



GÖTEBORGS UNIVERSITET

ILSA-the subordination of Education to
Economic Theory
(ILSA-Wirtschaftstheorie statt Bildung)

Anders Lindqvist

Programme/course: International Masters of Educational Research/PDA184

Master Thesis: 30 ECTS **Level:** Advanced level

Semester/Year: Winter/2016

Supervisor: Prof. Dr. Girma Berhanu

Acknowledgements

The study was made through the helpfulness of all those that participated and who most generously shared their time and knowledge with me. Especially I thank Prof. Dr. Richard Münch for the inspiration to an idea, Dr. Sigrid Hartong for the development of an idea and Prof. Dr. Ewald Terhart for very valuable contacts.

Furthermore I am in debt to Dr. Joachim Wuttke without whose help this study had not been possible and last, my very patient and ever encouraging supervisor Prof. Dr. Girma Berhanu, I owe you.

Abstract

Narrative: In the beginning of the 21st century a trend became apparent; the increasing production of education monitoring data. It was the outcome of international large scale assessments (ILSA) breakthrough; quantitative comparative education (QCE) studies said to be able to give important information on the state of education systems. Germany, which until then had not taken part, became an ardent advocate of the testing, the reason was the poor results in TIMSS 1995 backed by PISA 2000 three years later. However, by the participation in ILSA concepts with origin in economic theory, concepts such as quality, standard and assessment, are brought into the education context in the partaking land. Although these terms have a history in education, they change meaning and content, they become concepts for Human Capital theory. The reason is that by focusing on comparisons based on scales (test results) students become comparable numbers, *valueated*¹ by competitiveness, or in other words, by their thought Human Capital. This study tries to empirically demonstrate the consequences of this conceptual transfer by a case study in form of a German Land, Bavaria. It deals with how these concepts become articulated and reconstructed, the effect on the different levels in an education system and on the working practice. Furthermore, the study gives a frame narrative on how QCE became a dominant research field in education.

Methodology: The research methodology was foremost qualitative, though supplemented with quantitative data in the form of descriptive statistics. The data was collected by interviews and literature studies. Pierre Bourdieu's field working theory was the main tool in the analytical work because of the possibilities and flexibility it offered, to explain and display collective and relational positional changes associated with altered societal norms.

Main findings: Bavaria was the main proponent of the introduction of ILSA in Germany. The main reasons were a combination of an already economized education system and close links between the political and economic fields. The introduction meant that the notion of education as the prime savior of future success was established and by that, a rush and demand of best secondary schooling was created. This demand was handled by the dominant within Bavarian politic, CSU, through the establishment of transitional paths between the different forms of secondary schooling. However, instead of raising the number of transitions to Gymnasium, the majority of transitions have been departures from Gymnasium. This is a consequence of the increased number of students that mean a more heterogeneous mass of students in Gymnasium with the outcome of an increasingly complex work situation for teachers. The exemