outcome of the exam takers. The inclusion of this theme as a social or cultural competence is reasonable, as the stakeholders of the SAT hold the assumption that certain learned dispositions of exam takers are at the basis of predicting higher education success. In some curricular documents such is stated as a certain exposition to English, Math, Science and Social Studies or History education.

The following themes are conceptually related, as they both express embodied social properties exam takers are required to have in order to succeed.

Table 4.8. Summary of Narrative Synthesis: Social Factors for the SAT

<b>SAT</b>	<b><i>Alignment with Authority, Degree of Social Assimilation</i></b>	<b><i>Score</i></b>
	<p>The SAT measures skills, which are highly relevant to the working of the current society and postsecondary achievement of its exam takers.</p> <p>The SAT is able to measure the amount of academic rigor displayed by an exam taker through its academic rigor index (ARI).</p>	659
<b>SAT</b>	<b><i>Social Factors of Learning / Social Stratification</i></b>	<b><i>Score</i></b>
	<p>Academic rigor is related to social upbringing and parental income of the exam takers.</p> <p>Success in the SAT often reflects parental income, social class and whether or not students had the opportunity to partake in extracurricular activities.</p>	659

Table 4.8 first summarizes the theme I call ***Alignment with Authority, Degree of Social Assimilation***. The idea often expressed in curricular documents on the SAT is that the exam is able to measure academic rigor in relevance to postsecondary demands. If this coding result is put into a framework of social and cultural competence, the assumption dawns, that students who



scored high in the SAT have acquired competencies, which are considered highly relevant by social stakeholders in higher education or the workforce.

The second theme described in Table 4.8 is interesting, as it is drawn from a criticism on exam outcomes, but even more so relevant for my analysis. I called the theme *Social Factors of Learning / Social Stratification*, and like the previous theme, it is the assumed embodiment of social properties through students' competences. Curricular documents link the assumption of academic rigor with student socioeconomic upbringing and the students' previous access to extracurricular activities. If formulated as a competence, this theme would best be described as the embodiment of parental income and socioeconomic class. This theme is also the result of a critical perspective by the stakeholders on the SAT, which I encountered during coding. Such was seemingly absent in curricular documents relating to the other two exams.

### **Results: Narrative Synthesis per Exam**

The following sections focus on the minutia pertaining to the summaries, which I presented above, and quote some excerpts from the curricular documents for each theme presented. I also briefly complete the discussion on my sampling strategies, which I began in Chapter 3. The reason for including these sampling steps is that I already relied on several analytical steps to get to the final sample size from which the themes were eventually drawn. The rigid stratification and the systematic reduction of the sample size helped me to introduce further rigor in my analysis given the fact that I cannot provide for inter-coder reliability as is often done in the sampling for qualitative meta-synthesis. It is debatable to what degree sampling already involves analysis. Analytical steps were needed to reduce the sample size, and those were de facto already inspired by my theoretical framework, more precisely assumptions on conceptual proximity of certain texts to the exams over others. Figure 4.1 serves as a reminder of the sample

size and the text types which came to be included in each subsample for each exam respectively, after I systematically reduced the sample size as laid out in the previous chapter.

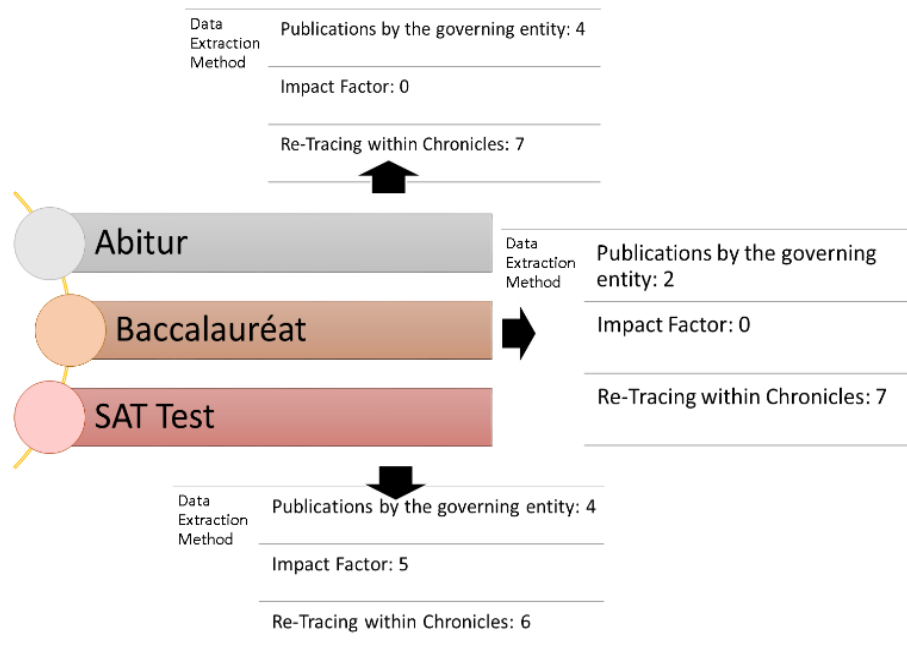


Figure 4.1. Narrative Synthesis: Sample Size for Each Exam

There are significant differences among the types of texts included for each exam and the numbers of inclusions in each category. These differences served as a point of reflection later, when I discuss my findings. Despite their categorically different sample make-ups, all included texts for each exam fit the defined rationale, meaning that I assume that they are similarly significant for each exam in terms of the factors dominion and impact. Based on its timeline of publication the sample is stratified as follows:

Table 4.9. Narrative Synthesis: Reference Table for Sampled Texts

Year	Abitur	Baccalauréat	SAT Exam
1788	Edict vom 23. December 1788 über die Prüfung der auf die Universität gehenden Schüler.		
1812	Einführung des Abiturs in Preußen: Edikt unterzeichnet von König Friedrich Wilhelm III., Staatskanzler Hardenberg und Friedrich von Schuckmann.		
1834	Prüfungsordnung Deutscher Bund.		
1840		1. Circulaire relative à l'exécution des arrêtés du 2 octobre 1840 2. Instruction relative à l'exécution de l'arrêté du 14 juillet 1840 3. Instruction relative à l'exécution de l'arrêté du 25 août 1840	
1874		Circulaire relative à l'application du nouveau plan d'études des lycées, prescrit par arrêté du 23 juillet 1874.	
1902		Programme des examens du Baccalauréat de l'enseignement secondaire. Décret et Arrêté du 31 Mai 1902.	
1919			College Board: Nineteenth Annual Report of the Secretary.
1925			College Board: Twenty-fifth Annual Report of the Secretary.
1926	Neue Reifeprüfungsordnung Weimar.		
1927		Réorganisation du baccalauréat de renseignement secondaire. Rapport au Président de la République.	College Board: Twenty Seventh Annual Report of the Secretary.
1932			College Board: Thirty-Second Annual Report of the Secretary.
1933			College Board: Thirty-Third Annual Report of the Secretary.

1958	Tutzinger Gespräche I		
1959	Tutzinger Gespräche II		
1960	Tutzinger Gespräche III		
1972 - 2018	Vereinbarung zur Gestaltung der gymnasialen Oberstufe und der Abiturprüfung.		
1974			Criterion-Referencing, Norm-Referencing, and the SAT.
1977			The Predictive Value of the SAT.
2006			SAT Standards: 1. English 2. Math 3. Sciences
2007			A Historical View of Subgroup Performance differences
2011			College Readiness Benchmark
2012	Bildungsstandards: 1. Fremdsprache 2. Deutsch 3. Mathe		
2014		Charte relative à l'élaboration, à la mise en œuvre et au suivi des programmes d'enseignement ainsi qu'aux modalités d'évaluation des élèves dans l'enseignement scolaire	
2015			Test Specifications for the Redesigned SAT.
2018		Note d'analyses et de propositions sur les programmes du lycée et sur les épreuves du baccalauréat.	

Blue – From Chronicles

Black – Impact Factor Research

Red – From Governing Entity

First, it must be noted, that some of the texts I included do not have APA conventions available, especially if published by a long gone monarchichal government and administration in the case of Germany and France. When citing or referring to these documents and consequently (for uniformity's sake) also to the newer entries of Table 4.9, I just mentioned the year and the

title of the publication. Table 4.9 serves as a reference aid to locate the texts. The complete table with an APA reference or, if not available, a source description is available in Appendix 1.

While each row is reserved for publications linked to a single exam, Table 4.9 serves to provide a visual overview of the timely distribution of publications across the three exams. Furthermore, the colors help visualize the distribution of the three categories of publications collected in the context of each exam. The line, which divides the table into two distinct parts represents a practical distinction I came to make during the data analysis; it serves to distinguish between texts I considered *modern* and texts I considered *historical*. While each exam today, most likely, carries expressions and elements of its genesis over time within, it remains to be proven to what extent that holds true. While listing the results of my data analysis in the following subchapters, I therefore asked for each identified theme, whether its textual evidence aligns with both modern and historical publications alike.

A close reading of the collected 100 word pairs with the strongest keyness score per exam in context of the coding rationale, permitted me to exclude some pairs and include others for further analysis through their location within and across texts. As a reminder, the coding rationale was dictated by reflecting on the assumptions I underlaid definitions of social and cultural capital.

### **Findings for the Abitur**

Each of the following sections is dedicated to one exam, wherein I present the resulting themes of the inductive analysis, followed by a description of the theme based on its included key terms with their total keyness value. The tables thereby represent the results of inductive coding with the help of Sketchengine. The summaries, which follow immediately after, are text passages located surrounding one or more of the key terms listed, which I found to contribute

significant and substantive information to delineate each theme. Depending on the distribution of significant information, I choose to manually search the sampled texts for passages. I determined it necessary to add so collected evidence, when a theme turned out to be one-sidedly represented by evidence drawn from either the *modern* or the *historical* portion of the sampled texts.

The translation for each foreign language term combination is provided in the table. For citations from the source material, I provide an indirect translation directly under the quote, which captures the most important information, while providing a summary and positioning the information in regards to the overarching theme and with regard to the my coding rationale.

### ***Communicative Competence***

*Communicative Competence* as a theme includes terms with a combined keyness value of 544, one or more of the included term combinations are represented in the modern publications. The theme captures information from 64 % of the publications, where one or more of the term combinations are mentioned.

Table 4.10. Abitur: Communicative Competence

<b><i>Communicative Competence</i></b>	<b><i>Translation</i></b>	<b><i>64 % / yes / keyness:</i></b>
<i>kommunikative Kompetenz</i>	communicative competence	109.330
<i>funktionale kommunikative Kompetenz</i>	functional communicative competence	75.360
<i>sprachliches Mittel</i>	linguistic means	64.300
<i>interkulturelle kommunikative Kompetenz</i>	intercultural communicative competence	58.840
<i>kommunikative Strategie</i>	communicative strategy	57.400
<i>sprachliche Handlung</i>	speech act	33.910

(table cont'd.)

<i>Communicative Competence</i>	<i>Translation</i>	<b>64 % / yes / keyness:</b>
<i>zweite Fremdsprache</i>	second foreign language	78.110
<i>fortgeführte Fremdsprache</i>	continued foreign language	41.280
<i>kurze Übersetzung</i>	short translation	25.790

All terms included in this category pertain to contexts of communicating in a foreign language. The term combination *kurze Übersetzung* (short translation) represents the oldest inclusion:

*(...) einer kurzen Uebersetzung aus dem Deutschen ins Griechische, wobei etymologische und syntaktische, und überhaupt grammatische Richtigkeit in jeder Hinsicht, in Betracht kommen. (1812, Edict)*

Drawn from the 1812 Edict, translation describes the ability to orally translate from German into Greek, whereas special attention is given to etymology, proper syntactical sequence and correct grammar overall.

Another important passage pertains to the keywords *zweite Fremdsprache* and is drawn from 1959, *Tutzingen Gespräche II*:

*Liege allerdings die gründliche Schulung in einer romanischen Sprache vor, so könne man allein in diesem (eher seltenen) Fall die Streichung der zweiten Fremdsprache – wenn auch mit Bedenken – erwägen; propagieren solle man sie jedenfalls nicht. (1959, Tutzingen Gespräche II)*

The passage refers to the fact, that a second foreign language has been a standard to each student's Abitur and describes a peculiar exception to that commonality. If a student opts to specialize in a Romance language, it might be considered to except the student from having to enter a second foreign language track. With Research Question 1 in mind in the context of the prior passage, it points towards an exceptional position that the romance languages have always

had in the German Abitur.

Another outstanding aspect of this theme is the function of communicative competence as a means to bridge intercultural difference. The key term intercultural competence speaks to that, while the passage revealed by the key term communicative strategy (*kommunikative Strategie*), elaborates on the ideals connected to that:

*(...) entsprechendekommunikative Strategien einsetzen, um adressatenrelevante Inhalte und Absichten in der jeweils anderen Sprache zu vermitteln bei der Vermittlung von Informationen gegebenenfalls auf Nachfragen eingehen Inhalte unter Nutzung von Hilfsmitteln, wie z. B. Wörterbüchern, durch Kompensationsstrategien, wie z. B. Paraphrasieren, und gegebenenfalls Nutzung von Gestik und Mimik adressatengerecht und situationsangemessen sinngemäß übertragen Erhöhtes Niveau Die Schülerinnen und Schüler können darüber hinaus für das Verstehen erforderliche Erläuterungen hinzufügen zur Bewältigung der Sprachmittlung kreativ mit den beteiligten Sprachen umgehen. (2012, Bildungsstandards Fremdsprache)*

Following the quote above, communication in a foreign language needs to be relevant to the situation of the listener, first and foremost. The speakers need to be able to comprehensively communicate content and their conversational intention towards the situation. In order to be able to pass across both, students are ideally required creativity, proper compensation mechanisms for lacking ideal foreign language expressions, that can include, but are not limited to mimicking and gesturing. This modern passage clearly shows a progression from the ideals mentioned in the 1812 Edict, but instead of invalidating the historical ideals cited therein, it adds several layers of complexity and nuance.

### ***Production, Interpretation and Manipulation of Text***

*Production, Interpretation and Manipulation of Text* as a theme includes terms with a combined keyness value of 598; one or more of the included term combinations are represented in the modern publications. The theme captures information from 82 % of the publications, where one or more of the term combinations are mentioned.

Table 4.11. Abitur: Production, Interpretation and Manipulation of Text

<b><i>Production, Interpretation and Manipulation of Text</i></b>	<b><i>Translation</i></b>	<b><i>82 % / yes / Keyness:</i></b>
<i>schriftliche Arbeit</i>	written exam	140.450
<i>pragmatischer Text</i>	pragmatic text	91.520
<i>deutscher Aufsatz</i>	German essay	50.370
<i>literarischer Text</i>	literary texts	46.720
<i>schriftliche Prüfungsarbeit</i>	examined writing	40.920
<i>schriftliche Abiturprüfung</i>	written Abitur exam	40.570
<i>eigene Textproduktion</i>	production of proper text	33.780
<i>komplexer Text</i>	complex text	32.240
<i>Text Unterschiedlicher Medialer Form</i>	texts of different medial form	25.790
<i>bedeutende Schrift</i>	significant writings	25.580
<i>mediale Form</i>	medial form	44.640
<i>mittelschwere Prosa</i>	medium difficulty prose	25.790

An example for the theme is provided in a historical publication on the word combination

*German essay:*

*Die schriftlichen Prüfungsarbeiten bestehen: 1. In einem deutschen Aufsätze , welcher vorzüglich die Bildung des Verstandes und der Phantasie beurkunden soll, wie auch, in seiner Abfassung, die Kenntniß der deutschen Sprache und die Gewandtheit in deren Gebrauch. (1812 Edict Abiturreglement)*



The eloquence of the use of German language is one of the key measures, which allows a successful exam taker to display imagination, while at the same time being coherent.

Another example is a collocation of the word pair *significant writings* and the quote stems from a 1960s publication on the conception of the Abitur:

*Verständnis einiger Meisterwerke der deutschen Literatur, und zwar audi soldier aus dem Umkreis der klassischen Literaturepodie sowie bedeutender Schriften sowohl philosophisch als auch literarisch wertvoller Prosa und Verständnis einiger 155 grundlegend wichtiger Meisterwerke der Weltliteratur, vor allem auch der Antiken. (1960, Tutzinger Gespräche III)*

Significant writings consist of Greek classics and derive value from philosophical depth, philological intricacy and a global recognition. The quote above identifies 155 foundational masterpieces of prose of global format, with a focus on Greek antiquity.

Another pair of rather generic terms; *literary texts*, I found to have a high keyness score and therefore merited further definition through a related passage. The quote below is taken from a modern curricular document.

*Sich mit literarischen Texten auseinandersetzen Die Schülerinnen und Schüler erschließen sich literarische Texte von der Aufklärung bis zur Gegenwart und verstehen das Ästhetische als eine spezifische Weise der Wahrnehmung, der Gestaltung und der Erkenntnis. (2012 Bildungsstandards Deutsch)*

The quote expresses what students require to analyze literary texts. An understanding of aesthetic properties is mentioned. The skill to recognize such is proof of a desired, unique, and special form of perception, necessary for the successful creation of literary texts.

### ***Modes of Reasoning***

*Modes of Reasoning* as a theme relating to the Abitur can be more precisely summarized as *Mathematical / Scientific Reasoning*. It includes terms with a combined keyness value of 424; one or more of the included term combinations are represented in the modern publications. The

theme captures information from 73 % of the publications, where one or more of the term combinations are mentioned.

Table 4.12. Abitur: Mathematical, Scientific Reasoning

<b><i>Mathematical, Scientific Reasoning</i></b>	<b><i>Translation</i></b>	<b><i>73 % / yes / Keyness:</i></b>
<i>mathematische Kompetenz</i>	mathematical competence	75.810
<i>allgemeine mathematische Kompetenz</i>	general mathematical competence	75.360
<i>analytische Geometrie</i>	analytical geometry	47.130
<i>bezogenes mathematisches Sachgebiet</i>	related mathematical subject matter	34.050
<i>mathematische Arbeit</i>	mathematical work	33.650
<i>mathematisches Modell</i>	mathematical model	29.120
<i>Kenntnis der Elementarmathematik</i>	knowledge of elementary math	25.790
<i>stochastische Situation</i>	stochastic situation	25.790
<i>mathematisch-naturwissenschaftlich-technisches Aufgabenfeld</i>	mathematical scientific subject matter	25.680
<i>ebene Trigonometrie</i>	flat trigonometry	25.580
<i>anschauliche Natur</i>	observable nature	25.680

An interesting term combination which appeared with significant keyness is *flat trigonometry*, which appears in a historical curricular document:

*(...) in der Mathematik wird erfordert Kenntniß der Rechnungen des gemeinen Lebens nach ihren auf die Proportions-Lehre gegründeten Principien, des Algorithmus der Buchstaben, der ersten Lehre von den Potenzen und Wurzeln, der Gleichungen des ersten*

*und zweiten Grades, der Logarithmen, der Elementar-Geometrie (so weit sie in den sechs ersten und dem 11ten und 12ten Buche des Euklides vorgetragen wird), derebenen Trigonometrie und des Gebrauchs der mathematischen Tafeln (sic). (1812, Edict Abiturreglement)*

Mathematical reasoning pertains to operations considered most relevant for the everyday life of the exam takers. Mentions here would today be subsumed under the mathematical fields of Trigonometry and Geometry. The quote below is taken from a modern publication and relates to the word combination *analytical geometry*.

*Die inhaltsbezogenen Kompetenzen werden jeweils übergreifenden Leitideen zugeordnet, die nicht auf bestimmte klassische mathematische Sachgebiete (Analysis, Lineare Algebra/Analytische Geometrie, Stochastik) begrenzt sind. (2006, Bildungsstandards Mathe)*

The quote makes general assumptions on mathematical reasoning. Mathematical reasoning is of utmost importance, especially when comprehending the natural properties of the world that surrounds us. Aside analytical and applied geometry, a proper historico-developmental framework for math within scientific endeavors is also of the essence.

### ***Oral Mastery***

*Oral Mastery* as a theme includes terms with a combined keyness value of 253, one or more of the included term combinations are represented in the modern publications. The theme captures information from 55 % of the publications, wherein one or more of the term combinations are mentioned. This category has the least key terms of all which have relatively high keyness values. It also still manages to reach a similar source coverage, compared to other categories, which speaks for its significance.

Table 4.13. Abitur: Oral Mastery

<i>Oral Mastery</i>	<i>Translation</i>	<i>55 % / yes / Keyness:</i>
<i>mündliche Prüfung</i>	oral exam	176.370
<i>mündliche Prü-fung</i>	oral exam	25.790
<i>mündliche Prüfungen</i>	oral exams	25.790
<i>Gegenstand der Mündlichen Prüfung</i>	substance of the oral exam	25.680
<i>die mündliche Prüfung</i>	oral exam (with article)	149.72

*Oral exam* appears in abundance in the curricular texts, which points towards definitions of oral mastery as social and cultural competence, that require further dissemination in its earlier mentions. *Oral exam* appears to have similarities with the oral defense of a piece previously written by the student in its earliest mentions:

*2. Die Prüfung geschieht theils mündlich , theils schriftlich ; doch muss die letztere einige Tage vor der mündlichen erfolgen , damit die Prüfungs arbeiten bei der mündlichen Prüfung zum Grunde gelegt werden können. (sic) (1788 Erstes Abiturreglement)*

The exam is in equal parts held orally and in writing; the written exam needs to happen take place several days before the oral, so the examiners can base the oral evaluation of the student on their prior written performance. In this curricular document, the oral expression is logically defined as an elaboration on a previous written exam. Another historical passage from the time of the German Weimar Republic lays a focus on the *oral exam catering* to the individual strengths of the student.

1. Die mündliche Prüfung kann ich grundsätzlich auf alle Unterrichtsfächer sowie die Praktischen Übungen und die Freien Arbeitsgemeinschaften erstrecken, an denen der Prüfling in der Prima teilgenommen hat.
2. Der Prüfling darf sich ein Fach auswählen, in dem er eine besondere Leitungsfähigkeit nachweisen will (vgl. 8 6, 4). Der Prüfungsausschuß bestimmt im Anschluß an die Vorschläge gemäß 8 17, 1, ob und in welchen Fächern der Schüler außer dem Fache einer Wahl noch geprüft werden soll. Dabei wird es weit zweckmäßiger sein, den Schüler in den Fächern zu prüfen, in denen er seine eigentümlichen Kräfte entfalten kann, als in solchen, in denen er geringere Leistungen auszuweisen hat. (sic) (1926, Neue Reifeprüfungsordnung Weimar, 289)

The quote above states that the oral exam can potentially include all subjects the students elected. It can draw from all practical exercises, as well as from voluntary consortia, in which the students participated during their secondary school years. The idea that the students have the freedom to choose is strongly expressed here. As such, the examinee is free to choose a subject in which they have to be tested in depth, in special capacity. The passage stresses that such should be subjects in which the student managed to express a natural talent. It is the purpose of this examination to focus on the students' strengths and not reveal their weaknesses.

### ***Classic Competence***

*Classic Competence* as a theme includes terms with a combined keyness value of 304; one or more of the included term combinations are represented in the modern publications. The theme captures information from 55 % of the publications, where one or more of the term combinations are mentioned.

Table 4.14. Abitur: Classic Competence

<b><i>Classic Competence</i></b>	<b><i>Translation</i></b>	<b><i>55 % / yes / Keyness:</i></b>
<i>alte Sprache</i>	ancient language	45.790
<i>lateinischer Aufsatz</i>	Latin essay	42.310
<i>griechisches Scriptum</i>	Greek script	34.050
<i>griechischer Schriftsteller</i>	Greek author	32.800

(table cont'd)

<b><i>Classic Competence</i></b>	<b><i>Translation</i></b>	<b><i>55 % / yes / Keyness:</i></b>
<i>lateinische Sprache</i>	Latin language	32.210
<i>hebräische Sprache</i>	Hebrew language	29.580
<i>griechische Sprache</i>	Greek language	27.820
<i>Gebrauch der Lateinischen Sprache</i>	application of the Latin language	33.650
<i>homerischer Gesang</i>	Homerian song	25.790

A historical document elaborates on the role the *Hebrew language* plays in this theme of classic competence:

*Ein Circulare vom 8. Juli 1785 an alle Consistorien fordert , dass die gelehrten Schulen diejenigen , welche sich der Theologie widmen , zu gründlicher Erlernung der griechischen und hebräischen Sprache ernstlich anhalten sollen. (sic) (1788 Erstes Abiturreglement)*

The quote above is from the oldest published curricular document on the Abitur and states that classic competence and the learning of ancient Greek and the Hebrew language are an integral part of the curriculum for students who aspire to study Theology and need to be tested accordingly.

### ***Sitte / Mental-Cultural Ideals***

*Sitte / Mental-Cultural Ideals* as a theme includes terms with a combined keyness value of 475; one or more of the included term combinations are represented in the modern publications. The theme captures information from 55 % of the publications, where one or more of the term combinations are mentioned.

Table 4.15. Abitur: Mental Cultural Ideas

<b>“Sitte” / Mental-Cultural Ideals</b>	<b>Translation</b>	<b>55 % / yes / Keyness:</b>
<i>besondere Lernleistung</i>	outstanding learning achievement	90.800
<i>Arbeitsweise der Oberstufe</i>	operating principles in upper secondary	50.570
<i>unbedingte Tüchtigkeit</i>	natural efficiency / competence	34.050
<i>sittliche Führung</i>	moral leadership, proper customs	33.910
<i>geistige Arbeit</i>	mental work	35.710
<i>besondere Arbeitsweise</i>	outstanding modes of work	33.220
<i>bedingte Tüchtigkeit</i>	limited efficiency / competence	25.790
<i>vorhergegangene Präparation</i>	previous preparation	34.050
<i>begründete Kenntniss</i>	grounded knowledge	25.790
<i>enthaltene Bestimmung</i>	contained providence	77.640
<i>wissenschaftspropädeutische Bildung</i>	propaedeutic science education	33.910

The first term combination on which I would like to focus is *sittliche Führung*, which would be best translated moral leadership or proper customs. The term *Sitte* being an idiom, it cannot be properly translated but is described through the quote below:

*Diejenigen, welche ein ausländisches Gymnasium besucht haben , oder aus Privat-Unterricht, und nicht unmittelbar von einem Gymnasium zur Universität über gehen , haben die Prüfung ihrer Kenntnissreife unter Einreichung der Zeugnisse ihrer bisherigen Lehrer, über ihre Studien und ihresittliche Führung bei der Prüfungs-Commission eines inländischen Gymnasiums, dessen Wahl den Eltern oder Vor mündern überlassen bleibt, schriftlich auf die in S. 6 . bestimmte Art nachzusuchen, und sich den Anordnungen dieses Reglements zu unterwerfen. (sic) (1834, Prüfungsordnung Deutscher Bund)*

Taken from a historic curricular document, the quote above expresses behaviorial objectives as important competencies for any exam taker of the Abitur. Moral leadership for example, is

valued as much as academic excellence and needs to be accessed with any student before he or she can be admitted into higher education.

The next term combination is *geistige Arbeit*, best translated as mental work. The quote below is taken from a 1960s curricular document:

*Der Schüler der Oberstufe soll sich die Methoden der geistigen Arbeit aneignen und das Zusammenspiel von Arbeitshaltung und Arbeitstechnik üben. (1960, Tutzinger Gespräche III)*

The quote above expresses the importance of attitude, without defining that further. Work attitude is as important, as the knowledge of proper academic methods to succeed and graduate from upper secondary school.

The next term I found needing exemplification is “propaedeutic science education”, *wissenschaftspropädeutische Bildung*, especially as it appears in one of the more modern curricular publications as well as in a 1960s curricular document:

*Vermittelt werden eine vertiefte Allgemeinbildung, allgemeine Studierfähigkeit sowie wissenschaftspropädeutische Bildung. So werden die Grundlagen für fachliches und überfachliches Handeln mit Blick auf Anforderungen von Wissenschaft und beruflicher Bildung geschaffen. Bildungstheoretische Grundlagen des Mathematikunterrichts sind der Allgemeinbildungsauftrag wie auch die Anwendungsorientierung des Unterrichtsfaches Mathematik. (2012 Bildungsstandards Mathe)*

*8. Propädeutik: Verständnis für die philosophischen Einleitungsfragen, besonders für die anthropologischen, ausgehend von Platon oder Descartes oder Kant. 9. Orientierung über die Christenlehre, die kirchengeschichtlichen Hauptereignisse und Einführung in die ethischen Grundfragen. (sic) (1960, Tutzinger Gespräche III)*

The quote from the 2012 publication defines this competence as the ability to relate academic knowledge to its practical applications. Aside the ability to conduct higher studies, students receive a propaedeutic education, especially with regard to the applicability of their knowledge within practical scientific contexts. The 1960s document defines propaedeutic as being the understanding of philosophical and anthropological questions of essence and existence. Ethics here stem from the Greek Classics, as well as from the Christian doctrine.



## ***Reife***

*Reife* as a theme includes terms with a combined keyness value of 559; one or more of the included term combinations are represented in the modern publications. The theme captures information from 55 % of the publications, where one or more of the term combinations are mentioned. This word is a German idiom which is difficult to translate, but socially and historically very representative of the Abitur.

Table 4.16. Abitur: *Reife*

<b>“Reife”</b>	<b>Translation</b>	<b>55 % / yes / Keyness:</b>
<i>Zeugniss der Reife</i>	certificate of maturity (older German)	224.080
<i>allgemeine Hochschulreife</i>	general degree of maturity	160.500
<i>Zeugnis der Reife</i>	certificate of maturity	106.600
<i>Begriff der Hochschulreife</i>	terminology for higher education maturity	34.050
<i>volle Hochschulreife</i>	full higher education maturity	33.910

The first term combination I would like to focus on is *Zeugnis der Reife*, which could be best translated as certificate of maturation. The following quote is taken from a historic curricular document:

*(...) statt durch ein hastig zusammenge rafftes Wissen verwirrt und erdrückt zu werden , sicher und gründlich vorgebildet mit frischer Kraft, mit freudigem Muthe und mit freier Umsicht zur letzten Prüfung stellen können . Während das Reglement, wie es sein Zweck erfordert, die aus dem Gymnasial-Unterricht sich erge benden Gegenstände der schriftlichen und mündlichen Prüfung aufzählt , und für jeden das mittelst dieses Unterrichts zu er reichende ideelle Ziel feststellt, unterscheidet dasselbe diese letzteren Bestimmungen , welche ausdrücklich den Prüfenden nur bei der Schlussberathung zur*

*leitenden Richtschnur für die Ertheilung des Zeugnisses der Reife dienen sollen. (sic)*  
(1834, Prüfungsordnung Deutscher Bund)

The quote expresses the idea that academic mastery springs from the continued long-term endeavor to acquire knowledge. A certificate of maturity (a historical name for Abitur, wording still present in Austria) is not given to anyone displaying a motly assortment acquired in mandatory and hasty studies. Maturity is hence a competence, which is disconnected from the pure academic ability of a student.

The next term combination is *Volle Hochschulreife*, which translates into full higher education maturity, the quote is taken from a 1958 curricular publication:

*Gerade weil nun aber der Begriff der Hochschulreife seit den Schuldebatten der zwanziger Jahre mit dem der „Lebensreife“ so eng verbunden und zum Teil bis zur Ununterscheidbarkeit mit ihm vermischt worden sei (cf. Referat Professor LICHTENSTEIN, S. 21 f), tue eine klare Differenzierung zwischen beiden Begriffen heute dringend not. Erst wenn die Unterscheidung zwischen Hochschulreife als spezifischer „Studierfähigkeit“ und Lebensrcife als allgemein menschlichem Erfahrungs- und Entscheidungsniveau klar sei, könne man dann auch wieder nach ihrem Zusammenhang wie nach ihren eventuellen Kompensationsmöglichkeiten fragen. (sic) (1958, Tutzingen Maturitätskatalog)*

The quote expresses the notion that higher education maturity should be distinguished from a social, cultural maturity of the individual within society. Previously used nearly synonymously, the relationship these terms share is marked through commonalities, as well as through differences. While sharing similar properties, higher education maturity must be disconnected from a general lived experience of the student within the society.

### **Findings for the Baccalauréat**

Challenges for coding the textual data from curricular documents relating to the Baccalauréat arose through orthographical errors and an apparent shortcoming of optical character recognition for aspects of French orthography. Despite the challenge this posed, no key term combination has a significant keyness score solely due to its being an morphologic

aberration. The ideal morpheme for each aberration included in the following has a significant keyness score as well, which permits to add the significance of the aberrations to the significance of the ideal form, which I eventually did for this section.

### ***Communicative Competence***

*Communicative Competence* as a theme includes terms with a combined keyness value of 1610, one or more of the included term combinations are represented in the modern publications. The theme captures information from 89 % of the publications where one or more of the term combinations are mentioned.

Table 4.17. BAC: Communicative Competence

<b>Communicative Competence</b>	<b>Translation</b>	<b>89 % / yes / Keyness:</b>
<i>langue vivante</i>	modern language	302.470
<i>langues vivante</i>	modern languages	211.620
<i>langue vivante</i>	modern language	62.520
<i>explication française</i>	Analysis of a text	331.570
<i>interrogation</i>	Short test (oral or written)	104.300
<i>sufficiente articulation</i>	sufficient articulation	41.050
<i>langue vivante étrangère</i>	modern foreign language	327.990
<i>langue étran</i>	foreign language (spelling error)	42.400
<i>langue étrangère</i>	foreign language	42.320
<i>langue vivante étrangère</i>	modern foreign language (spelling error)	39.520

The first term combination on which I focused is *langue vivante*, which is a term to include all modern, “living languages”. The following quote is taken from a historical document, where the author first argued to include modern foreign languages in the Baccalauréat:

*(...) l'enseignement classique et l'enseignement des langues vivantes, en les mettant à leur véritable place, et en les disposant dans l'ordre qui leur convient. Ce plan constituera, je l'espère, les études de nos sur des fondements durables. Il préviendra la confusion des idées, la dispersion et l'affaiblissement de l'attention ; car le meilleur système d'études ne consiste pas à enseigner le plus de choses à la fois, mais à enseigner chacune d'elles en son temps. (...) Le nouveau programme exercera l'intelligence des élèves par une culture mieux appropriée à chaque âge et à la nature de l'esprit humain ; et il est permis d'espérer qu'il préparera à l'État et au pays des générations plus fortes (sic). (1840, Instruction relative à l'exécution de l'arrêté)*

The author argues that modern foreign languages carried an importance similar to that of the ancient languages, which appear to have been given preference until that time. The author went on to stress that thorough and age-appropriate foreign language study is in the interest of a vital citizenship and would help educate the spirit of the students.

Another interesting passage is about the term combination *explication française*, which could be translated as the analysis in French of the content of a foreign language text.

*Explication d'un texte français dit moyen Age ou du dernier siècle ;  
2\* Explication d'un texte français classique ou moderne ;  
3\* Explication d'un texte de langue vivante étrangère ;  
4° Explication d'un texte d'une seconde» langue vivante étrangère. (1927, Décret Baccalauréat)*

The above passage reveals the importance of modern foreign languages, an importance at least equal to the mastery of the French and of ancient, classic language. The quote is just a listing, but the decree stresses the importance for students to have the skill to come to a cogent analysis in French of a foreign language text. The practice taken from French literary studies consists of closely examining the language of a text to determine its meaning. The application of the practice unto foreign language texts could be seen as marking the increased importance of foreign languages for the French secondary school of that time.

### ***Production, Interpretation and Manipulation of Text***

*Production, Interpretation, and Manipulation of Text* as a theme includes terms with a combined keyness value of 1143; one or more of the included term combinations are represented

in the modern publications. The theme captures information from 67 % of the publications, where one or more of the term combinations are mentioned.

Table 4.18. BAC: Production, Interpretation and Manipulation of Text

<b>Production, Interpretation and Manipulation of Text</b>	<b>Translation</b>	<b>55 % / yes / Keyness:</b>
<i>composition française</i>	short French essay	431.470
<i>texte français</i>	French text	173.070
<i>dissertation philosophique</i>	Long philosophical essay	142.640
<i>dictée française</i>	French dictation	63.230
<i>cohérence</i>	Coherence	42.400
<i>dissertation française</i>	Long French essay	42.150
<i>question littéraire</i>	question on literature	40.580

The first term combination on which I chose to focus is “question littéraire”, which roughly translates into “question on literature”. The following quote is taken from a historic publication:

*Les questions littéraires du nouveau programme forment un ensemble de connaissances diverses que doit posséder tout élève de rhétorique qui a profité des leçons qu’il a reçues ; mais ces connaissances, il doit les avoir recueillies successivement dans les deux grands exercices dont se compose une classe de rhétorique, la lecture et l’explication des grands maîtres de l’art d’écrire , et leur imitation par la composition quotidienne. (1840, Instruction relative à l’exécution de l’arrêté)*

The author of the quote talks about inclusions into a new program for the Baccalauréat, one being rhetoric mastery, which is at the heart of literary theses, and is honed through the reading and in-depth study of great Classics, as well as through their imitation.

Another term combination which provides for an interesting passage was *dictée française*, translated as “French dictation” and *composition française*, short French essay. The following passage addresses both and is taken from a historical publication :

*Une dictée française suivie de questions sur certaines parties du texte dicté permettant de constater chez les candidats la connaissance de la langue et l'intelligence du texte ; 2 une composition française ou une composition sur une des matières du cours (histoire, géographie, sciences) (sic). (1902, Programme des examens du Baccalauréat)*

Through a dictation, students reveal their mastery of the spelling of the French language, as well as insight into the ingenuity of the text. The passage reveals an important link between the understanding of a text and the mastery of the native language, French. The latter is expressed through the mention of the French short essay. The connection between both is considered inherent in the document. For this theme, I found that passages in modern publications barely reflect on the competencies required for textual production, despite the high occurrence of terms related to the theme throughout. Though the insight into the relationship between form and mastery is revealing here, the French dictation is a practice which is probably not practiced in the Modern French school system anymore.

### ***Modes of Reasoning***

*Modes of Reasoning* as a theme includes terms with a combined keyness value of 482; one or more of the included term combinations are represented in the modern publications. The theme captures information from 64 % of the publications, where one or more of the term combinations are mentioned.

Table 4.19. BAC: Modes of Reasoning

<i>Modes of Reasoning</i>	<i>Translation</i>	<b>64 % / yes / Keyness:</b>
<i>nature scolaire</i>	scholarly nature	62.980
<i>façon cohérente</i>	coherent manner	47.360
<i>scolarité obligatoire</i>	Mandatory schooling	46.520
<i>monde contemporain</i>	contemporary world	43.430
<i>théorie scientifique</i>	scientific theory	42.930
<i>attente fondamentale</i>	fundamental expectation	42.070
<i>contrôle continue</i>	continuous assessment	41.920
<i>d'aptitude</i>	(of) aptitude	38.940
<i>formation intellectuelle</i>	intellectual development	36.690
<i>pensée complexe</i>	complex thought	36.160
<i>théorie scientifique</i>	scientific theory	42.930
<i>cohérence globale</i>	overall coherence	51.79

The first term combination on which I would like to focus is *façon cohérente*, which translates as “coherent manner”. The following passage is taken from a modern curricular document:

*(...) leurs finalités doivent être présentées, de façon cohérente tout au long de leur parcours, les choix inhérents aux programmes doivent être dans toute la mesure du possible explicités et les fonctions, langages et méthodes des différentes disciplines*

*doivent être d'autant plus clairement présentés que les mêmes objets d'enseignement peuvent relever de plusieurs d'entre elles.* (2014, Charte relative à l'élaboration, à la mise en œuvre et au suivi des programmes)

The text is focused on competencies students need to express during their final exams. Students' final examinations coherently reflect their individual educational track and subject choice. Here the focus is on the fact that the students choose to specialize into subject matters and acquire sufficient skill to communicate about their subjects. Their language, as well as their methods display refinement in their respective specializations.

Another example here is the term combination *formation intellectuelle*, which translates as “intellectual education” or “intellectual development”. The following quote is taken from another modern source:

*Notre ambition est de proposer des programmes qui contribueront à la formation intellectuelle et civique des jeunes générations et leur donneront les clés pour comprendre le monde dans lequel ils vivent.* (2018, Note d'analyses et de propositions sur les programmes)

Programs are designed to form students intellect and integrity, which helps them in turn to best understand their lived realities. The text expresses the idea that intellectual development needs to happen with a focus on practical application of knowledge.

### ***Oral Mastery***

*Oral Mastery* as a theme includes terms with a combined keyness value of 739; one or more of the included term combinations are represented in the modern publications. The theme captures information from 56 % of the publications, where one or more of the term combinations are mentioned. Like in German, this category has the least number of key terms of all, which have relatively high keyness values. It also still manages to reach a similar source coverage, compared to other categories, which speaks for its significance.



Table 4.20. BAC: Oral Mastery

<i>Oral Mastery</i>	<i>Translation</i>	<i>56 % / yes / Keyness:</i>
<i>épreuve orale</i>	oral exam	163.280
<i>épreuves orale</i>	oral exams	141.810
<i>orale terminale</i>	oral final exam	83.970
<i>épreuve orale terminale</i>	oral final exam	83.970
<i>examen oral</i>	oral exam	58.600
<i>autre orale</i>	other oral examination	42.320
<i>nouvelle épreuve orale</i>	new oral exam	41.980
<i>épreuves orale</i>	oral exams	41.570
<i>nouvelle épreuve orale terminale</i>	new final oral exam	42.490
<i>competence orale</i>	oral competence	38.940

The first term combination I would like to focus on here is *épreuve orale*, which translates into oral exam. The following is taken from a historical publication:

*Épreuves orales.*

*1° L'explication d'un texte latin;*

*2° L'explication d'un texte français;*

*Ces textes sont choisis dans les ouvrages*

*Inscrits aux programmes des classes de seconde et de première de la section C de l'enseignement secondaire.*

*3° Une épreuve de langue vivante étrangère, allemand, anglais, espagnol ou italien, au choix du candidat, d'après le mode déterminé par l'instruction annexée au présent décret;*

*4° Une interrogation sur l'histoire, d'après le programme de la classe de première de la*

*section C de l'enseignement secondaire;*

*5° Une interrogation sur la géographie, d'après le même programme;*

*6° Une interrogation sur les mathématiques, d'après le même programme;*

*7° Une interrogation sur la physique, d'après le même programme;*

*8° Une interrogation sur la chimie, d'après le même programme. (sic) (1902, Programme des examens du Baccalauréat de l'enseignement secondaire. Décret et Arrêté.)*

The passage above is taken from a curricular document of 1902 and describes modes and pathways to the oral exam. At the beginning the choice is given (to the student) to choose to have their oral exam done on a French and a Latin text, if the student has specialized into both of the given languages during grades 10 – 12. Another choice for an oral exam includes either German or English and another foreign language chosen by the student, depending on their specialization during secondary school. A third choice is given to students in a different track and includes an oral exam on topics treated in history class during grades 10-12. The options presented in four to seven stipulate the possibility for students to take the oral exam in non-language subjects, e.g. history, geography, maths, physics and chemistry. In one of these cases, the content of the oral final exam will also be taken from the materials covered during grades 10-12.

Another example is from the term combination *nouvelle épreuve orale terminale*, which translates into “new oral final exam”. The following passage is taken from a modern publication :

*La nouvelle épreuve orale terminale : les conditions de réussite d'une épreuve de nature scolaire. La création d'une épreuve orale terminale constitue l'une des grandes avancées du futur baccalauréat : elle accorde une importance inédite à la construction et à l'évaluation de la compétence orale, jusqu'alors minorée dans l'enseignement secondaire. Savoir s'exprimer à l'oral (comme à l'écrit), dans un français correct, avec des arguments pesés et organisés, est en effet essentiel tant pour les études ultérieures que pour la vie personnelle et professionnelle : il conviendra donc que les différents enseignements, et notamment les enseignements de spécialité, envisagent des temps spécifiques à la construction et à l'exercice de cette compétence. (2018, Note d'analyses et de propositions sur les programmes du lycée et sur les épreuves du baccalauréat)*

The above passage is taken from a modern curricular document and describes today's modalities on the final oral exam within the Baccalauréat. At the beginning of the quote, the increased

importance of an oral exam element is stated for the newly designed Baccalauréat (2020 and beyond under normal circumstances). The quote continues to underline the importance of a fluent expression in spoken French of well thought and well-structured arguments for student success in all professional, as well as personal spheres in society. Teachers in all subjects, but foremost in the specializations, are called to put aside time for the specific preparation of students' oral competence in their respective subjects.

### ***Classic Competence***

*Classic Competence* (Latin, Ancient Greek, Study of the Classics) as a theme includes terms with a combined keyness value of 1143; one or more of the included term combinations are represented in the modern publications. The theme captures information from 67 % of the publications, where one or more of the term combinations are mentioned.

Table 4.21. BAC: Classic Competence

<b><i>Latin, Ancient Greek, Classic Studies</i></b>	<b><i>Translation</i></b>	<b><i>67 % / yes / Keyness:</i></b>
<i>version latine</i>	Latin version	389.900
<i>explication latine</i>	Explication of Latin text	187.320
<i>version grecque</i>	Greek version	133.160
<i>étude classique</i>	classical studies	103.690
<i>langue ancienne</i>	ancient language	82.590
<i>explication grecque</i>	Explication of Greek text	63.230
<i>texte latin</i>	Latin text	60.540

(table cont'd)

<i>Latin, Ancient Greek, Classic Studies</i>	<i>Translation</i>	<i>67 % / yes / Keyness:</i>
<i>lca grec ancien</i>	“LCA” Ancient Greek	42.490
<i>I latin</i>	Latin (with artifact)	42.490
<i>culture antique</i>	Classical culture	37.830

The first term combination here is *étude Classique*, which translates as “classical studies” and the following passage is taken from a historical publication:

*1 ° Les portions d’enseignement scientifique réparties depuis la sixième jusqu’à la rhétorique inclusivement ne produisent aucun bon résultat ; 2 ° cet enseignement accessoire, infructueux en lui – même, nuit considérablement aux études classiques. Il a donc fallu, monsieur le recteur, chercher une autre combinaison. C’est un principe reconnu que les études doivent être proportionnées à l’âge des élèves. C’est un autre principe également reconnu que, dans un même âge , toutes les études doivent être analogues pour produire une impression forte et durable. Voilà pourquoi l’expérience générale a place d’abord les études classiques, si bien appelées humanités, parce qu’elles forment l’homme, et cultivent à la fois la mémoire , l’imagination , l’esprit et le cœur. (sic). (1840, Instruction relative à l’exécution de l’arrêté)*

In this quote from one of the historical documents, the author criticized a singular understanding of education as being science education. A singular focus on science education eventually hurts the development of children. As a consequence, the author proposed to introduce classical studies as a complement according to the children’s grade level. In a humanist sense, classical studies help develop a combination of memory, imagination, spirit, and soul (heart).

The next term combination is *version grecque*, which translates as “Greek version” and refers to the translation of a text into the mother tongue in the following historical passage:

*Pour les épreuves écrites, sauf pour la version latine, la version grecque et l’épreuve de langues vivantes, il est donné trois sujets, entre lesquels les candidats ont le droit de choisir. Art. 12. — L’analyse de tout ou partie d’un des textes portés au programme de l’examen, une question d’histoire ou de géographie ne peuvent constituer un sujet de composition écrite. (1902, Programme des examens du Baccalauréat.)*

The above quote from a historical document expresses the idea that an examination of mastery of the Latin or Ancient Greek languages should be combined with subject matter from other disciplines. The exam takers are allowed to choose to have their written exam in Latin or in Ancient Greek in one of 3 topics given to them from those subjects (except from history or geography).

The last excerpt relates to the term combination *langue ancienne*, which refers to ancient languages and is taken from a modern curricular document:

*Par l'étude fine de la langue ancienne et sa confrontation avec les langues européennes, ils enrichiront leur compréhension et leur maîtrise des systèmes linguistiques et de leur propre langue : la confrontation des systèmes de la langue ancienne et de la langue française leur permettra d'en saisir les finesses et les différences. (2018, Note d'analyses et de propositions sur les programmes)*

The above quote taken from a modern document refers to the study of old languages as a means to come to a deeper understanding of the linguistic structures which underlie those modern languages, modern French and other modern European languages. The study of ancient languages is carried out in comparison with modern languages.

### ***Culture and Morals***

*Culture and Morals* as a theme includes terms with a combined keyness value of 649; one or more of the included term combinations are represented in the modern publications. The theme captures information from 56 % of the publications, where one or more of the term combinations are mentioned.

Table 4.22. BAC: Culture and Morals

<b><i>Culture and Morals</i></b>	<b><i>Translation</i></b>	<b><i>56 % / yes / Keynes:</i></b>
<i>culture scolaire</i>	School culture	143.920
<i>enseignement moral</i>	moral education	101.110
<i>culture commune</i>	Common core culture	84.560
<i>rite municipale</i>	district rite	63.230
<i>enseignements communs conciliant culture</i>	Common core teachings linking culture (with..)	42.490
<i>culture humaniste</i>	humanist culture	69.400
<i>nature scolaire</i>	Nature of schooling	62.980
<i>attente fondamentale</i>	fundamental expectation	42.070
<i>pratique scolaire</i>	scholarly practice	39.230

This theme includes terms that are often mentioned but rarely reflected upon. During my search for collocations, which would help give a definition to this theme as a social or cultural competence, I found the following passage on *culture scolaire*, which translates as “school culture” from a modern publication:

*Cette culture scolaire commune doit respecter les exigences suivantes : sa constitution est fondée sur des disciplines historiquement construites et en évolution, établies sur un certain nombre de concepts, d’objets et de méthodes ; les concepts considérés par les groupes comme fondamentaux pour chaque enseignement seront clairement mentionnés dans les projets de programme ; elle est fondée sur l’étude, c’est-à-dire sur un travail personnel de l’élève qui acquiert peu à peu son autonomie ; elle s’appuie sur l’acquisition des savoirs, favorise la réflexion et l’esprit critique ; elle enrichit la personne dans sa vie intellectuelle, morale, civique et sociale ; elle aborde l’histoire des savoirs ainsi que leurs enjeux contemporains ; elle donne des repères aux élèves pour les aider dans leur choix d’orientation ; elle exprime un nécessaire équilibre et une cohérence d’ensemble entre culture humaniste et culture scientifique et technologique ; elle permet aux élèves d’inscrire leur compréhension du monde dans le temps long et*

*dans la pleine conscience des grandes évolutions scientifiques et technologiques de notre temps. (2018, Notes d'Analyses et de Propositions)*

The school culture consists of historically evolved fundamental concepts. These concepts need to be communicated in the respective curricula of the programs and in their subjects (in secondary school, leading to the Baccalauréat). Furthermore, the quote states that school communities have the authority to decide upon their fundamental tenets, and it is the aim of the student to appropriate these fundamental tenets through intellectual growth, as much as it is the task of the teacher to instruct the students accordingly. Through this growth, the students refine their intellect, moral decision making and build social integrity. A school culture takes into account scientific reasoning, as well as humanism.

### **Findings for the SAT Test**

Like in the previous sections, this section focuses on themes that are shared, to some degree, across as many of the sampled texts as possible. Unlike in the previous two sections, I have to take into account the unique relationship the College Board entertains with its past, that the organization has continuously, scientifically reflected upon. As a consequence, I used primarily modern sources to illustrate sample terms. The keyness overall remains an important measure, as despite all the historical reinventions, some terms keep reappearing in older as in modern publications by the College (Entrance Examination) Board. Unlike in the previous sections, I will not have to translate the here quoted passages. For a description and a reflection on the findings, please refer to section 1 of this chapter, where I compare the results of the narrative synthesis.

#### ***Communicative Competence***

*Communicative Competence* as a theme includes terms with a combined keyness value of 417; one or more of the included term combinations are represented in the modern publications.

The theme captures information from 72 % of the publications, where one or more of the term combinations are mentioned.

Table 4.23. SAT: Communicative Competence

<i>Communicative Competence</i>	<i>72 % / yes / Keyness:</i>
word choice	94.270
language use	47.150
Reading	34.780
effective language	30.400
reading test	30.400
effective language use	25.500
vocabulary knowledge	24.560
English language	24.120
English proficiency	22.500
language ability	21.960
reading comprehension	20.760
written English grammar	20.600
rhetorical purpose	20.480

The first term combination I exemplify here is *word choice*. The following quote is taken from a modern publication:

(...) [students] make careful and considered use of evidence as they read and write; demonstrate skill in analyzing data, including data represented graphically in tables, graphs, charts, and the like, in reading, writing, and math contexts; and reveal an



understanding of relevant words in context and how word choice helps shape meaning and tone. Evidence such as this of what's truly important for college and career readiness pervades the work of the College Board in both instruction and assessment and will be the focus not just of the board but also of the models of student work that we will partner with educators to develop (sic) (...). (2015, Standards Redesigned SAT)

A second example is taken from the term combination *effective language* and is taken from a modern publication:

The basic aim of the redesigned Writing and Language Test is to determine whether students can demonstrate college and career readiness proficiency in revising and editing a range of texts in a variety of content areas, both academic and career related, for development, organization, and effective language use and for conformity to the conventions of Standard Written English grammar, usage, and punctuation. The test comprises a series of high-quality multiparagraph passages and associated multiple-choice questions. (2015, Test Specifications for the Redesigned SAT)

***Production, Interpretation and Manipulation of Text***

*Production, Interpretation, and Manipulation of Text* as a theme includes terms with a combined keyness value of 419; one or more of the included term combinations are represented in the modern publications. The theme captures information from 64 % of the publications, where one or more of the term combinations are mentioned.

Table 4.24. SAT: Production, Interpretation and Manipulation of Text

<b><i>Production, Interpretation and Manipulation of Text</i></b>	<b><i>64 % / yes / Keyness:</i></b>
text complexity	164.740
written analysis	43.910
writing assessment	38.800
writing act	35.260
essay task	35.230
reading test	30.400

(table cont'd)

<b><i>Production, Interpretation and Manipulation of Text</i></b>	<b>64 % / yes / Keyness:</b>
literary reading	29.630
clear written analysis	20.600
written English grammar	20.600

The following passage is on the term combination *written analysis* and taken from a modern publication:

(...) texts that examine in an accessible way ideas, debates, trends, and the like in the arts, the sciences, and civic, cultural, and political life. In response to these passages, students must produce a clear and cogent written analysis in which they explain how the author of a text builds an argument to persuade an audience through the use of evidence, reasoning, stylistic and persuasive elements, and/or other features the students themselves identify. (2015, Test Specifications for the Redesigned SAT)

Another interesting passage with the coding rationale in mind is from the same publication on the term combination *literary reading*:

Young people do not understand the ideals of citizenship (...) and their appreciation and support of American democracy is limited. It is probably no surprise that declining rates of literary reading coincide with declining levels of historical and political awareness among young people. (2015, Test Specifications for the Redesigned SAT)

### ***Modes of Reasoning***

*Modes of Reasoning* as a theme includes terms with a combined keyness value of 1701, one or more of the included term combinations are represented in the modern publications. The theme captures information from 94 % of the publications, where one or more of the term combinations are mentioned. Based on the sum of keyness scores and the distribution of the theme across the source texts, this theme has by far the highest significance of all listed.

Table 4.25. SAT: Modes of Reasoning

<i>Modes of Reasoning</i>	<i>94 % / yes / Keyness:</i>
critical reading	306.060
scholastic aptitude	289.440
evidence-based reading	170.440
text complexity	164.740
critical reading section	98.600
reasoning test	62.090
mathematical reasoning	49.410
source analysis	47.980
reading challenge	46.160
written analysis	43.910
deep reading	40.040
textual evidence	36.810
evidence use	34.910
science literacy	34.770
reasoning ability	28.700
ed range of text complexity	25.500
range of text complexity	25.500

(table cont'd)

<i>Modes of Reasoning</i>	<i>94 % / yes / Keyness:</i>
constructing meaning	25.000
quantitative literacy	24.560
full context	23.500
clear written analysis	20.600
complexity example	20.600
text complexity example	20.600
metalinguistic understanding	20.580
linking knowledge	20.370
psychological examination	20.130

The first term combination for which I found an interesting passage is *critical reading* and the following excerpt is taken from a modern publication:

For example, Marilyn Jager Adams, reviewing in the research literature on the challenges students face reading complex texts, helped collect a range of scholarly evidence documenting a decadeslong decline in – text complexity even as college and career readiness demands on students’ reading skills remained high. The disparity between high school and postsecondary expectations for text complexity has left too many students underprepared for the rigors of reading in college and careers. One sign of this problem can be found in recent test data, which indicate that only about half of all test-takers attained a score on the Critical Reading test high enough for them to be considered college ready. (2015, Test Specifications for the Redesigned SAT)

The next passage is a collocation of the term combination *mathematical reasoning* and is taken from an older publication:

As most competent psychometrists have known for many years, single-IQ tests of the Stanford Binet variety are of little value for differentiating among cognitive abilities. Even the Wechsler tests are not as useful for educational placement and facilitation as are the sat and, for example, the Differential Aptitude Test battery. Mathematical score does. In its talent searches, sMpy has regularly found quite a few youths who are far better at

mathematical reasoning than in reading comprehension and vocabulary knowledge, and sometimes vice versa. Often we learn more of educational relevance about a particular student in 150 minutes of large-group SAT testing than could be determined by individual administration of the Stanford-Binet (...). (1977, The Predictive Value of the SAT)

Another insightful collocation concerns the term combination *science literacy* and is taken from a modern publication:

knowing science requires individuals to integrate a complex structure of many types of knowledge. These knowledge types include the ideas of science, the relationships between the ideas, the reasons for these relationships, and the ways to use these ideas to complete the following tasks: explain and predict other phenomena, interpret situations, solve problems and participate productively in science practice and discourse. Much of the supporting research is summarized in synthesis reports and research handbooks, such as (...). (2006, Standards Science)

### ***Readiness***

*Readiness* as a theme includes terms with a combined keyness value of 997; one or more of the included term combinations are represented in the modern publications. The theme captures information from 94 % of the publications, where one or more of the term combinations are mentioned.

Table 4.26. SAT: Readiness

<b><i>Readiness</i></b>	<b><i>94 % / yes / Keyness:</i></b>
postsecondary readiness	20.450
student readiness	20.100
college readiness	298.440
career readiness	290.180

(table cont'd)

<b><i>Readiness</i></b>	<b><i>94 % / yes / Keyness:</i></b>
readiness benchmark	127.880
college readiness benchmark	49.950
readiness status	39.880
academic preparation	38.540
career readiness proficiency	30.400
readiness proficiency	30.400
college readiness status	30.400
career readiness assessment	20.600

The first term combination is one with the highest keyness score, being readiness benchmark, for which I found two insightful passages from modern publications:

As would be expected, there is a strong relationship between the SAT College Readiness benchmark and these measures of high school performance. For example, when looking at HSGPA, approximately 9 to 12 percent of students with a HSGPA of C (C+, C, or C-) or lower met the benchmark, compared to over 57 to 84 percent of those with a HSGPA of A (A+, A, or A-). College Board Research Reports 15 SAT Benchmarks College readiness also varied considerably with academic course-taking behavior. One measure of college preparation is a core curriculum which consists of four years of English and three years each of mathematics, science, and social science/history. Half of the students who took a core curriculum met the benchmark, compared to 29 percent of those who did not. (2011, College Readiness Benchmark)

While the SAT college readiness benchmark can be an accurate indicator of the academic preparedness of students, it does not consider other noncognitive factors such as motivation and persistence, which are also linked to success in college. Due to the omission of noncognitive factors and other measures of high school achievement, the benchmark is designed to evaluate

the aggregate readiness of a group of students rather than the individual student. When evaluating the individual student, the SAT should not be the only piece of information considered in making decisions on readiness for college. (2011, College Readiness Benchmark)

The next passage I found here is a collocation to the term combination *college and career Readiness* and is taken from a modern publication:

Each test within the redesigned model is designed to collect evidence from student performance in support of a broad claim about what students know and can do (...), and each claim is aligned to the exam's primary purpose of assessing college and career readiness. The exam is, at its core, a postsecondary admission and guidance exam, and as such is designed to be a strong predictor of postsecondary success as measured by first-year, retention to second and subsequent years, and overall completion of postsecondary education. (...) the exam's predictive validity. (2015, Standards Redesigned SAT)

#### ***Alignment with Authority and Degree of Social Assimilation***

*Alignment with Authority and Degree of Social Assimilation* as a theme includes terms with a combined keyness value of 659; one or more of the included term combinations are represented in the modern publications. The theme captures information from 88 % of the publications, where one or more of the term combinations are mentioned.

Table 4.27. SAT Alignment with Authority and Degree of Social Assimilation

<b><i>Alignment with Authority and Degree of Social Assimilation</i></b>	<b><i>88 % / yes / Keyness:</i></b>
college-level study	20.370
workforce training	121.350
essential knowledge	52.940
test competence	35.300

(table cont'd)

<i>Alignment with Authority and Degree of Social Assimilation</i>	<i>88 % / yes / Keynes:</i>
academic rigor	28.930
nonacademic success	25.500
core curriculum	24.960
powerful knowledge	24.850
own curriculum	23.120
benchmark attainment	20.600
mathematics benchmark	20.600
writing benchmark	20.600
ccssm content standard	20.600
nontraditional study	20.600
criterion-referenced testing	20.580
institutional selectivity	20.560
college-level science	20.480
college-level study	20.370
aptitude test	153.060

The first passage is related to the term combination *workforce training* and is taken from a modern curricular document:

We can't continue to allow vast numbers of our country's students to fall behind academically. It's therefore critical that we do everything possible to ensure that all students are on a trajectory to gain meaningful access to postsecondary courses and workforce training programs, complete degrees and certifications, and participate



successfully in an increasingly competitive and fluid global economy. (2015, Standards Redesigned SAT)

Another passage relating to the coding rationale is on the term combination *academic rigor* and taken from a modern publication:

One measure of college preparation is a core curriculum which consists of four years of English and three years each of mathematics, science, and social science/history. Half of the students who took a core curriculum met the benchmark, compared to 29 percent of those who did not. The academic rigor index (ARI), which measures the challenge associated with high school course work, is highly related to the core curriculum. (2011, College Readiness Benchmark)

### ***Social Factors of Learning / Social Stratification***

*Social Factors of Learning / Social Stratification* as a theme includes terms with a combined keyness value of 659; one or more of the included term combinations are represented in the modern publications. The theme captures information from 59 % of the publications, where one or more of the term combinations are mentioned. This theme has the weakest overall significance based on a combination of keyness scores and the percentage of distribution across publications.

Table 4.28. SAT: Social Factors of Learning

<b><i>Social Factors of Learning / Social Stratification</i></b>	<b><i>59 % / yes / Keyness:</i></b>
individual intelligence	20.330
parental education	27.530
fair assessment	23.590
income group	21.980

(table cont'd)

<b><i>Social Factors of Learning / Social Stratification</i></b>	<b><i>59 % / yes / Keyness:</i></b>
income level	20.920
low parental income	20.600
annual parental income	20.580

An interesting passage in terms of a question for the revelation of cultural or social competence of the exam taker is a collocation of the term combination *parental education* and is taken from a modern publication:

Camara, Kobrin, and Sathy (2005) showed that academic rigor, or the amount of advanced or AP courses taken by a student in a number of different subject areas in high school, is related to parental income and parental education level. Therefore, the greater academic rigor of the higher SES students' course work may explain their better performance on the SAT. Two additional explanations are commonly offered for performance differences by income level on standardized admissions tests (Zwick, 2004). The first explanation for this relationship is that ingrained in the content of these tests is a white, middle-class way of thinking that disadvantages lower-income students who think differently. (2007, A Historical View of Subgroup Performance Differences)

Another term combination which produced an interesting collocation to the coding rationale is *income level* and the following is taken from another modern publication:

She found that even when the content of a test was more focused on material learned in the classroom there were still comparable differences in average score by income level. Also, after reviewing the work of Camara and Schmidt (1999) and Owings, McMillen, and Burkett (1995), Zwick demonstrated that students' HSGPA, and their ability to meet selective admissions criteria (based on SAT score, HSGPA, participation in at least two extracurricular activities, and positive teacher perceptions) were also correlated with parental income level. Therefore, it is not only standardized test scores expressing the educational disadvantage of students from lower-income families, but many other factors related to academic achievement. (2007, A Historical View of Subgroup Performance Differences)

## **Results Network Text Analysis**

The following section is dedicated to the results of several analytical procedures, each contributing to a deep and thorough understanding of exam content and the networks that content is embedded into. I particularly seek to answer Research Question 2: How are social or cultural capital embedded within and revealed in comparison across the Baccalauréat, Abitur and SAT Test? I further elaborated on the potential of each analytical step to respond to certain aspects of social and cultural capital. As laid out in Chapter 3, all representations of choice are categorical numbers. I added the code tables with the actually calculated numbers in Appendix 2.

### **Findings: Representation of Choice**

An important first result towards RQ1, when comparing the three exams, comes from the fact that all subjects in the Abitur and the French exam are subject to tracking schemes, as the German and French exams feature a significantly larger number of subjects than the SAT; at the same time they cannot all be combined freely with one another. Furthermore, unlike the SAT, the German and French exams offer multiple degrees of specialization, which add a layer of complexity to the broad assumption of choice. It is also important to note that the SAT has the same representation across all dimensions, as it lacks any form of specialization, and its subjects can be freely combined with one another (no tracking) and since it offers no form of answer choice for any of the subjects. Nevertheless I chose to represent it as a reference in all diagrams.

Following are all the subjects, which the three exams have in common and which therefore can be used for comparison. During the analysis, four types of representations emerged, which I discuss in the following sections.

### ***Type 1: Mathematics***

The subject Mathematics stands out, as it has one of the most interesting representations across all three exams and one of the highest overall scores, generated despite several differences between the three exams.

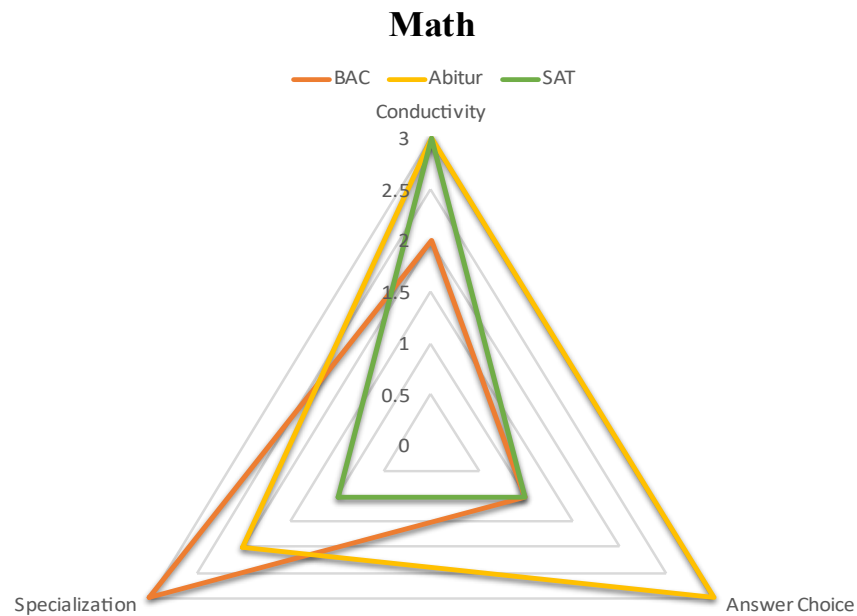


Figure 4.2. Representation of Choice: Math

Figure 4.2. reveals that in terms of conductivity, Math has one of the highest subject scores in the Abitur, which indicates that it is in an essential part in all subject tracks. The Baccalauréat has a slightly lower conductivity score in this subject, which means that it is not among the most numerous appearing subjects in the exam and can even be left out [note: this is true for the SAT and the *terminale* of the Baccalauréat].

A fact that leads to the dimension of specialization, where the Baccalauréat and Abitur majorly differentiate in the aspect that the BAC allows for Math as a subject to be left out in its final exam [not in *primaire*, but in *terminale*], which is impossible in any track for the Abitur. In terms of Answer Choice, the Abitur was the only exam, which offered sets of questions to choose from.

In summary, Math is a subject that has the overall highest degree of choice all dimensions taken together, and accounted for all differences between the exams. The representation is also unique, as it is the only subjects which displays these structural differences across the Abitur, Baccalauréat and SAT.

### ***Type 2: Native Language Literacy***

Native Language literacy has a unique representation as shown below in Figure 4.3:

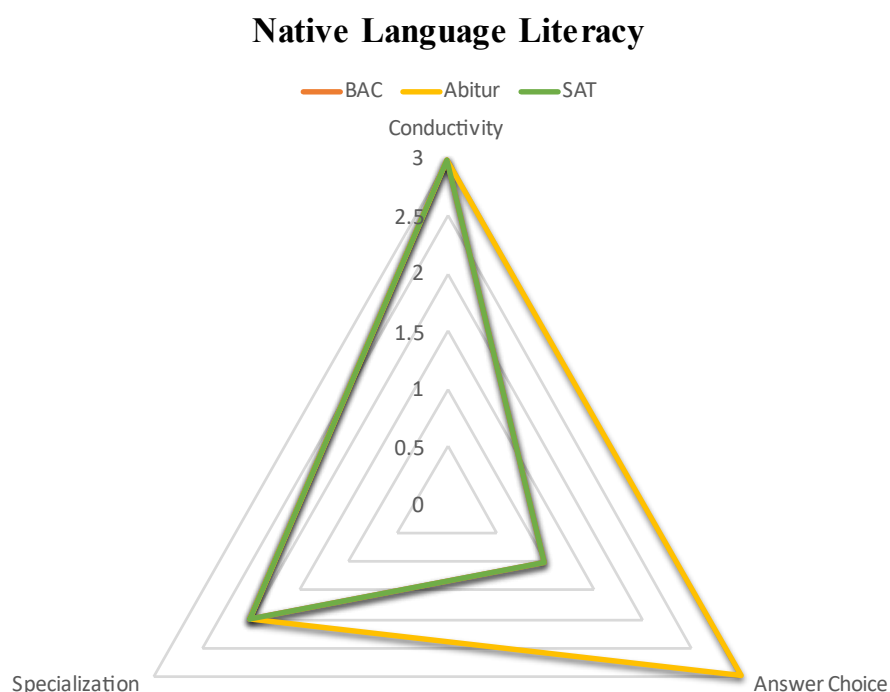


Figure 4.3. Representation of Choice: Native Language Literacy

While the high conductivity score for all exams indicates the importance of this subject in all subject tracks, the lower specialization score is a representation of a de-facto absence of choice, which is born from the subject's importance: it cannot be left out in either of the exams (Baccalauréat or Abitur), which allow for specialization. This fact makes the score coincide with that of the SAT, where no specialization is possible, but in turn the subject can be left out. Both scenarios lead to a specialization score of 2 out of 3, while the Abitur and the Baccalauréat

receive their subtraction due to a different reason. Furthermore, the representation stands out due to the Abitur and the Baccalauréat having the highest score in terms of answer choice here, because both exams offer their students three sets of questions to choose from.

### ***Type 3: History***

The representation of the subject History across all three exams is unique as shown in Figure 4.4:

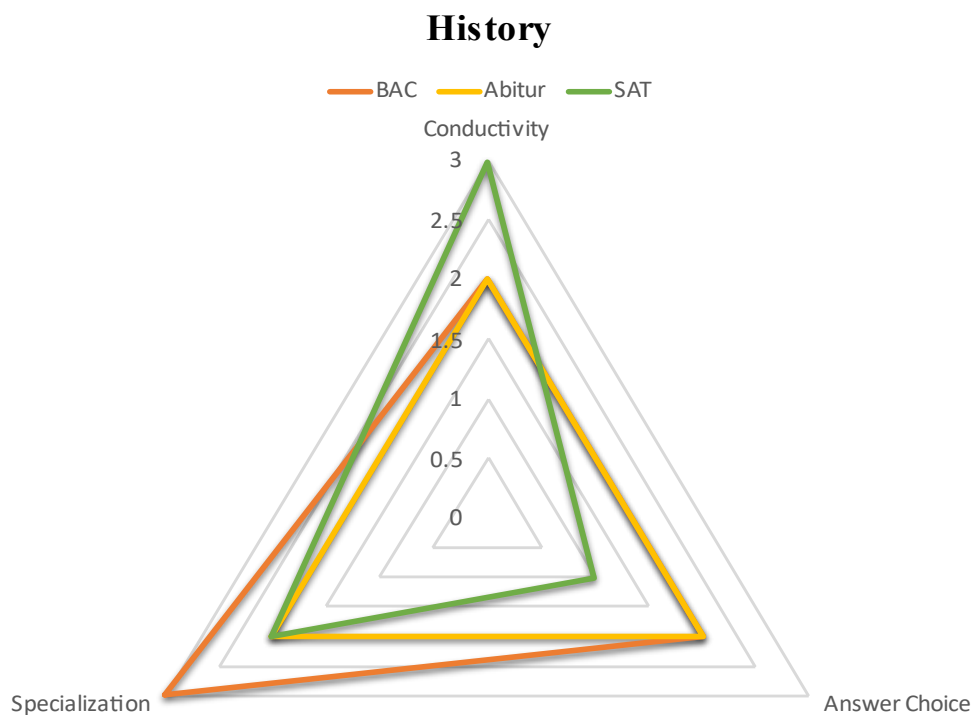


Figure 4.4. Representation of Choice: History

While the subject History has a lower conductivity score across the Abitur and the Baccalauréat, it can only be left out in the Baccalauréat according it a Specialization score of 3. The fact that the subject has a lower conductivity score in the Abitur, but at the same time the subject History appears in all subject tracks, is related to restrictions that apply when the subject History is taken as a specialization, concerning the choice of other specialization or lower proficiency subjects to choose from.

#### ***Type 4: Sciences, Foreign Languages and Latin***

All remaining subjects in the group Sciences, except Math, share the same network representation with Spanish and Latin. These subjects also have the highest overall degree of choice as determined per my rationale, given the fact that some scores actually diminish the degree of choice, the higher they become, like conductivity. Figure 4.5 depicts the network representation of the Sciences group:

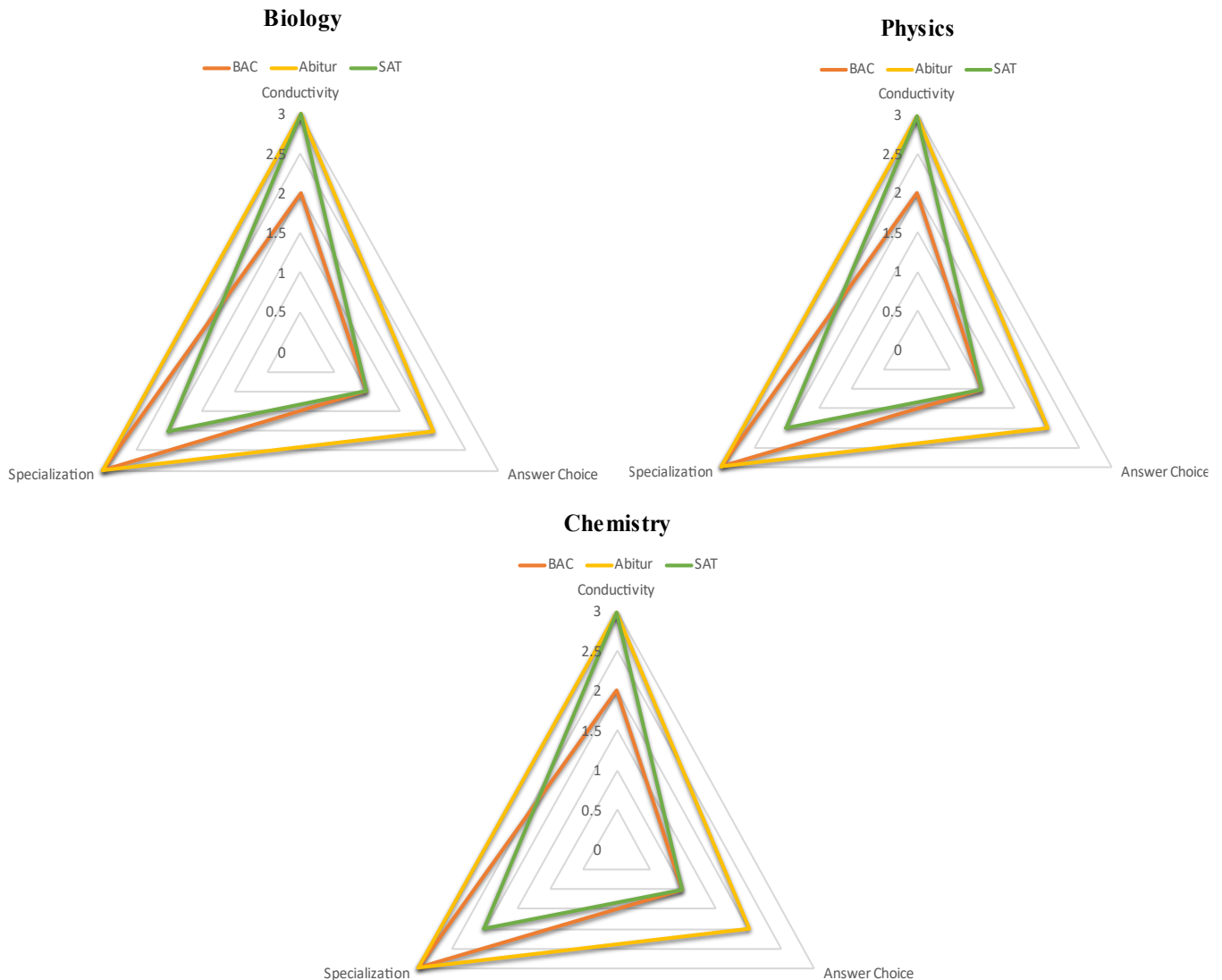


Figure 4.5. Representation of Choice: Sciences

All these representations are similar in terms of specialization, meaning all subjects included have the highest degree of choice in the Baccalauréat and the Abitur, meaning a student can choose to specialize in any of them, take them with a lower requirement, or leave them out completely. Like with Math, the conductivity score for the Baccalauréat is lower, while higher for the Abitur. This points towards different strategies in subject tracking between the two exams, and consequently, it might also point to different understandings of the roles that all these subjects play in the French and the German societies respectively.

The same applies to Figure 4.6. Spanish (as a representative of the only foreign language that all exams have in common) and ancient Latin are represented through the same network as the sciences:

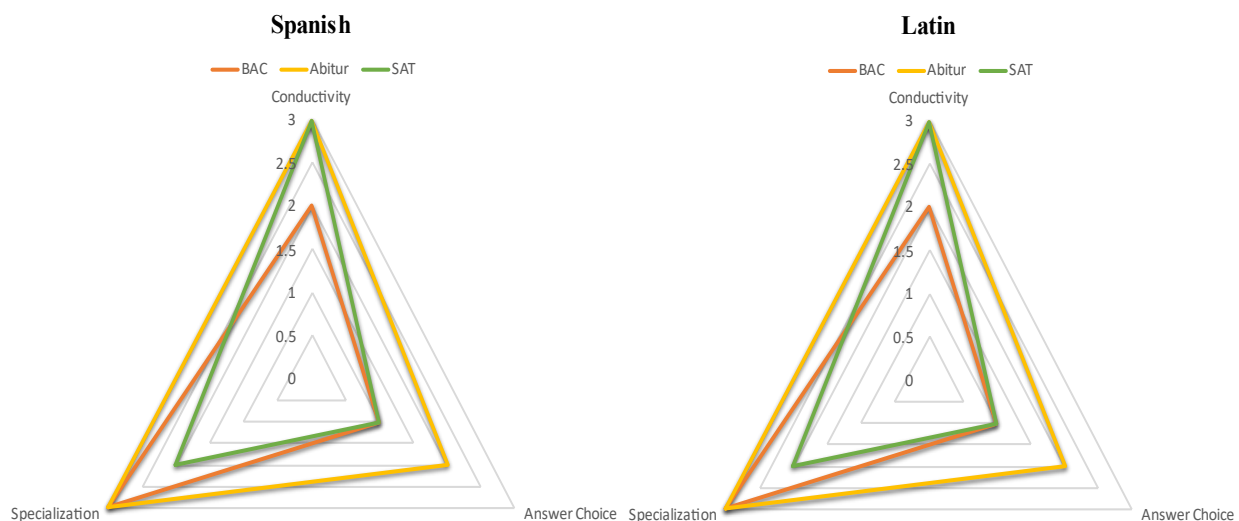


Figure 4.6. Representation of Choice: Foreign and Classic Languages

### Explanations: Availability and Degree of Specialization

All possible subjects of all three exams combined are listed in Table 3.5 in Chapter 3, the total available number ( $N$ ) from all three exams combined being 48 subjects. “Availability” is a network representation based on the table presented in Chapter 3; also here, exam content is



tabulated by subject with all three exams in mind. Table 3.5. in chapter 3 only manages in providing an overview of general availability.

It also needs to be mentioned that I focused only on the general stream of the BAC, for reasons of feasibility and comparability. This dimension is less relevant for the SAT as all its subjects are treated equally and therefore all SAT subjects received the same score. I made the choice to include the SAT in this dimension nevertheless, for reasons of better reference with the other two exams.

The degree of specialization is a score I gave based on the idea that a subject, which can be left out, specialized into or taken as a lower requirement version, receives the highest possible score. If one of the three statements was not true, I subtracted a point. This means that the degree of specialization score can agree between two exams for different reasons. This approach puts subjects like creative arts and IT to the top, as they can appear in all three formats (specialized, lower requirement, absent), hence have the highest representation of specialization as per my rationale, while subjects considered most important often received a lower score due to consequent lack of choice.

Given the structure of the SAT, all subjects available neither permit for specialization, nor do they permit for declassification, and each subject test of the SAT can potentially be left out. Hence the degree of specialization is a measurement which can only be applied to the BAC and the Abitur.

### **Explanations: Conductivity and Answer Choice**

Another element represented in Figures 3.2 to 3.6 is the amount of question combinations from which students can choose for each exam and for each subject. As such, answer choice is the realization of choice on the textual level of the exam sheets. The score given here is

straightforward as all question sheets in the Abitur and Baccalauréat either had (1) no choice, (2) two sets of questions to choose from, (3) or rarely also three sets of questions to choose from. For the SAT, I found no exam sheet with a choice in that matter for the students.

The Conductivity score is an expression of choice on the subject level, as it expresses the conductivity of one subject in combination with all others. For the rationales underlying choice, this actually could mean that the higher the conductivity of an exam subject, the less choice is given to a student and an exam taker to exclude the subject or not specialize in it. The results of this representation as depicted in Figures 4.2 to 4.6, focus on the total sum of possible subject combinations an infinite amount of test takers could potentially face. For that purpose, I assumed it best in Chapter 3 to calculate a value for each subject matter. Subjects with a lower conductivity score will potentially appear in less of the total sum of possible exam iterations. The score itself, though, refers to the maximum amounts of subjects with which subject in focus could possibly appear in combination. Finally to understand the restrains on choice measured through the conductivity score, it also became necessary for me to look at the different systems of subject tracking in the Baccalauréat and the Abitur.

For example, the Abitur guidelines for the exam in Lower Saxony, where my exam sample is taken from, know fifteen possible foreign languages and group three to four subjects as sciences, where sometimes, IT is or is not included. It recognizes two – three subjects in the category of Music and Art (theater is or is not included), and groups a total of 6 possible subjects under the category Religion / Philosophy/ Ethics, where students must choose one. The exam guidelines in Lower Saxony offer 5 distinct tracks, wherein each dictates the inclusion or exclusion of certain categories of subjects. The Baccalauréat as a comparison, groups 8 possible foreign languages (general stream) and 3 distinct tracks (L, ES, S), which each to some degree

excludes certain subject combinations. Although the new Baccalauréat (2021 onwards) does not track subjects in that manner anymore, the exam data I had access to still operated with the older rationales. Whether the new Baccalauréat agrees more choice and imposes less tracking cannot be determined here. The result in form of numerical scores on conductivity are included in Annex 3.

### **Findings: Representation of Content per Subject**

The second coding method of the exam data, network text analysis, focuses on the actual content of the exams. The following visualizations represent the adaptations of Carley's (1996) methodological framework. I created a thesaurus as a projection of already existing networks of exam content and as a provisional structure. For the thesaurus creation, I relied on a subject classification system called Library of Congress (*LOC*) classification, which is globally considered the most frequently used to categorize and describe subjects and subject content (Harris, 2013; Kumbhar, 2011).

Furthermore, the ideas underlying semantic networks as expressed by Popping (2000, 2003), and specifically the type – token relationship, find a logical representation in form of the tabulation of exam questions and exam responses. The dimensions refer to the themes of the questions as well as the mental operations of the responses. I located both in my methodology by underlying both either the representation of cultural capital (overt, cf. Bourdieu, 1986) or by underlying the representation of a form of mediated cognition (cf. Vygotsky, 1977).

### **Visualization of Semantic Networks**

With the following representations, the type, i.e. the subject is marked in red, while the token, i.e. the responses are marked in blue. Both representations carry a number value, which represents the number of times the topic appeared in questions or the number of times a certain

category of responses was required. Questions represent topics (social capital), while responses represent certain mental operations or artefacts, which are encoded through expectations for a correct response (artefact mediated cognition).

The following data is extracted from past exam samples, and all diagrams include citations or examples from the exam data, or the answer keys, to provide more transparency to the rationales, which informed the categorization of the content. For the SAT, I relied on The College Board (2013) which provided the official question sheets and responses for the SAT subject tests. For the Abitur, I relied on *Niedersächsisches Landesinstitut für schulische Qualitätsentwicklung* (2020, January 20), a public state website providing past exams in their entirety and official answer keys. For the Baccalauréat, I relied on *Ministère de l'Education Nationale et de la Jeunesse* (2020a), a public government website, for past exam sheets and on Scribr (2021, March 21) for answer keys.

### ***Native Language Literacy***

The following is a representation of the semantic networks I discovered in the form of a Sankey Diagram. To create this diagram and all the following ones I used the free software SankeyMatic (2021). As I described in Chapter 3, I took the high-profile subject test, where available (Abitur, Baccalauréat).

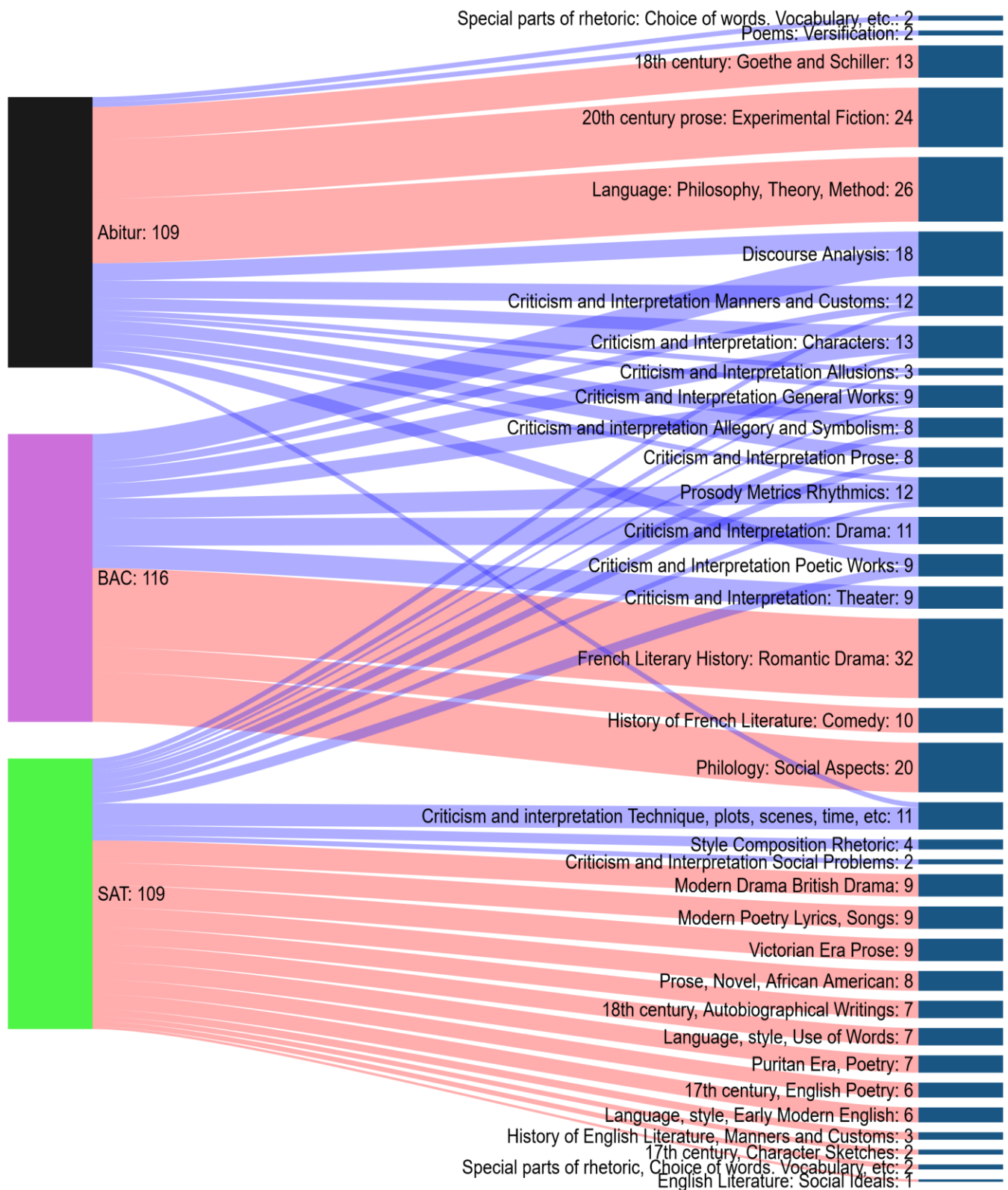


Figure 4.7. Representation of Content: Native Language Literacy

The direct comparison of the visual network representation of Native Language Literacy, reveals that the SAT appears much more fragmented than its two counterparts. Themes and operations are at the same time more numerous, as well as smaller in intensity. The questions in the Abitur and Baccalauréat exam sheet required many more mental operations to be carried out, which explains the higher intensity given to each of the exam topics.

### ***The Topics (red)***

The diagram in Figure 4.7 reveals that no topics were shared across the three exams. While the French exam sheet asked students to analyze literary production in the context of drama and arts, the German exam sheet focused on modern and 18th century prose.

The SAT united topics largely related to British literary history, with some impact of American contemporary political writing. In the context of 20th century political developments, exam takers were asked to analyze fragments of speeches.

### ***Mental Operations (blue)***

Several mental operations are similar across all three exams. To provide coding transparency, I am exemplifying some mental operations based on excerpts from the exam samples and answer keys.

### ***Criticism and Interpretation: Manners and Customs***

This mental operation represented as the most frequent in the Abitur exam and also appears several times within the SAT subject test. The following is an example of an exam question (question 37) from the SAT pertaining to this mental operation:

37. The simile of the hand (line 7) suggests
- (A) acceptance of change in social worlds.
  - (B) a rationale for a segregated social system
  - (C) a symbol of racial pride
  - (D) hard work as the basis for economic prosperity
  - (E) a physical basis for similarities and differences
- (SAT, English Subject Test)

Exam takers were required to read and interpret a short excerpt from modern prose and asked to interpret the events based on their understanding of the simile of the hand. The response to this question required exam takers to be fluent with the meaning and interpretation of manners and social customs at the time and in the social context in which this text was written in. The following excerpt is from a question in the Abitur and relates to the same operation. Unlike for the SAT question, the question in the Abitur was not multiple choice, hence I here quoted from the answer key. Within the document, the response relates to *Deutsch EA, Aufgabe 2, Frage 1*:

*(...) formulieren eine aufgabenbezogene Einleitung, etwa:*

*- Thematisierung des Erlebnisses des DDR-Bürgers C., der den Aktionismus von Konsumenten in einer westdeutschen Großstadt schildert; kritische Reflexion des Konsumexzesses.*

*(...) formulate an introduction related to the general topic of the question, for example:*

*- Thematize the experiences from the perspective of the fictional former East-German citizen C., who elaborates on consumer excess of Western capitalist consumers in a Western German metropolis; reflect critically on consumer excess.*

In the case of the Abitur question, an exam taker is required to have in depth knowledge of the disparities between socialism and capitalism, and between former Eastern and Western German societies in order to interpret correctly the fictional former East German character named C. and elaborate on his assumed experiences in a West German metropolis.

### ***History***

This analysis focuses on history as the exam subject in the SAT, as well as in the Abitur. Though the SAT subject test contains two entries of history, US American and World History, I failed to properly analyze the World History exam, for reasons that I will explain after comparing the exam questions of the US American subject test with those included in the Abitur. Furthermore, operations outnumber themes in this visualization as some questions and the steps to their respective solutions required more than a single operation at a time.

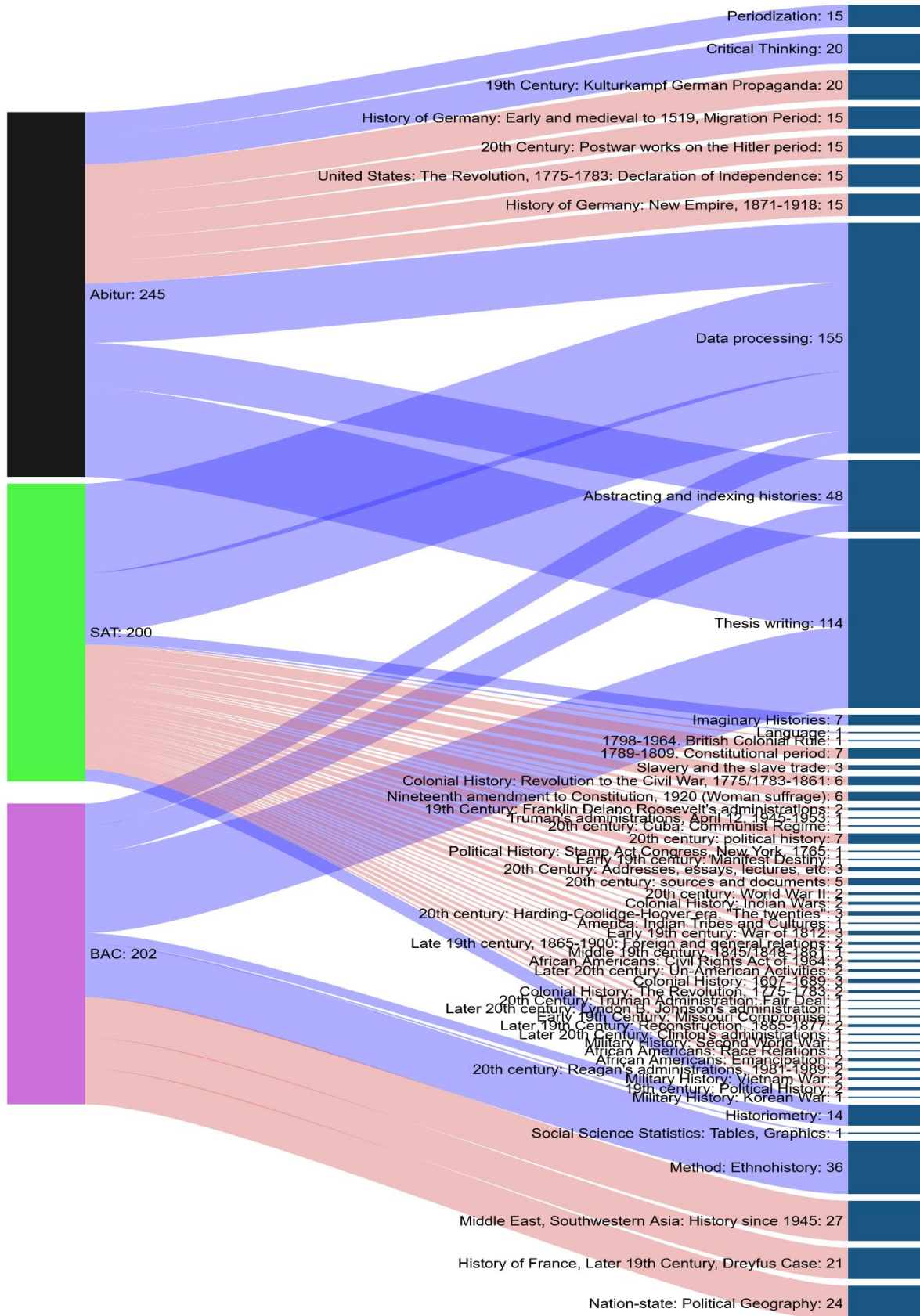


Figure 4.8. Representation of content: History



### ***The Topics (red)***

In this subject area no topics directly overlap in a comparison across all three exams. The topics of the Abitur are categorized into two sets of questions. The first set focuses on the ideas of nationalism throughout the German history from early medieval history, and the Germanic migration period to the rise of fascism under the Nazis. Interestingly, this elaboration on historical nationalism is juxtaposed with the rise of nationalism in the United States, which is expressed in Figure 4.8 through the category: *United States: The Revolution 1775 – 1783: Declaration of Independence*. A similar pattern of reflection on nationalism is depicted through the categories, which arise from the Baccalauréat questions. The exam questions focus on events related to antisemitic nationalism in France, the persecution of Jews and the creation of the Jewish nation state. At the same time, the questions also require students to have some historical geographical knowledge of Europe and the Middle East.

Making summative statements for the SAT is more difficult, as its topics are highly fractured and barely repeat or stretch over several question. A quick look at network representation of the SAT reveals, that the topics are far more numerous than with its German and French counterparts. That also leads to a high degree of fragmentation which could be linked to the nature of the SAT being 100% multiple choice, as opposed to the BAC and Abitur. The most, which can be said about the SAT topics would be that all of them focus on the modern standard history of the United States.

While the BAC and the Abitur require narrated responses, the length of which may vary from paragraph to full essays of multiple pages in length, the fragmented nature of the SAT exam does not allow for such depth.

### *Mental Operations (blue)*

For the sake of coding transparency, I reflect on the mental category *data processing* as shown in Figure 4.8., which is by far the most numerous represented across all three exams. Data processing is an operational category which I identified with exercises, that require the student to work with historical sources, primary sources, or snippets of such. Unlike another common SAT operational category *Abstracting and indexing history*, which does little more than ask the students to recall memorized information, Data processing requires a significantly more intense mental effort. The following is question 43 in the SAT exam sheet. In the original it starts with the depiction of a cover of the 1905 issue of *McLure's Magazine*, which I cannot show here for copyright reasons. Based on the Title and the abstract of contents depicted on the cover, the exam takers are required to respond to the following question:

The articles appearing in this 1905 issue of *McLure's Magazine* illustrate all of the following trends in the early twentieth-century United States EXCEPT:

- (A) Popular magazines were beginning to turn their attention to the issues of reform.
- (B) Reform of municipal city governments was a growing concern.
- (C) Exposure of monopolistic business practices was beginning to draw public attention.
- (D) Scientific methods were increasingly called on to lend credibility to all sorts of theories.
- (E) Reformers of both government and society received widespread support among the leading industrialists.

That question exemplifies the concept behind this mental operation, i.e. exam takers are required to abstract or index information provided to them in different forms, based on their historical knowledge. The following is an excerpt from the Abitur pertaining to the same mental operation from *Geschichte EA Aufgabe 1, Aufgabe 2*:

*Entstehung des Materials M1 nach der Reichseinigung „von oben“ (1871); national-konservative Position des im Material vertretenen Geschichtsbildes („Vaterlandsliebe“, Z. 19 und „Liebe zu Kaiser und Reich“, Z. 29) in Abgrenzung zu liberalen und demokratischen Vorstellungen, die ebenfalls das nationale Denken in Deutschland im 19. Jahrhundert prägten.*

Explain how material M1 came to be in context of the empire's pacification "from above" (1871); Mention the national-conservative position of the historical understanding

expressed in the material (“Love for the nation”, line 19 and “Love for emperor and empire”, line 29) and delineate this specific form of thought to liberal and democratic ideas, which also prevailed in 19<sup>th</sup> century Germany.

Like in the SAT example, that response to the Abitur question requires the exam taker, to abstract on a source material provided in the exam and to check their responses with their knowledge of particular aspects of German history, that are valued by the exam’s stakeholders.

### ***Mathematics***

For the math exams, the tables look slightly different, as operations heavily outweighed content as per the established definition. The relevance of content can arguably not be as well established, as for the other subjects, which I am going to discuss in Chapter 5.

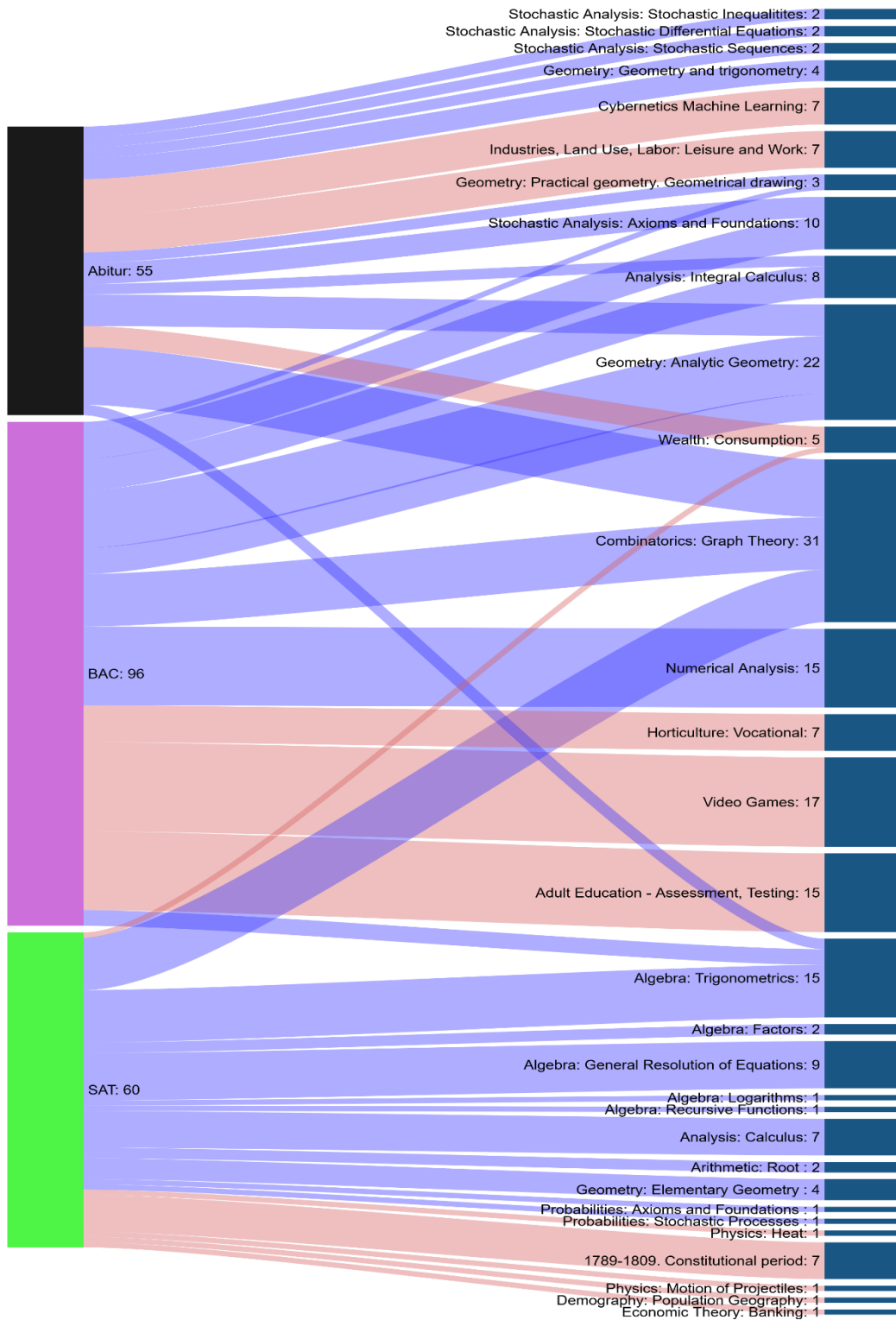


Figure 4.9. Representation of Content: Mathematics

### ***The Topics (red)***

A quick look at Figure 4.9 shows that no topics overlap in a comparison across all three exams. Unlike in the previous two subject tests, the Math exam has many exercises which do not relate to a topic and are purely made up of the operations, which predetermine its correct response. This fact is true across all three exams. The highest degree of topic embedding is still found in the Baccalauréat which contextualizes its required Math operations with allusions to Video Game problems, to Horticulture, and to Assessment in Adult Education.

### ***Mental Operations (blue)***

A look at the mental operations in Figure 4.9 shows that the Baccalauréat and the Abitur share a larger quantity of operations between each other than either is sharing with the SAT. Apart from that the analysis of the Math exam remains the least conclusive in relation to my coding rationale, as operations heavily outnumber the topics. What that means for the assumption of cultural or social capital embedded within the exam remains to be discussed. The following section will focus on analytical perspectives, which can be drawn from the ratio or the discrepancies of theme or topics versus mental operations.

### **Complexity Values for Each Exam**

An opportunity of analysis arose through measuring the ratio between the number of themes and the number of mental operations, which emerged through coding. I am naming this ratio *complexity value* as it directly relates to differences, which arise from designing an exam in pure multiple choice format, as opposed to free response format. While content is a remnant representation of objectified capital, the operations express differences in demand for mental operations. The latter aligns well with Vygotsky's (1978) concept of artefact, as theorized multiple times before. The exam responses turn into a societal "gesture" (Vygotsky et al., 1978),

which especially fulfills the control function of Vygotsky's theorem. The approach also meant that I sorted all presented exam content in two groups, the operational group and the content group, which enabled me to look at ratios of content and operational requirement.

The ratio is calculated by simply dividing the number of different operational requirements by the number of different content per exam. The result can be used as a signifier of depth and complexity of the exercises posed (see Chapter 3).

Table 4.29. Complexity Values for Each Exam

		Baccalauréat	Abitur	SAT
<b><u>History</u></b>	operational	6	5	5
	content	3	5	38
	<b>ratio</b>	<b>2</b>	<b>1</b>	<b>0.13</b>
<b><u>Native Language</u></b>	operational	7	11	13
	content	3	3	11
	<b>ratio</b>	<b>2.33</b>	<b>3.67</b>	<b>1.18</b>
<b><u>Mathematics</u></b>	operational	7	10	13
	content	3	3	6
	<b>ratio</b>	<b>2.33</b>	<b>3.33</b>	<b>2.17</b>

The generated score is directly proportionate to the degree of fragmentation, which I discussed earlier as being a property that arises from a consequence of an exam being multiple choice. The higher the ratio, the more mental operations are required to solve the question, which also means that larger parts of the exam are *concealed*. In reference to Vygotsky's (1978) artifact mediated cognition, this means that exams, which rely on free responses and which have a higher complexity value, are also imposing higher expectations on the cultural and mental developments of their students, as exam takers are required to discern the totality of what is required for the perfect grade based on the response. I discuss the exact implications of the complexity score in the following Chapter 5, while also I also argue for the importance of the complexity score and the coding rationale which I underlaid.

## **Chapter Five. Discussion and Further Research**

The results of the previous chapters revealed textual and meta-textual information about each of the Abitur, the SAT and the Baccalauréat, with reference to one another and to the societal contexts in which they were designed and executed. While the narrative synthesis discovered institutional capital held by curricular authorities linked to each of the exams Abitur, Baccalauréat and SAT, the network text analysis complemented textual reading of exam content with a constructive analysis of concealed networks. Earlier in Chapter 3, I termed these networks “hidden curriculum” with reference to Pinar (2004). Doing so also facilitated a new application of network text analysis, which required to translate previously established methodologies on the analysis of text to the analysis of textual and non-textual information. In the following, I will discuss three significant findings of both methods, narrative synthesis and network text analysis in relation to each other, before I discuss further insights I gained through coding, as well as the limitations of my analyses.

In general, it needs to be said that there are many more results, which would deserve a discussion, beyond what I can provide in this chapter. On one side, I focus on findings that coincide as a result of both analyses, as I believe such correlations are valuable as they give validity and add significance to the methodological mix that I employ here. Beyond that I will also reflect on further realizations, which I gained through coding and could not express as a response to either of the research questions. These sections will focus on limitations and possibilities for further research.

### **Abitur: Insights into Maturity and Choice**

While coding, I realized an interesting connection between the narrative theme of *Reife* in curricular documents and certain findings of the analysis of representation of choice in the

context of the formal representations of choice. I have a unique, personal understanding of how the idea of maturity translates into the life of school children and teenagers and I would like to assert my positionality while shortly discussing this link:

Maturity or *Reife* has received extensive historical and contemporary attention in the German speaking literature (Eberhardt, 2006; Kruse, 2017; Pöggler, 1964) and arguably even has a German literary genre linked to it, which is called *Bildungsroman* (Selbmann, 2016; Schrader, 2019). The *Bildungsroman* is a categorization of prose, which includes works written during the literary period of German Romanticism. A classic publication on the characteristics of the German *Bildungsroman* is provided by Selbmann (1994), where he defines some of the core characteristics of what maturity and the path towards adulthood mean in this context on page 2. He explains that reaching adulthood in this literary genre is closely linked to a journey, which helps an adolescent learner to discover and express their position in the world and in society and allows the subject to discover and specialize according to their natural talents and capacities (Selbmann, 1994, pp. 2 – 6).

As a former pupil of the German school system, I had the feeling that many of the unique assumptions which are expressed through the concept of journeying and finding a place in society do still persist, albeit in different forms and not necessarily through a journey as depicted in a *Bildungsroman*, for which a prominent example would be *Wilhelm Meisters Lehrjahre I – III*, by Johann Wolfgang von Goethe. Until today, reaching adulthood in Germany means to choose a path or a track very early on, which might arguably be best symbolized through the fact that Germany has a school system which categorizes children into three groups. After grade 4 children are recommended into a further educational track, either *Hauptschule*, *Realschule* or *Gymnasium* [Note: the three tracks are meant to classify students according to inherent academic



ability]. Out of the three, only the *Gymnasium* permits to take the Abitur and consequently pursue university education. Specialization and tracking play a large role in the German education system, and I perceive some results of both, the narrative synthesis, as well as the network text analysis related to this notion.

When faced with the translation of the theme *Reife* into English, I realized how difficult the task was, if alone for the discrepancies between what the passages of the narrative synthesis revealed, in comparison to how much I actually understood about the implications reading the documents. I provide some of the reflections on this topic above. As a coder, I could not interpret more into the results than what my coding rationale permitted, but as a researcher, I felt that the curricular documents, especially when the theme *Reife* was concerned remained too vague and too implicit, warranting a deeper reading and a better translation, than what I could provide.

On the other hand, based on the results of the representations of choice of the network text analysis in chapter 4, where the Abitur appeared as the exam with the most extensive subject tracking and the highest specialization scores for all subjects, it stood out in that aspect across the three exams. A high degree of specialization coincides with the narrative theme *Reife*, but also related to further expectations, known to me concerning the ideas of maturity in German society. This correlation between the results of both methods would warrant and inspire further research into the extent to which ideas of maturity in the form of *Reife* are present as capital in German educational institutions and their published curricular documents.

### **Baccalauréat: The Importance of Writing**

In this section I discuss another striking correlation, that presented itself to me when reflecting on the coding results of the narrative synthesis in comparison with the network text analysis concerning the Baccalauréat. The semantic representation in Figure 4.8 reveals the outstanding character of an operation I named *written thesis* in the History exam. On the other

hand, the narrative synthesis in Chapter 4 revealed, that the cultural competence I named *Production, Interpretation and Manipulation of Texts*, is by far more significantly represented in curricular documents of the Baccalauréat, with a keyness score of more than double that of its counterparts.

When coding the exam sheet of the 2019 Baccalauréat, I realized that it was unique, as it did not include any form of supplementary information. More than half of the exam sheet did not even ask a specific question to guide the student. The following excerpt represents roughly half of the Baccalauréat exam sheet in the subject History, the other half being about historio-geographical knowledge as shown in the corresponding representation of content in chapter 4.

The direct translation is provided below in italic:

## **HISTOIRE**

### ***Composition***

*Le candidat traite l'un des sujets proposés.*

***Sujet 1*** – *Le Proche et Moyen-Orient, un foyer de conflits depuis la fin de la Seconde Guerre mondiale.*

***Sujet 2*** – *En vous appuyant sur les exemples étudiés au cours de l'année, vous traiterez le sujet suivant : médias et opinion publique dans les grandes crises politiques en France depuis l'Affaire Dreyfus.*

(Ministère de l'Education Nationale et de la Jeunesse, 2020a)

## **HISTORY**

### **Written Exam**

The exam taker treats one of the two proposed topics.

**Topic 1** – The Near and Middle East, a powder keg for conflicts since the end of World War II.

**Topic 2** – Based on the examples, you studied during the past academic year, you are to treat the following topic: relationship between media and public opinion during times of major political crises in France since the Dreyfus Affair.

The extract above reveals that the exam sheet does provide themes instead of questions, and beyond these topics and statements no supplemental material is given. Beyond having the factual knowledge necessary to respond, students need to display considerable skills in structuring their written response. The problem of coherence arises, and students need to find ways to compose a

text, which is factual, but at the same time stringent and presents the required elements of the response in a chronological order. Topic 1 resembles a hypothesis more than it does an exam question, a hypothesis, which could be the basis for a format of essay commonly called “position paper”. With only three topics, relating to three questions, the BAC expects students to write a literal thesis to respond to one of the two questions, a fact, which also translated into the highest complexity score in the subject history.

Repositioning these finding back into the framework of my research, the correlating results above and the excerpts from the source material support the notion that an analysis of social and cultural capital can lead to relevant results, and further to explanations of structures, which are considered axiom. I did not find expressed a justification or a reflection on why this specific exam adds such complexity to the subject of history, but I was here able to reflect on some of the origins for these decisions through narrative synthesis of curricular documents over time. This approach might continue to produce relevant information and serve to break apart curricular certainties, which are considered axiom in our current educational environments.

### **SAT: Fractured Knowledge or Accessibility?**

The assumption pertaining to the network category of complexity, which I employed in Chapter 4, is that it expresses how much of the necessary knowledge for the ideal response is mediated through the questions, and how much of it is concealed. I believe it is not by accident, that the exams with the highest per subject complexity scores (the Abitur and the Baccalauréat) also turned out both having hard to translate mental-cultural components in the narrative synthesis of their curricular texts. The difficulty with translation did not simply arise from finding correct descriptions, but also from the lack of space and time to reflect on several implications. Having grown up or having studied in those respective societies I am aware of that, but I cannot express it through the excerpts to which I am limited through coding rationales.

Across the board, the SAT exam had the lowest complexity scores, which might make it the most accessible. Vygotsky (1976) implied that educational artefacts are the results of continued acculturation and socialization of the child (p. 36 ff.). A fact that also permits the society to control the learning of its children through a monopoly on the tools which students need to unlock relevant knowledge, and control of the ways in which they develop their own perspective over the course of their adolescence. A higher complexity score comes with a higher degree of concealed expectation for exam takers, towards the possession of a larger arsenal of “tools” (as used by Vygotsky, 1976).

Given that a higher degree of complexity would also point towards a higher degree of behavioral policing through the stakeholders, the fact that the SAT exam has the lowest complexity score might make it the most accessible to students of different educational and socio-cultural backgrounds. This assumption correlates with some of the narrative themes expressed through the analysis of curricular documents to the SAT in chapter 4, particularly the social competences *Readiness* and *Social Assimilation* (cf. Table 4.8). These categories are different from the other themes in that they express a social competence of succeeding in the SAT exam, but posit this competence not from a position of authority, but from one of self-criticism. The following quote is one of the examples I presented in Chapter 4 on *Readiness*:

We can’t continue to allow vast numbers of our country’s students to fall behind academically. It’s therefore critical that we do everything possible to ensure that all students are on a trajectory to gain meaningful access to postsecondary courses and workforce training programs, complete degrees and certifications, and participate successfully in an increasingly competitive and fluid global economy. (2015, Standards Redesigned SAT)

That quote expresses an assumed ability of the SAT exam, one that could be summarized as being able to place students on meaningful trajectories. On the one hand, the students taking the exam must hence display that they are ready for the trajectory, for which they apply, but on

the other hand, the exam must be able to allow the recognition of as many student trajectories as possible. The latter means that the exam must have a wide-ranging accessibility, to meet minority students and students with as many educational backgrounds as possible. Similarly, the theme *Social Factors of Learning / Social Stratification* as described in Table 4.28, recognizes students with certain socioeconomic backgrounds, but also that the test needs to meet students with as many diverse backgrounds as necessary.

Returning to the results of the network representation *complexity*: The fact that the SAT is completely designed as a multiple-choice test and therefore low in terms of complexity, might actually help face the issues expressed above. Exam questions, which do not conceal excessive amounts of implications as the example from the Baccalauréat above, it might help meet students from more diverse socioeconomic backgrounds that display larger differences in acculturation. Whether or not these findings correlate would need to be further discussed, beyond what is possible here. The conclusion of the last three sections do not exhaust the possibilities of comparison and discussion which arise from my results. More research is necessary and warranted.

### **Different Approaches to Discovering Cultural Capital and Cultural Bias**

In the previous three sections, I mentioned the network representation complexity, and I spelled out a few implications which come with a high complexity score in relation to the discussion. Next, based on the discussion in the previous sections, I discuss whether my approach can help reveal cultural bias and to what extent each exam might suffer from bias.

The complexity value was a final analytical step I decided to include, when looking at the stark differences on first sight across the exams in terms of length and question content. Sometimes the Baccalauréat just consisted of one to two questions, while exam takers,

depending on the subject, had more than 300 minutes (5 hours) to respond. In comparison, the SAT, depending on the subject, accords 60 minutes for 60 to 90 exam questions consisting of multiple choice.

Considering the theorem of artifact mediated cognition, a higher complexity score then reveals something about social control. Like an iceberg, much more is hidden under the ocean's surface, when we have high complexity values. These hidden parts are neither realized nor expressed in the question, but require the students to nonetheless apply the correct arsenal of tools in order to acquire the perfect grade. It is not by accident, that the exams with the highest per subject complexity scores (the Abitur and the Baccalauréat) also turned out both having a strong and hard to translate mental-cultural component in the narrative synthesis of their curricular texts.

With the concerns expressed on cultural bias in Chapter 2 of this research in mind, it could be argued, that the exams, which conceal large portions of their required response behind culturally mediated artefacts, are those with the highest cultural bias. Societal, cultural control could be assumed to be much more expressed here, as hints towards the selection of tools are scarce. Additional difficulty, especially for children that are less socially or culturally affluent?

On the other hand, I could look at the representation of content and ask how similar are the Baccalauréat, the Abitur and the SAT in terms of their representation of content? Given the data visualization design I chose, it becomes clearly visible that there are differences between the three exams, not just in terms of their complexity value, but also in terms of the topics, which are represented through the exam questions. I revealed topics based on the semantic representation of content as being reflected by each exam question. Like previously (Chapter 3) explained, the semantic representation I employ here operates with the assumption that exam data is the

question posed, as well as the complete response to the question which is deemed perfect by the exam's authority. Following that rationale, my results position questions as type, while the necessary operations to get to the ideal response are positioned as token, in application of Popping's (2000, 2003) methodology.

Figures 4.7 to 4.9 revealed that the exams share close to no topics with one another. Figure 4.7 described native language literacy, and the topics for each exam are within their expected national frameworks. The SAT topics somewhat deviate in their sum from that statement, as they arguably depict a measure of diversity through the inclusion of different genres, from prose, to poetry to political speeches, through the inclusion of older British and US-American themes with some exercises treating of political texts on the African American civil rights movement and women's suffrage. The SAT is able to include so many topics in one exam through its multiple-choice format, where the Abitur and the Baccalauréat are very limited through their treatment of a few topics, but those in-depth.

The idea of including as many topics as possible was most prevalent with the SAT History exam. Figure 4.8 describes all topics included in the SAT, and they all are confined within US or British history. At this point I need to state that the SAT offers two separate subject tests, one titled US History and the other titled World History. I was unable to include the exam on World History, as the coding with my rationales proved to be impossible. The exam on World History of the SAT did not compare with the other exam sheets on history, as it appeared to belong to a different discipline entirely. I had to stop coding the exam sheet, when the majority of questions revealed topics relating to cultural- or social studies. Furthermore, a semantic representation of the exam sheet would have not fit into any visualization such as I used in Figures 4.7 to 4.9 since the next to no topic appeared twice. The SAT World History exam sheet

appeared as an assortment of 90 questions, scratching the surface of contemporary and historical topics relating to the whole globe.

It remains questionable, whether the College Board's approach on diversifying their history exam serves its purpose if none of the people and cultures discussed therein is given nearly enough attention and complexity (in the sense of the network representation of complexity), to remotely compare it to the "real" history exam within the SAT or with the history exams of the Baccalauréat and the Abitur. Figure 4.8 reveals that the Baccalauréat exam sheet introduces a focus on the historical development of certain aspects of the French society, in relation to historical and political developments in the Near and Middle East. With the highest degree of complexity and the highest sums of operation required to respond to each exam question, it can be expected that the exam takers' knowledge on international history will be limited but extremely focused on some aspects of a people and a culture that is not the French.

Another aspect of the Baccalauréat's Math exam is revealed through the network representation in Figure 4.8. Looking at the three diagrams for the three subjects I analyzed, the visualization of Mathematics distinguishes itself significantly from History and Native Language Literacy across all three exams. Topics assume, if any, marginal positions within the Math exam and the focus completely lies on the mental operations. Though none of the three exams appears to value Math in application to content matter, the Baccalauréat exam sheet revealed the most relevance to popular culture, albeit in a limited fashion. Research is done on linking Math exams with sociocultural content and there is already considerable literature on the benefits of such, for example "social justice math" or "culturally relevant mathematics" (Bartell, 2013; Gutstein, 2003; Gutstein & Peterson, 2005; Leonhard et Al., 2010). There is unfulfilled potential across all three exams to create Math problems going beyond the bare operations.



## Contributions Towards a Pedagogy of Assessment

The previous sections lead to several assumptions, about how exams reflect the ways in which the society from which they originate, thinks about assessment. The institution of assessment in general is called under scrutiny through the wholistic perspective, which I impose on the understanding of exams. Many of the findings I already presented enrich certain categories from the framework on pedagogy of assessment by Müller and Schmidt (2009), which locates assessment in a variety of different structures outside of the exam paper (see Chapter 2). The sampling procedure of curricular texts for narrative synthesis would not have been possible without reason to assume that assessment represents ways through which a society controls the access to social, cultural and economic resources of its adolescent generations. The notion is expressed by Müller and Schmidt (2009) in the context of *initiation, distribution of status* and *legitimization* (translated from German, pp. 23-28).

The reasonable assumptions on assessment expressed here made the search for social and cultural capital within exam texts worthwhile in the first place, as I could hope to learn more about what hidden curricula are working behind institutionalized assessment. The method of network text analysis further succeeded in providing me with a toolkit to reveal some of the open and covert relationships which the exam entertains with assessment as a social, educational institution. I came to the realization that the way in which stakeholders group and categorize subjects together, holds more analytical potential, than I am able to reveal at this point.

One particular question remains unanswered here: What are subjects and why are some subjects categorized into tracks in the Baccalauréat and the Abitur? Based on the network representation alone, I came to understand subjects in exams as being controlled by certain properties relating to choice and content, but I could not reveal, why the stakeholders for each

exam decided to track certain subjects the way they did or found it important to let students specialize or leave out subjects, while they could not do so with others. I cannot provide a conclusive response, but the notion of control as expressed in the previous section provides a lead.

### **SAT: the English Subject Test and Recent Developments**

This section continues a discussion I began earlier, when I questioned the forms through which cultural bias can be revealed as a result of the network representation of exam subjects. In this section, I focus on some results of the narrative synthesis, more specifically on *Communicative Competence*. In Table 4.2 I compare how curricular documents express difference when this competence is concerned. While the Abitur and the Baccalauréat both abundantly stress the importance of foreign, classic and native language studies to hone Communicative Competence, curricular documents related to the SAT refer exclusively to the command and mastery of English as standard British or American English.

This difference becomes more significant in light of recent developments. As stated on the College Board Website (College Board, 2019a), the SAT subject tests will no longer be offered after June of this year, 2021 for American students. As stated on their website, the subject tests will continue throughout this year for international students, as they are “(...) are used internationally for a wider variety of purposes” (College Board, 2021). Though the College Board has not published a statement concerning the years beyond 2022, they do make reference to the popularity of the SAT subject tests internationally. Henceforth international students will be subjected to these exams exclusively. In light of the focus on standard English as communicative competence, these questions are raised: Will subject tests continue to be used internationally beyond 2021 and if so would they be adapted to measure communicative competence differently? Furthermore, the College Board will continue to administer the SAT in

a different form for American students. To what extent will they use the already existing SAT subject tests?

These questions matter for further discussion in relation to my analyses as the SAT and the high numbers of foreign exam takers as laid out in Chapter 2. The SAT already serves as a gatekeeper for migration with a high number of international exam takers. There is a danger that the exam might lack even more transparency in its newest developments. Adding to that concern is the fact that subject tests apparently have not been revised since 2013. When looking for most recent exam materials, I relied on a 2013 publication by the College Board for my sampling as these are even today, the recommended preparation materials for the exam.

### **Insights and Limitations: 250 Years of Texts Across Three Languages**

The narrative synthesis of curricular writing relied on a theoretical framework of cultural and social capital developed by Bourdieu (1986) in which he describes capital to become embodied by agents and members of a society through the repeated exchange and formalization of skill (pp. 240-250). Once social or cultural capital has gained an institutionalized endorsement, it can be exchanged like a currency, but less so questioned, discerned or reflected. Inspired by the idea to deconstruct assessment in this research, I decided to sample curricular documents on all three exams based on a longer period of time, hoping to be able to discover patterns, which repeated and persisted in one form or another.

During the sampling, I found rationales which helped me identify curricular documents related to the Abitur beginning from the 18th century and to the Baccalauréat beginning from the early 19th century. The inclusion of these documents posed several conceptual and practical challenges. The German nation did not exist until 1875, so I relied on the assessment of chronologists like Bölling (2010) to identify and retrace the continuum leading to the Abitur as it

is administered today. Similar problems did not exist with the curricular documents relating to the Baccalauréat, as the earliest inclusions were already published by direct predecessors of political entities, which persist in today's France.

When coding the inclusions of my sample, I was faced with limitations relating to the state of some of the documents and limitations of optical character recognition software.

Although I restored some of the text by hand, I did not succeed in doing so for all the content of all the documents included in my sample. The tool Sketchengine, which I used for coding, produced some false results, linked to artefacts or misspellings of words, which I eventually cleared from the final list of keyness.

Unlike with the SAT data, the inductive coding of the Abitur sample was met with a few additional challenges as well, one of which is linked to the statistics underlying keyness.

Whenever a term or a grouping of terms occurs more frequently in the sample corpus than within the reference corpus, these terms receive a higher score, hence have higher keyness. That mechanism statistically relies on log-likelihood procedures (Chapter 3) and assures high distinctive value of the results, pertaining to the content of the texts which are included in the sample. The problem I faced is due to the standard German language undergoing significant shifts in grammar and vocabulary since the 18th century. As a result, most terms or term combinations, which are no longer used or different in modern standard German (of which the reference corpus draws its words from), made it into the highest positions. Those terms and term combinations received the highest keyness scores due to their statistically established “uniqueness” in comparison with the modern standard German reference corpus.

An interesting insight, which warrants discussion is the ratio of types of texts I came to include for each exam. In Figure 4.1 in Chapter 4, I depicted the types of texts, which eventually

made it into my final sample after systematically reducing the sample size. I categorized the texts into three groups, one being published by the governing entity, one a result the impact factor search, and another guided by search within the aforementioned chronicles. Comparing the exams with one another, it becomes apparent, that only the curricular documents relating to the SAT included results from the impact factor search. The reason for that being that statements about the Abitur and the Baccalauréat are to some extent owned by the stakeholder, being the government administrative body, the French Ministry of Education or the German *Kultusministerkonferenz der Länder / des Bundes*.

Though the French Ministry of Education has its own publishing house (of sorts, cf. Ministère, de l'Education Nationale et de la Jeunesse, 2020b), the documents which I found relevant at first sight, did not provide relevant information for a reflection on the structure and the purpose of the exam within society. The German stakeholders do not maintain a publishing service; they utilize a third party, scientific peer review, if any at all. Furthermore, all contemporary documents I included in the final sample for the Baccalauréat and for the Abitur, did not agree to many conventions of transparent scientific writing, as they neither included, nor referenced outside research. Their writing was authoritative, which made me initially refer to them as authoritative documents at the beginning of chapter 3. Some of the observations I describe here lead to several incentives for further research, one being the question of transparency and access, which also posed itself when sampling for exam documents and answer keys.

### **Insights and Limitations: Access and Authority**

Access was an issue while searching and sampling for complete exam data and answer keys for all three exams. Sampling for each exam was met with different challenges. The College

Board does not freely publish complete exam sheets of previous SAT exams. The website provides samples of authentic exam questions and responses, but none of which was presented in the form of a complete test, as test takers would have in front of them (College Board, 2021b). The practice tests did not include instructions for exam takers, they did not include a mention of the time given or the number of total questions, which needed to be treated. Although the latter might not be necessary to prepare for the exams, these information are key to my analysis, which goes beyond the expressed structures of content. To receive complete exam sheets, I had to resort to a publication from 2011/2013 which is still referenced as the most recent complete preparatory material on the College Board (2021b) website. Beyond that, the website refers students to private, paid exam preparation classes, offered by different institutions around the world. I could not find out, whether these private preparatory institutions have access to more recent exam material, nor could I find scientific publications treating specifically these matters of access here presented. Further critical inquiry would be warranted, especially as the role of the SAT remains elevated internationally, as doorkeeper to resources and educational access as laid out in section 1 of Chapter 2.

The Abitur is designed to be a public good, it is funded and paid for by tax money (cf. Kultusministerkonferenz, 2020c) and should as such be available and accessible to the public. Nevertheless, when sampling my exam data, I faced several issues. The Abitur does follow the same guidelines throughout Germany, but the responsibility for the concrete execution of these guidelines lies with the authority of each German state, as well as with the ownership of each annual iteration of the exam. When sampling, I came across large differences in the treatment of exam data across different states. Of all 16 German states, only four (Bavaria, Northrhine-Westfalia, Hesse and Lower-Saxony) made past exam data available to future and current exam

takers. The problem is that access there was given solely to students and teachers of middle and high-schools. Without an affiliation to a German public school, I would have had to buy past Abitur exam sheets from a private publishing company. It appears common for German states and the educational stakeholders therein, to sell off publishing rights to past exam data and answer sheets to private publishing firms. The following is taken from an email conversation I had with a representative (name removed) of the State Ministry of Education of the state Northrhine-Westfalia from 09.21.2020:

*Allerdings stellt das Ministerium für Schule und Bildung des Landes NRW auf der Grundlage vertraglicher Vereinbarungen mit dem Dachverband der Bildungsmedien (VBM) den Schulbuchverlagen nach Abschluss des Prüfungsverfahrens Prüfungsaufgaben in einigen Fächern zu Publikationszwecken zur Verfügung. Sie finden daher viele der Prüfungsaufgaben bei den gängigen deutschen Lehrbuchverlagen.*

The Ministry of Education of the state Northrhine-Westfalia entertains publishing agreements with the “Dachverband der Bildungsmedien” (VBM) [a union of German educational publishing houses] on past finished Abitur exam sheets. You will find many of the requested exam sheets with established German educational publishers.

The conversation and the fact, that a public, tax funded good is trademarked for the economical gain of state agencies is highly problematic on different levels. For the purpose of my discussion here, it suffices to say that it symbolizes obstacles to access, albeit not always for students, but for the public and for researchers who want to reflect upon, analyze, and question the exams and the exam content. Further critical research needed to be done as to what extent this limitation to access is intentional and a defense mechanism towards change is possible. The only German state, which eventually provided both exam sheets and answer keys publicly without restrictions to access was the state of Lower-Saxony, where I obtained my exam samples.

The Baccalauréat is a public exam in France and as the German counterpart, it is funded through tax money under the jurisdiction of the French Ministry of Education. As the exam is completely centrally organized, which I discussed in Chapter 3, I did not have to navigate state

by state regulations as in Germany. Unlike in Germany as well, past exam data is freely available and published on the Ministry of Education's website on which I relied for one half of my sample. Where the problem arose though, was getting access to the other half, i.e. to official answers. Though guidelines on grading for teachers exist, the Ministry appears to treat answer keys differently from the exam sheets, which are made available to the public (Ministère, de l'Education Nationale et de la Jeunesse, 2020b). Although, as here established, an official, endorsed response is just as much part of exam data as the question sheet, I was not able to find and sample official answer keys. I finally received answer sheets to the Baccalauréat exam questions of 2019, but these answer sheets did not identify an author, nor were they from a reliable source, as I received them from an educational files sharing service, Scribbr (Scribbr, 2020, December 8). Although they appeared thorough and similarly detailed as the official answer sheets to the Abitur, this fact presented a major limitation to my sampling and analysis. I discussed this in Chapter 3; and that lack of reliability also led me to the decision of only including exam subjects, where I felt confident enough to critically verify the information given on the answer sheets, based on my own subject knowledge. Given the fact that France's open access to past exam data conceals the crucial part of the correct, endorsed response, warrants further research into the implications. Given the theoretical framework of artefact mediated cognition, one could assume that this leads to a disproportionate bias for students from non-academic backgrounds, minorities, or lower socio-economic strata of the society.

### **Insights and Limitations: Contradicting Axis of Choice?**

Another limitation, which deserves attention at this point, is related to the results of the representation of choice, which I presented in Chapter 4. The diagrams I employed in Chapter 4 (Figure 4.2 to Figure 4.6) to represent choice as a social and cultural capital, express seemingly



contrasting notions. The higher the conductivity score for each subject, the less degree of freedom (of choice) do students have. In concrete terms that means that all students of the Baccalauréat will have French at native language literacy represented in their exams, no matter their specializations; and that means for all exam takers of the Abitur, that they have no combination available that allows them to opt out of Math.

The other two dimensions presented in Figure 4.2 to 4.6 (specialization and availability, answer choice) actually express degrees of freedom, therefore one would assume here, that higher scores represent higher degrees of choice. Although it remains to be discussed whether that is correct, given choice as social capital. With Boudieu's (1978, 1986) concept of institutionalized social and cultural capital in mind, I can argue for the inclusion of all three categories into a diagram together. Especially when looking at section 1 of this chapter, choice cannot always be universally defined as freedom.

In the context of the German theme of *Reife*, or maturity, choice is not freedom when it relates to the consequences of specialization. In setting children and teenagers in front of a decision to follow a schooling track early on, or a subject specialization later in their middle school, the essence of this choice is not comprehended through the freedom to have it, but through the severity of the consequences it brings with it. The consequences attributed to certain choices in the tracked German educational system are severe, school tracks taken after grade 4 often determine whether the child has any hope of acquiring an academic degree at an ordinary university and choosing subjects for the Abitur in relation, determine possible subjects of study at a university.

Hence the essence of choice in this context is not the freedom to choose, but the freedom that is taken away through a certain choice to specialize, or not follow an educational track. The

notion of taking away the freedom to choose is also expressed through the impact, which tracking subjects into groups and categories has on the mathematical formula which I used to calculate conductivity (see Chapter 3). Tracking led to a reduction of the conductivity score, just as much as grouping one subject together with a number of other subjects into, for example: foreign languages, sciences, religion/ethics etc. Given the rationales I presented, it remains to be further discussed what choice actually entails and how the network representation of choice as a cultural or social competence reveals layers of definitions, which were not visible previously.

### **Différance**

At the end, I would like to discuss and return back to the theoretical implications of deconstruction, which I mentioned at the beginning of my research, and which informed the title of this work, a “deconstruction of assessment”. As stated at the beginning, I came to title my research a deconstruction, for lack of better ways to describe what I am doing here. The methods are empirical in nature and the results lead to several different implications, which I did not even succeed to discuss in their entirety because of the limitations for this research project. Depending on the reading of my results, several more implications are imaginable, and this is where I was remembered of Derrida and some of the aspects of the concept on deconstruction, which continue to be developed and discussed after his death, until today (Derrida, 2020). One mention, which reoccurs with Derrida (1972) is the notion of *différance*. Based on his discussion (Derrida, 1972, p. 3) of the term, he primarily devised it as an oxymoron, which expressed challenges he faced in the literary reading and sensemaking, which eventually came to challenge ideas of essence, of truth, and interpretation. The term is grammatically oxymoronic, as it expresses an impossible (or incorrect) nominalization of a participial progressive form of the verb *to differentiate*. The term not just came to express contradictions of essential readings of texts and

unavoidable dissolutions of essence thereof, when embedded into poststructuralism, but also to help Derrida (1972) to express the necessity to do the contradictory, despite all.

Given my analysis of exam data, I was faced at first with a seemingly impossible task to come to a critical understanding of hidden or concealed structures, which inform assessment as an educational institution, only based on textual material, which does not intend to critically analyze assessment, but to give it essence and structure. It required that I adapt methods, initially designed for a structural analysis of text and society, to produce results that go beyond structure; furthermore I was also required to apply theoretical frameworks, which have never been applied in such a sense before. I managed to create provisional structures, whose existence can only be validated if certain assumptions of the theoretical framework, namely on the creation and forms of cultural and social capital, and on the socio-mental workings of child development, prove to be valid. I created an empirical network analysis, which relied on the constant reading of the present, as much as on the reading of the concealed in the same cognitive sphere. This research and its many implications which I have not been able to fully discuss, embody the first perceived paradoxical exercises, which also led Derrida (1972) to come up with the notion of *différance*, hence I dared to include deconstruction in the title, in lack of a better way to describe it.

## Appendix A. Sample Table for Text included in the Narrative Synthesis.

Year	Abitur	Baccalauréat	SAT Exam
1788	<p>Edict vom 23. December 1788 über die Prüfung der auf die Universität gehenden Schüler.</p> <p><i>The edict was commissioned in the name and absolute authority of King Wilhelm III. of Prussia, without mentions of actual authors.</i></p>		
1812	<p>Einführung des Abiturs in Preußen: Edikt unterzeichnet von König Friedrich Wilhelm III., Staatskanzler Hardenberg und Friedrich von Schuckmann.</p> <p><i>This edict was published under the absolute authority of King W. III. of Prussia like above, but included the mention of a King's chancellor, Hardenberg and an advisor named Friedrich von Schuckmann.</i></p>		
1834	<p>Prüfungsordnung Deutscher Bund.</p> <p><i>Commissioned by not further defined bodies of the German confederation and published with authority in all constituent areas of the German confederation (Prussia, Austria, Rhine states)</i></p>		
1840		<ol style="list-style-type: none"> <li>1. Circulaire relative à l'exécution des arrêtés du 2 octobre 1840</li> <li>2. Instruction relative à l'exécution de l'arrêté du 14 juillet 1840</li> <li>3. Instruction relative à l'exécution de l'arrêté du 25 août 1840</li> </ol> <p><i>All three texts are descriptive documents to explain the execution of several related orders to restructure access to universities, and the Baccalauréat</i></p>	

Year	Abitur	Baccalauréat	SAT Exam
		<i>commissioned by the 2<sup>nd</sup> French empire's Ministry of Education. Beyond the royal authorities, authors are not mentioned.</i>	
1874		<p>Circulaire relative à l'application du nouveau plan d'études des lycées, prescrit par arrêté du 23 juillet 1874.</p> <p><i>A missive of the newly installed Ministry of Education under the 3<sup>rd</sup> French Republic. It explains and organizes major changes to plans of study in French lycées and to the Baccalauréat. No author is given beside the authorities of the ministry.</i></p>	
1902		<p>Programme des examens du Baccalauréat de l'enseignement secondaire. Décret et Arrêté du 31 Mai 1902.</p> <p><i>This document is an order and the explication of its execution by the French Ministry of Education, directly relating to new programs for the Baccalauréat. Also here, no author is given beside the authorities of the ministry.</i></p>	
1919			<p>College Board: Nineteenth Annual Report of the Secretary.</p> <p><i>(College Entrance Examination Board (1919). Nineteenth Annual Report of the Secretary. New York: College Board.)</i></p>
1925			<p>College Board: Twenty-fifth Annual Report of the Secretary.</p> <p><i>(College Entrance Examination Board (1925). Twenty-Fifth Annual Report of the Secretary. New York: College Board.)</i></p>

Year	Abitur	Baccalauréat	SAT Exam
1926	<p>Neue Reifeprüfungsordnung Weimar.</p> <p><i>This document is formalized curriculum on fundamental changes to the Abitur. Also here, no author is given beside the authorities of the Ministry of Education of the German Weimar republic.</i></p>		
1927		<p>Réorganisation du baccalauréat de renseignement secondaire. Rapport au Président de la République.</p> <p><i>This document is a letter from the French minister of education of the time, Mario Roustan (identified as author), to the French president. He summarizes several discussions on how the Baccalauréat needs to change. This document is important, as it seemingly serves as the most concise description of the late 1920s changes to the Baccalauréat.</i></p>	<p>College Board: Twenty Seventh Annual Report of the Secretary.</p> <p>(College Entrance Examination Board (1927). <i>Twenty-Seventh Annual Report of the Secretary</i>. New York: College Board.)</p>
1932			<p>College Board: Thirty-Second Annual Report of the Secretary.</p> <p>College Entrance Examination Board (1932). <i>Thirty-Second Annual Report of the Secretary</i>. New York: College Board.</p>
1933			<p>College Board: Thirty-Third Annual Report of the Secretary.</p> <p>(College Entrance Examination Board (1933). <i>Thirty-Third Annual Report of the Secretary</i>. New York: College Board.)</p>
1958	Tutzinger Gespräche I		
1959	Tutzinger Gespräche II		
1960	Tutzinger Gespräche III		

Year	Abitur	Baccalauréat	SAT Exam
	<p><i>These are notes on decisions made in the span of 3 years as a result of debates and discussions on meaning and purpose of the Abitur, between the German ministry of education and school directors from the East and Western occupational zones.</i></p> <p>Reference: Scheuerl, H. (1962). <i>Probleme der Hochschulreife: Bericht über die Verhandlungen zwischen Beauftragten der Ständigen Konferenz der Kultusminister und der Westdeutschen Rektorenkonferenz; 1958-1960, 'Tutzingen Gespräche'I-III. Heidelberg: Quelle &amp; Meyer.</i></p>		
1972 - 2018	<p><b>Vereinbarung zur Gestaltung der gymnasialen Oberstufe und der Abiturprüfung.</b></p> <p><i>A modern form of decree, similar to the historical documents, no author is identified, but the documents is commissioned under the full authority of the current German Ministry of Education. Steadily revised, latest version from 2018. (section "Kultusministerkonferenz")</i></p>		
1974			<p>Criterion-Referencing, Norm-Referencing, and the SAT.</p> <p>(Angoff, W. H. (1974). <i>Criterion-Referencing, Norm-Referencing, and the SAT</i>. College Board Review.)</p>
1977			<p>The Predictive Value of the SAT.</p> <p>Stanley, J. C. (1977). The predictive value of the SAT</p>

Year	Abitur	Baccalauréat	SAT Exam
			for brilliant seventh-and eighth-graders. <i>College Board Review</i> , 106, 31-37.
2006			<p><b>SAT Standards:</b></p> <ol style="list-style-type: none"> <li>1. English</li> <li>2. Math</li> <li>3. Sciences</li> </ol> <p>College Board (2006). <i>College Board Standards for College Success. English Language Arts</i>. New York: College Board.</p> <p>College Board (2006). <i>College Board Standards for College Success. Mathematics and Statistics</i>. New York: College Board.</p> <p>College Board (2006). <i>College Board Standards for College Success. Science</i>. New York: College Board.</p>
2007			<p>A Historical View of Subgroup Performance differences</p> <p>Kobrin, J. L., Sathy, V., &amp; Shaw, E. J. (2007). A <i>Historical View of Subgroup Performance Differences on the SAT Reasoning Test™</i>. Research Report No. 2006-5. College Board.</p>
2011			<p>„College Readiness Benchmark“</p> <p>Wyatt, J., Kobrin, J., Wiley, A., Camara, W. J., &amp; Proestler, N. (2011). <i>SAT Benchmarks: Development of a College Readiness Benchmark and Its Relationship to Secondary and Postsecondary School Performance</i>. Research</p>



Year	Abitur	Baccalauréat	SAT Exam
			<i>Report 2011-5. College Board.</i>
2012	<p><b>Bildungsstandards:</b></p> <ol style="list-style-type: none"> <li>1. Fremdsprache</li> <li>2. Deutsch</li> <li>3. Mathe</li> </ol> <p><i>A modern form of decree, similar to the historical documents, no author is identified, but the documents is commissioned under the full authority of the current German Ministry of Education. (section “Kultusministerkonferenz”)</i></p>		
2014		<p>Charte relative à l’élaboration, à la mise en œuvre et au suivi des programmes d’enseignement ainsi qu’aux modalités d’évaluation des élèves dans l’enseignement scolaire.</p> <p><i>Commissioned under the authority of the French Ministry of Education, this document discusses assessment and instruction in French schools and secondary to higher education. No author is identified beyond the authority the document is accorded by the legislative.</i></p>	
2015			<p>Test Specifications for the Redesigned SAT.</p> <p><i>College Board. (2015). Test specifications for the redesigned SAT. New York: College Board.</i></p>
2018		<p>Note d’analyses et de propositions sur les programmes du lycée et sur les épreuves du baccalauréat.</p> <p><i>Like above, this document is published under the authority of the French ministry of education and does not specify authors. It is an assembly of guidelines to</i></p>	

Year	Abitur	Baccalauréat	SAT Exam
		<i>structure French public secondary school education and the Baccalauréat.</i>	

Blue – From Chronicles   Black – Impact Factor Search   Red – From Governing Entity

## Appendix B. Coding Tables for Network Representation of Content.

History				
SAT US American History	Multiple Choice (No Choice in Response)	Abitur (Leistungskurs)	Free Response (Choice: 4 or 4)	Matched for comparison:
1798-1964. British colonial rule	1	19th Century: Kulturkampf German Propaganda	4	20
1789-1809. Constitutional period	7	History of Germany: Early and medieval to 1519, Migration Period	3	15
Slavery and the slave trade	3	20th Century: Postwar works on the Hitler period	3	15
Colonial History: Revolution to the Civil War, 1775/1783-1861	6	United States: The Revolution, 1775-1783: Declaration of Independence	3	15
Nineteenth amendment to Constitution, 1920 (Women suffrage)	6	History of Germany: New Empire, 1871-1918	5	25
19th Century: Franklin Delano Roosevelt's administrations, 1933-April 12, 1945	2	Abstracting and indexing histories	6	30
Truman's administrations, April 12, 1945-1953	1	Counterfactuals. Imaginary		
20th century: Cuba: Communist Regime	1	Data processing	8	40
20th century: political history	7	Periodization	3	15
Political History: Stamp Act Congress, New York, 1765	1	Critical Thinking	4	20
Late 19th century, 1865-1900	6	Preparation of thesis: Thesis writing	12	60
Early 19th century: Manifest Destiny	1			
20th Century: Addresses, essays, lectures, etc	3	<b>Baccalaureat (ES) (Histoire Geographie)</b>	<b>Free Response (Choice: 2 or 2)</b>	<b>Matched for comparison:</b>
20th century: sources and documents	5	Middle East, Southwestern Asia: History since 1945	9	27
20th century: World War II	2	History of France, Later 19th Century, Dreyfus Case	7	21
Colonial History: Indian Wars	1	Nation-state: Political Geography	8	24
20th century: Harding-Coolidge-Hoover era. "The twenties"	3	HA31 Theory and method of social science statistics: tables, graphics	10	30
America: Indian Tribes and Cultures	1	Abstracting and indexing histories	6	18
Early 19th century: War of 1812	3	Data processing	5	15
Late 19th century, 1865-1900: Foreign and general relations	2	Historiometry.	2	6
Middle 19th century, 1845/1848-1861	1	Methodology: Ethnohistory	12	36
African Americans: African American suffrage		LB 2369 Preparation of thesis: Thesis writing	18	54
African Americans: Civil Rights Act of 1964	2			
Later 20th century: Un-American Activities	1			
Colonial History: 1607-1689	3			
Colonial History: The Revolution, 1775-1783	2			
20th Century: Truman Administration: Fair Deal	1			
Later 20th century: Lyndon B. Johnson's administration	1			
Early 19th Century: Missouri Compromise	1			
Later 19th Century: Reconstruction, 1865-1877	2			
Later 20th Century: Clinton's administrations	1			
Military History: Second World War	1			
African Americans: Race Relations	1			
African Americans: Emancipation	2			
Later 20th century: Reagan's administrations, 1981-1989	2			
Military History: Vietnam War	2			
19th century: Political History	2			
Military History: Korean War	1			
Abstracting and indexing histories	60			
Data processing	14			
Imaginary histories	7			
Language	1			
Historiometry.	8			

Native Language Literacy				
SAT English Language and Literature	Multiple Choice (no choice in response)	Abitur (Leistungskurs)	Free Response (Choice: 2 or 2 or 2)	Matched for comparison:
Modern Drama: British Drama	9	Criticism and interpretation: Allegory and Symbolism	2	4
Modern Poetry: Lyrics, Songs	9	Criticism and Interpretation: Allusions	1	2
Victorian Era: Prose	9	Criticism and Interpretation: Characters	2	4
Prose: Novel, African American	8	Criticism and Interpretation: General Works	1	2
18th century: Autobiographical Writings	7	Criticism and Interpretation: Manners and Customs	3	5
Language, style: Use of Words	7	Criticism and Interpretation: Prose	2	4
Puritan Era: Poetry	7	Criticism and interpretation: Technique, plots, scenes, time, etc	1	2
17th century: English Poetry	6	Criticism and Interpretation: Poetic Works	2	4
Language, style: Early Modern English	5	Prosody: Metrics: Rhythmics	1	2
History of English Literature: Manners and Customs	3	Special parts of rhetoric: Choice of words, Vocabulary, etc.	1	2
17th century: Character Sketches	2	Style, Composition, Rhetoric		0
Special parts of rhetoric: Choice of words, Vocabulary, etc.	2	Poems: Versification	1	2
English Literature: Social Ideals	1	Discourse Analysis	3	5
Criticism and interpretation: Technique, plots, scenes, time, etc	9	18th century: Goethe and Schiller	9	16
Criticism and Interpretation: Poetic Works	4	20th century prose: Experimental Fiction	16	29
Style, Composition, Rhetoric	4	Language: Philosophy, Theory, Method	17	30
Criticism and interpretation: Allegory and Symbolism	3			
Criticism and Interpretation: Prose	3	Baccalaureat (L)	Free Response (no choice)	Matched for comparison:
Criticism and Interpretation: Characters	2	Criticism and Interpretation: Theater	3	11
Criticism and Interpretation: Manners and Customs	2	Criticism and Interpretation: Characters	2	8
Criticism and Interpretation: Social Problems	2	Criticism and Interpretation: General Works	2	8
Prosody: Metrics: Rhythmics	2	Criticism and Interpretation: Manners and Customs	1	4
Criticism and Interpretation: Allusions	1	Criticism and Interpretation: Drama	4	15
Criticism and Interpretation: General Works	1	Prosody: Metrics: Rhythmics	3	11
		Discourse Analysis	4	15
		PQ580 French Literary History: Romantic Drama	10	38
		PQ563 History of French Literature: Comedy	3	11
		Philology: Social Aspects	6	23

Mathematics				
SAT Mathematics Level 2	Multiple Choice (no choice in response)	Abitur Mathematics (Leistungskurs)	Procedural Response (Choice: 3 or 3)	Matched for comparison:
Algebra: Factors	2	Combinatorics: Graph Theory	6	11
Algebra: General Resolution of Equations	9	Stochastic Analysis: Axioms and Foundations	2	4
Algebra: Logarithms	1	Analysis: Integral Calculus	1	2
Algebra: Recursive Functions	1	Stochastic Analysis: Stochastic Inequalities	1	2
Algebra: Trigonometrics	10	Stochastic Analysis: Stochastic Differential Equations	1	2
Analysis: Calculus	7	Algebra: Trigonometrics	1	2
Arithmetic: Root	1	Geometry: Analytic Geometry	3	6
Combinatorics: Graph Theory	10	Geometry: Practical geometry. Geometrical drawing	1	2
Geometry: Analytic Geometry	5	Stochastic Analysis: Stochastic Sequences	1	2
Geometry: Elementary Geometry	4	Geometry: Geometry and trigonometry	2	4
Probabilities: Axioms and Foundations	1	Cybernetics Machine Learning	4	7
Probabilities: Distributions	1	Industries, Land Use, Labor: Leisure and Work	4	7
Probabilities: Stochastic Processes	1	Wealth: Consumption	2	4
Physics: Heat	1			
Physics: Motion of Projectiles	1	Baccalaureat (S) Mathematics	Procedural Response (no choice)	Matched for comparison:
Industries, Land Use, Labor: Leisure and Work	1	Combinatorics: Graph Theory	7	10
Demography: Population Geography	1	Stochastic Analysis: Axioms and Foundations	4	6
Economic Theory: Banking	1	Analysis: Integral Calculus	4	6
Wealth: Consumption	1	Stochastic Analysis: Stochastic Inequalities		0
		Stochastic Analysis: Stochastic Differential Equations		0
		Algebra: Trigonometrics	2	3
		Geometry: Analytic Geometry	8	11
		Geometry: Practical geometry. Geometrical drawing	1	1
		Stochastic Analysis: Stochastic Sequences		0
		QA295 Mathematics: Numerical Analysis	11	15
		SB317.6 Horticulture: Vocational	5	7
		GV1469.3 Video Games	12	17
		LC5225.A75 Adult Education - Assessment, Testing	11	15

## Appendix C. Coding Tables with Raw Numbers for Network Representation of Choice

	Answer Choice		Conductivity		Specialization?	
	Answer Choice Abitur	Answer Choice BAC	Overall Abitur (sum of result of all subject tracks)	Overall Baccalauréat	Specialization Abitur	Specialization BAC
Math	3	1	330	26	2	3
Literacy Native Language	3	1	272	49	2	2
History		2	249	22	2	3
French			350		3	3
English	2	1	350	24	3	3
Biology	2	1	390	25	3	3
Spanish	2	1	350	24	3	3
Physics	2	1	390	25	3	3
Latin	2	1	350	25	3	3
Chemistry	2	1	390	25	3	3
Economy	2	2	146	26		3
Philosophy	2	3	271	49	3	2
Geography	2	1	52	26	3	3
Creative Arts	2	2	271	22	3	3
Modern Greek			350	24	3	3
Italian		1	350	24	3	3
Ancient Greek		1	350	25	3	3
Music	2		271	22	3	3
Theater		2	271	22	3	3
Dutch			350		3	3
Physical Education	2		248	26		3
Jewish Religion			261		3	3

	Answer Choice		Conductivity		Specialization?	
	Answer Choice Abitur	Answer Choice BAC	Overall Abitur (sum of result of all subject tracks)	Overall Baccalauréat	Specialization Abitur	Specialization BAC
Christian Religion	2		261		3	3
Ethics	2		271		3	3
IT			210		3	3
Russian			350		3	3
Hebrew			350		3	3
Japanese			350		3	3
Portuguese		1	350	24	3	3
Islamic Religion			261		3	3
Turkish			350		3	3
Chinese		1	350	24	3	3
Korean			350		3	3
Political Science	2		297	26		3
German		1		24		3
Cinema		2		22		3
Agriculture		1		25		3
Art History		1		22		3

## Appendix D. IRB Document



**TO:** Skinner, Kim  
LSUAM | Col of HSE | Education

**FROM:** Alex Cohen  
Chair, Institutional Review Board

**DATE:** 17-Sep-2020

**RE:** IRBAM-20-0070

**TITLE:** Pruefung - Interdisciplinary research of  
assessment across three languages and multiple  
frameworks of analysis.

**SUBMISSION TYPE:** Initial Application

**Review Type:** Exempt

**Risk Factor:** Minimal

**Review Date:** 17-Sep-2020

**Status:** Approved

**Approval Date:** 17-Sep-2020

**Approval Expiration Date:** 16-Sep-2023

**Re-review frequency:** Three Years

**Number of subjects approved:** 0

**LSU Proposal Number:**

**By:** Alex Cohen, Chairman

**Continuing approval is CONDITIONAL on:**

1. Adherence to the approved protocol, familiarity with, and adherence to the ethical standards of the Belmont Report, and LSU's Assurance of Compliance with DHHS regulations for the protection of human subjects\*
2. Prior approval of a change in protocol, including revision of the consent documents or an increase in the number of subjects over that approved.
3. Obtaining renewed approval (or submittal of a termination report), prior to the approval expiration date, upon request by the IRB office (irrespective of when the project actually begins); notification of project termination.
4. Retention of documentation of informed consent and study records for at least 3 years after the study ends.
5. Continuing attention to the physical and psychological well-being and informed consent of the individual participants, including notification of new information that might affect consent.
6. A prompt report to the IRB of any adverse event affecting a participant potentially arising from the study.



7. Notification of the IRB of a serious compliance failure.
8. **SPECIAL NOTE: When emailing more than one recipient, make sure you use bcc. Approvals will automatically be closed by the IRB on the expiration date unless the PI requests a continuation.**

*\* All investigators and support staff have access to copies of the Belmont Report, LSU's Assurance with DHHS, DHHS (45 CFR 46) and FDA regulations governing use of human subjects, and other relevant documents in print in this office or on our World Wide Web site at <http://www.lsu.edu/research>*

Louisiana State University  
131 David Boyd Hall  
Baton Rouge, LA 70803

O 225-578-5833  
F 225-578-5983  
<http://www.lsu.edu/research>

## References

- Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometrics. *Educational assessment*, 8(3), 231-257.
- Adorno, T.W. (1966). *Negative Dialektik*. Frankfurt am Main: Suhrkamp Verlag.
- Ahearn, L. M. (2011). *Living language: An introduction to linguistic anthropology*. John Wiley & Sons.
- Arneton, M., Bocéran, C., & Flieller, A. (2013). Les performances en mathématiques des élèves des départements d'Outre-Mer. *L'orientation scolaire et professionnelle*, 42(1), 1-29.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258–267.
- Baker, D., & LeTendre, G. K. (2005). *National differences, global similarities: world culture and the future of schooling*. Stanford, CA: Stanford Social Sciences.
- Batley, S. (2014). *Classification in Theory and Practice (Chandos Information Professional Series)* (2nd ed.). Chandos Publishing.
- Baumeister, R. F. (2003). Writing a literature review. In M. Prinstein & M. Patterson (Eds.), *The portable mentor: Expert guide to a successful career in psychology* (pp. 57-71). New York: Springer.
- Barber, B. R. (2011). *Jihad vs. Mcworld*. London: Corgi.
- Bartell, T. G. (2013). Learning to teach mathematics for social justice: Negotiating social justice and mathematical goals. *Journal for Research in Mathematics Education*, 44(1), 129-163.
- Behr, D., & Shishido, K. (2016). 19 The Translation of Measurement Instruments for CrossCultural Surveys. *The SAGE handbook of survey methodology*.
- Bennett, J. S., Driver, M.K. & Trent, S.C. (2017). Real or ideal? A narrative literature review addressing white privilege in teacher education. *Urban Education*, 54(7), 891–918.
- Bhabha, H. K. (1997). *The location of culture*. London: Routledge.
- Bhaskar, R., Murthy, R. S. R., Miglani, B. D., & Viswanathan, K. (1986). Novel method to evaluate diffusion controlled release of drug from resinate. *International journal of pharmaceutics*, 28(1), 59-66.
- Blaikie, N., & Priest, J. (2017). *Social research paradigms in action*. Cambridge, UK: Polity.

- Blum-Barth, Natalia (2016). Transkulturalität, Hybridität, Mehrsprachigkeit. Von der Vision zur Revision einiger Forschungstrends. *GFL German as foreign language*, 2016(1), 114-127.
- Bölling, R. (2010). *Kleine Geschichte des Abiturs*. München: Schöningh.
- Bourdieu, P. (1977). *Outline of a theory of practice*. New York: Cambridge University Press.
- Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory of research for the sociology of education* (pp. 241–258). Westport: Greenwood Publishing Group.
- Brezina, V. & Meyerhoff, M. (2014). A critical review of sociolinguistic generalizations based on large corpora. *International Journal of Corpus Linguistics*, 19(1), 1-28.
- Brown, A. I., & Ferrera, R. A. (1985). Diagnosing zones of proximal development. In J. Wertsch (Ed.), *Culture communication and cognition: Vygotskian perspective* (pp. 273-305). Cambridge, UK: Cambridge University Press.
- Buchmann, C., Condrón, D.J., Roscigno, V.J. (2010). Shadow Education, American Style: Test Preparation, the SAT and College Enrollment. *Social Forces*, 89 (2), p. 435–461.
- Bundeszentrale Politische Bildung (Ed.). (2016, December 09). *Das Tor zur Universität - Abitur im Wandel - Geschichte des Bildungssystems*. - Retrieved November 06, 2020, from <https://www.bpb.de/gesellschaft/bildung/zukunft-bildung/238795/abitur-im-wandel>
- Bußmann, H. (2002). *Lexikon der Sprachwissenschaft*. Michigan: Kröner.
- Carley, K. M. (1997). Network text analysis: The network position of concepts. *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts*, 4, 79-100.
- Carver, R. P. (1974). Two dimensions of tests: Psychometric and edumetric. *American Psychologist*, 29(7), 512.
- Charmaz, K. (2000). Grounded Theory. Objectivist and Constructivist Methods. In Y. S. Lincoln & N. K. Denzin (Eds.), *Handbook of Qualitative Research* (pp. 509–553). Thousand Oaks, Canada: SAGE Publications.
- Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on information theory*, 2(3), 113-124.
- Chomsky, N. (1995). *The minimalist program*. Michigan: MIT Press.
- Chouliaraki, L., & Fairclough, N. (1999). *Discourse in late modernity: Rethinking critical*

- discourse analysis*. Edinburgh University Press.
- Clark, M., Rothstein, J., & Schanzenbach, D. W. (2009). Selection bias in college admissions test scores. *Economics of Education Review*, 28(3), 295–307.
- Clifford, J. (1989). *The predicament of culture*. Harvard: Harvard University Press.
- College Board (2019 a). PSAT-related assessments – 2019 SAT suite of assessments program results – The college board. *College Board Program Results*. Retrieved September 2019 from: <https://reports.collegeboard.org>.
- College Board. (2019 b). SAT Program participation and performance statistics – Research. *College Board*. - Retrieved May 6, 2020, from <https://research.collegeboard.org/programs/sat/data>
- College Board. (2021a). *College Readiness*. - Retrieved April 6, 2021, from <https://research.collegeboard.org/programs/sat/data>
- College Board (2021b). SAT Subject Test Areas. – Retrieved April 10, 2021, from <https://collegereadiness.collegeboard.org/sat-subject-tests/subjects>
- Critchley, S., Derrida, J., Laclau, E., & Rorty, R. (1996). *Deconstruction and pragmatism*. Psychology Press.
- Delcroix, A., Forissier, T., & Anciaux, F. (2013). Vers un cadre d'analyse opérationnel des phénomènes de contextualisation didactique. In: Anciaux, F., Forissier, T., & Prudent, L.F. (Eds.), *Contextualisations didactiques : approches théoriques*, (pp.141-185). Paris: L'Harmattan.
- Delgado, R., & Stefancic, J. (2012). *Critical race theory: An introduction*. New York: New York University Press.
- Denzin, N. K., & Lincoln, Y. S. (2002). *Handbook of qualitative research*. Thousand Oaks, CA: Sage.
- Derrida, J. (1972). *MARGES DE PHILOSOPHIE (CRITIQUE)*. Paris : Minuit.
- Derrida, J., Allison, D. B., & Garver, N. (1973). *Speech and Phenomena: And Other Essays on Husserl's Theory of Signs (Studies in Phenomenology and Existential Philosophy)*. Northwestern University Press.
- Derrida, J. (1990). *L'archéologie du frivole*. Galilée.
- Derrida, J. (1993). *Spectres de Marx: L'état de la dette, le travail du deuil et la nouvelle Internationale (Collection La philosophie en effet) (French Edition)*. Editions Galilée.
- Derrida, J. (2020). *Deconstruction in a Nutshell: A Conversation with Jacques Derrida, With a*

- New Introduction*. Fordham University Press.
- Deutscher Akademischer Austauschdienst - DAAD. (n.d.). Retrieved November 2019 from: <https://www.daad.de/de/>.
- Diesner, J., & Carley, K. M. (2005). Revealing Social Structure from Texts. *Causal Mapping for Research in Information Technology*, 81–108.
- Dos Santos, S., Rakocevic, R., Lefresne, F., & Trosseille, R. (2012). *Les cadres d'évaluation en vue d'améliorer les résultats scolaires. Étude de l'OCDE - Rapport national de base pour la France*. Paris: Depp.
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 191, 61-74.
- Eberhard, V. (2006). *Das Konzept der Ausbildungsreife: ein ungeklärtes Konstrukt im Spannungsfeld unterschiedlicher Interessen*. Berlin: Bundesinstitut für Berufsbildung.
- Edwards, D., Coates, H., & Friedman, T. (2012). A survey of international practice in university admissions testing. *Higher Education Management and Policy*, 24(1), 1-18.
- Ehrhart, Sabine (2014). Le développement plurilingue et interculturel en milieu éducatif ouvert à la diversité – étude et bilan de trois projets universitaires avec la participation d'une écologie linguistique "à la luxembourgeoise". In : Abendroth-Timmer, D., & Hennig, E.-M. (Eds.) *Plurilingualism and multiliteracies*, (pp. 90-102). Luxembourg: Peter Lang AG.
- Engels, F. (2019). *Dialektik der Natur*. Verlag Neuer Weg.
- Erel, U. (2010). Migrating cultural capital: Bourdieu in migration studies. *Sociology*, 44(4), 642-660.
- Fairclough, N. (2000) Discourse, social theory and social research: the discourse of welfare reform: *Journal of Sociolinguistics* (4), 2. 163–195.
- Fairclough, N. (2001). Critical Discourse Analysis as a Method in Social Scientific Research. In: M. Meyer & R. Wodak (Eds.). *Methods of Critical Discourse Analysis* (pp. 121–138). Thousand Oaks, Canada: SAGE Publications.
- Fairclough, N. (2013). *Critical Discourse Analysis: The Critical Study of Language* (2nd ed.). New York: Routledge.
- Freeman, D. (1996). Derek Freeman: Reflections of a heretic. *The Evolutionist*, 1-5.
- Furr, R. M. (2017). *Psychometrics: an introduction*. Sage Publications.
- Getahun, A. (2018). Education in France. *World Education Service*. - Retrieved May 5, 2020,

- from <https://wenr.wes.org/2015/09/education-france>
- Geertz, C. (2008). Thick description: Toward an interpretive theory of culture. In *The cultural geography reader* (pp. 41-51). Routledge.
- Gök, Fatma (2010). Marketing hope: Private institutions preparing students for the university entrance examination in turkey. *International perspectives on education and society international educational governance*, 12, 123–134.
- Goetze, C. (2017). *The distinction of peace: a social analysis of peacebuilding*. Ann Arbor: University of Michigan Press.
- Gordon, E. W., & Terrell, M. D. (1981). The changed social context of testing. *American Psychologist*, 36(10), 1167-1189.
- Graham, P. (2000) Hypercapitalism: A political economy of informational idealism. *New Media and Society*, 2(2), 131–156.
- Harkness, J., Van de Vijver, F.J.R., & Mohler, P.Ph. (2003). *Cross-Cultural Survey Methods*. New Jersey: Wiley & Sons.
- Harris, C. (2013). library classification 2020. *Knowledge Quest*, 42(2), 14.
- Heckman, J. J. (2003). China's investment in human capital. *Economic Development and Cultural Change*, 51(4), 795–804.
- Hennig, K., & Kirova, A. (2012). The role of cultural artefacts in play as tools to mediate learning in an intercultural preschool programme. *Contemporary Issues in Early Childhood*, 13(3), 226–241.
- Hong, W., & Youngs, P. (2008). Does high-stakes testing increase cultural capital among low-income and racial minority students? *Education policy analysis archives*, 16(6), 2-18.
- House, R.J.; Hanges, P.J., Javidian, M., Dorfman, P.W., & Gupta, V. (2010). *Culture, leadership, and organizations. The GLOBE study of 62 societies*. Thousand Oaks: Sage.
- Huber, M. (1981). The renewal of curriculum theory in the 1970s: An historical study. *JCT*, 3(1), 14-84.
- Institute of International Education (2017). New international student enrollment, 2006/2007 – 2016/2017. *Open Doors Report on International Educational Exchange*. Retrieved November 2019 from <https://www.iie.org/Research-and-Insights/Open-Doors>.
- ITB Consulting (2019). *Maßnahmen der Qualitätssicherung beim TestAS – Ergebnisse wissenschaftlicher Studien*. Bonn: ITB.

- Jeantheau, J. P., Johnson, S., Garbe, C., Geron, S., Lafontaine, D., Mallows, D., Shiel, G. & Valtin, R. (2016). *Literacy in France. Country report. Short version*. Cologne: Elinet.
- Jones, K. (2004). Mission drift in qualitative research, or moving toward a systematic review of qualitative studies, moving back to a more systematic narrative review. *Qualitative Report*, 9, 95-112.
- Kamenetz, A. (2016). *The test: Why our schools are obsessed with standardized testing--but you dont have to be*. Philadelphia: PublicAffairs.
- Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis?. *American anthropologist*, 86(1), 65-79.
- Kilgariff, A., Baisa, V., Michelfeit, J., & Suchomel, V. (n.d.). *Language corpus management and query system*. Retrieved April 2019 from <http://www.sketchengine.eu/>
- Kuhn, K., & Fekete, P. (2019). *Deutsch-chinesische Studienangebote erfolgreich managen. Rahmenbedingungen und Erfolgsfaktoren aus interkultureller Perspektive*. Wiesbaden: Springer Gabler.
- Kultusministerkonferenz (2019). - Retrieved November 2019 from: <https://www.kmk.org/zab/central-office-for-foreign-education.html>.
- Kultusministerkonferenz. (2020). Abiturnoten. - Retrieved May 6, 2020, from <https://www.kmk.org/dokumentation-statistik/statistik/schulstatistik/abiturnoten.html>
- Kultusministerkonferenz (2020b). - Retrieved November 2020 from: <https://www.kmk.org/dokumentation-statistik/beschluesse-und-veroeffentlichungen.html>.
- Kultusministerkonferenz. (2020c). *Sekundarstufe II / Gymnasiale Oberstufe und Abitur*. – Retrieved April 2021 from: <https://www.kmk.org/themen/allgemeinbildende-schulen/bildungswege-und-abschluesse/sekundarstufe-ii-gymnasiale-oberstufe-und-abitur.html>
- Kumbhar, R. (2011). *Library classification trends in the 21st century*. Elsevier.
- Kramer, J., Neumann, M., & Trautwein, U. (2016). *Abitur und Matura im Wandel: Historische Entwicklungslinien, aktuelle Reformen und ihre Effekte (Edition ZfE) (German Edition)* (1. Aufl. 2016 ed.). Frankfurt, Germany: Springer VS.
- Kruse, A. (2017). *Lebensphase hohes Alter: Verletzlichkeit und Reife*. Springer-Verlag.
- Kruse, A. J. (2016). Cultural bias in testing: A review of literature and implications for music education. *Update: Applications of Research in Music Education*, 35(1), 23-31.

- Leamy, M., Bird, V., Le Boutillier, C., Williams, J., & Slade, M. (2011). Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. *The British Journal of Psychiatry*, 199(6), 445-452.
- Lemke, J. C. (2011). Factors that may impact the number of international students from China, Japan, South Korea, Thailand or Vietnam who study in the United States. Doctoral dissertation. - Retrieved November 2019 from <http://gradworks.umi.com/3491553.pdf>.
- Lentz, C. (2011). Der Kampf um die Kultur. Zur Ent- und Re-Soziologisierung eines ethnologischen Konzepts. *Soziale Welt*, 60, 305-324.
- Leonard, J., Brooks, W., Barnes-Johnson, J., & Berry III, R. Q. (2010). The nuances and complexities of teaching mathematics for cultural relevance and social justice. *Journal of Teacher Education*, 61(3), 261-270.
- Library of Congress. (2021). *Library of Congress Classification PDF Files*. Library of Congress Classification. Retrieved February 10, 2021, from <https://www.loc.gov/aba/publications/FreeLCC/freelcc.html>
- Marchand, P. (2009). Un chantier à ouvrir : L’histoire du baccalauréat, 1808-1940. *Histoire De L’éducation*, (94), 201-208. doi:10.4000/histoire-education.819
- Manning, W., Jackson, R. (1984). College Entrance Examinations - Objective Selection or Gatekeeping for the Economically Privileged. In: Reynolds, C. R., & Brown, R. T. *Perspectives on bias in mental testing*. (pp. 189 – 220). New York: Plenum Press.
- Marchand, P. (2009). Un chantier à ouvrir : L’histoire du baccalauréat, 1808-1940. *Histoire De L’éducation*, (94), 201-208. doi:10.4000/histoire-education.819
- Marx, K., Heyden G., Egorov, A., & Engels, F. (1975). *Karl Marx, Friedrich Engels Gesamtausgabe (Mega)*. Berlin, Ost: Dietz Verlag.
- McKiernan, E. C; Schimanski, L. A; Muñoz Nieves, C.; Matthias, L.; Niles, Meredith T.; Alperin, J. P (2019). Use of the Journal Impact Factor in academic review, promotion, and tenure evaluations. *eLife*, (8).
- Melucci, A. (1996) *Challenging Codes: Collective Action in the Information Age*. Cambridge: CUP.
- Meuriot, P. (1919). Le baccalauréat. Son évolution historique et statistique des origines (1808) à nos jours. *Journal De La Société Statistique De Paris*, Tome 60(1), 8-35.
- Ministère de l’Education Nationale et de la Jeunesse. (2020a). Le baccalauréat général. – Retrieved May 3, 2020, from <https://www.education.gouv.fr/le-baccalaureat-general-10457>
- Ministère, de l’Education Nationale et de la Jeunesse (2020b). Recherche. Retrieved November 09, 2020, from <https://www.education.gouv.fr/recherche>



- Misra, G., Sahoo, F. M., & Puhan, B. N. (1997). Cultural bias in testing: India. *European Review of Applied Psychology/Revue Européenne de Psychologie Appliquée*, 47(4), 309-317.
- Moje, E. B., Ciechanowski, K. M., Kramer, K., Ellis, L., Carrillo, R., & Collazo, T. (2004). Working toward third space in content area literacy: An examination of everyday funds of knowledge and Discourse. *Reading Research Quarterly*, 39(1), 38–70.
- Müller, A. & Schmidt, B. (2009). Prüfungen als Lernchance: Sinn, Ziele und Formen von Hochschulprüfungen. *Zeitschrift für Hochschulentwicklung*, 4(1), 23-45.
- NCES (2019). The Condition of Education - Postsecondary Education - Postsecondary Students - Immediate College Enrollment Rate - Indicator April. - Retrieved May 6, 2020, from [https://nces.ed.gov/programs/coe/indicator\\_cpa.asp](https://nces.ed.gov/programs/coe/indicator_cpa.asp).
- Niedersächsisches Landesinstitut für schulische Qualitätsentwicklung (2020, January 20). 2020 - *Niedersächsischer Bildungsserver*. - [https://www.nibis.de/2020\\_14733](https://www.nibis.de/2020_14733).
- Partnership, G. S. (2014). High-stakes test definition. Retrieved November 2019 from: <https://www.edglossary.org/high-stakes-testing/>.
- Petticrew, M. & Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. Hoboken, NJ: John Wiley.
- Pinar, W. F. (1989). A reconceptualization of teacher education. *Journal of Teacher Education*, 40(1), 9-12.
- Pinar, W. (2004). *What is curriculum theory?* Mahwah, NJ: Erlbaum.
- Pinar, W. F. (2006). *Understanding curriculum an introduction to the study of historical and contemporary curriculum discourses*. New York: Lang.
- Pöggeler, F. (1964). Der Mensch in Mündigkeit und Reife. *Eine Anthropologie des Erwachsenen*.
- van der Pol, H., & UNESCO Institute for Statistics. (2012). *International Standard Classification of Education - ISCED 2011*. - Retrieved 05/03/2021 from <http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf>
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., et al. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. *A product from the ESRC methods programme Version, 1*. – Retrieved 11/04/2020 from [https://www.researchgate.net/profile/Mark\\_Rodgers4/publication/233866356\\_Guidance\\_on\\_the\\_conduct\\_of\\_narrative\\_synthesis\\_in\\_systematic\\_reviews\\_A\\_product\\_from\\_the\\_ESRC\\_Methods\\_Programme/links/02e7e5231e8f3a6183000000/Guidance-on-the-conduct-](https://www.researchgate.net/profile/Mark_Rodgers4/publication/233866356_Guidance_on_the_conduct_of_narrative_synthesis_in_systematic_reviews_A_product_from_the_ESRC_Methods_Programme/links/02e7e5231e8f3a6183000000/Guidance-on-the-conduct-)

of-narrative-synthesis-in-systematic-reviews-A-product-from-the-ESRC-Methods-Programme.pdf

Popping, R. (2000). *New Technologies for Social Research: Computer-assisted text analysis*. London, : SAGE Publications, Ltd

Popping, R. (2003). Knowledge Graphs and Network Text Analysis. *Social Science Information*, 42(1), 91–106.

Rackin, H. (2020). A Comparison of Marriage and Fertility Schemas using Network Text Analysis. *Not published*. – Retrieved 05.03.2020 from: [https://www.lsu.edu/hss/sociology/people/faculty/faculty\\_profiles\\_new\\_2018/cvs/rackin\\_cv9-8-18.pdf](https://www.lsu.edu/hss/sociology/people/faculty/faculty_profiles_new_2018/cvs/rackin_cv9-8-18.pdf)

Reynolds, C. R., & Brown, R. T. (1984). *Perspectives on bias in mental testing*. New York: Plenum Press.

Ryan, G. & Bernard, R.H. (2002). Data management and analysis methods. In Denzin, N. K., & Lincoln, Y. S. (Eds.). *Handbook of qualitative research* (pp. 769-802). Thousand Oaks, CA: Sage.

Said, E. W. (2003). *Orientalism: Edward W. Said*. New York: Vintage Books.

Salaky, K. (2018). What standardized tests look like in 10 places around the world. Retrieved November 2019 from: <https://www.insider.com/standardized-tests-around-the-world-2018-9>.

Sandelowski, M., Docherty, S., & Emden, C. (1997). Qualitative metasynthesis: Issues and techniques. *Research in nursing & health*, 20(4), 365-371.

SankeyMATIC (2021). – Retrieved 05.01.2021 from: <http://sankeymatic.com/build/>

Scribbr. (2020, December 8). *Scribbr - Your path to academic success*. <https://www.scribbr.com/>

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2020). *Kultusministerkonferenz der Länder und der Bundesrepublik Deutschland*. Retrieved May 5, 2020, from <https://www.kmk.org>

Selbmann, R. (1994). *Der deutsche Bildungsroman*. Weimar: Metzler.

Schrader, M. (2019). *Mimesis und Poiesis: Poetologische Studien zum Bildungsroman* (Vol. 65). Walter de Gruyter GmbH & Co KG.

Selbmann, R. (2016). *Der deutsche Bildungsroman*. Springer-Verlag.

Shohamy, E. (2016). *Power of tests: A critical perspective on the uses of language tests*.

- London: Routledge.
- Shohamy, E., Or, I. G., & May, S. (2017). *Language Testing and Assessment (Encyclopedia of Language and Education)*. New York, USA: Springer.
- Stokman, F.N., de Vries, P.H. (1988) Structuring Knowledge in a Graph. In: van der Veer G.C., Mulder G. (eds) *Human-Computer Interaction* (186-210). Springer, Berlin, Heidelberg
- Takeuchi, Y. (1997). The self-activating entrance examination system – its hidden agenda and its correspondence with the Japanese “salary man”. *Higher Education*, 34, 183–198.
- Thompson, G. L., & Allen, T. G. (2012). Four effects of the high-stakes testing movement on African American K-12 students. *The Journal of Negro Education*, 81(3), 218-227.
- Tinnefeld, T. (2013). *Dimensionen der Prüfungsdidaktik, Analysen und Reflexionen zur Leistungsbewertung in den modernen Fremdsprachen*. Aachen: Shaker.
- Tinnefeld, T. (2002). *Prüfungsdidaktik: Zur Fundierung einer neuen wissenschaftlichen Disziplin - am Beispiel der modernen Fremdsprachen*. Aachen: Shaker.
- Valentine, John (1987). *The College Board and the School Curriculum. A History of the College Board's Influence on the Substance and Standards of American Education, 1900-1980*. New York: College Entrance Examination Board.
- Van der Veer, R., & Valsiner, J. (1994). *The Vygotsky reader*. Oxford: Blackwell.
- Vereinbarung zur Gestaltung der gymnasialen Oberstufe und der Abiturprüfung<sup>1</sup>, 1 Sekretariat der ständigen Konferenz (Kultusministerkonferenz 2016 (orig. 1972)).
- Varghese, N. V. (2008). *Globalization of higher education and cross-border student mobility*. Paris: IIEP/UNESCO.
- Veritas Health Innovation (2020). *Covidence systematic review software*. Melbourne, Australia.
- Vertovec, S. (2007). Super-diversity and its implications. *Ethnic and racial studies*, 30(6), 1024-1054.
- Vertovec, S. (2015). Formulating diversity studies In Vertovec, S. (Ed.), *Routledge international handbook of diversity studies* (pp. 1-21). New York: Routledge.
- Vygotsky, L. S., Cole, M., Stein, S., & Sekula, A. (1978). *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Waller, L. G. (2006). Introducing Fairclough's critical discourse analysis methodology for analyzing Caribbean social problems: going beyond systems, resources, social action, social practices and forces of structure or lack thereof as units of analysis. *Journal of Diplomatic Language*, (3)1. 1-20.

- WebPedago. (n.d.). *Partagez la connaissance*. - Retrieved April 13, 2021, from <https://www.webpedago.com>
- Wells, G. (1999). The zone of proximal development and its implications for learning and teaching. In Wells, G. (Ed.), *Dialogic inquiry: Towards a sociocultural practice and theory of education* (pp. 313-334). Cambridge UK: Cambridge University Press.
- Wertsch, J. V. (1998). *Mind as Action*. Oxford: Oxford University Press.
- Wheeler, J. B. (2013). *A campaign of quiet persuasion. How the College Board desegregated SAT test centers in the deep south, 1960-1965*. Baton Rouge: LSU Press.
- Wygotski, L. S. (1980): *Ausgewählte Schriften, Bd. 1: Arbeiten zu theoretischen und methodischen Problemen der Psychologie*. Berlin: Volk und Wissen.

## **Vita**

Thomas Benz, made Thomas Erich Benz by the power accorded in his F-1 Visa, arrived in the US in 2018 to begin a PhD in Curriculum and Instruction with the School of Education. Before his PhD he completed a Master of Arts in Frankfurt, Germany and a Bachelor of Arts in Mainz, Germany in Cultural Studies and Linguistics. During and between these degrees, he worked and studied in Germany, France, and Ivory Coast in diverse occupations from market research, translation to being a teacher and tutor of languages at private schools and academies. He worked for NGOs like AIESEC and companies like Fraport, Frankfurt Airport.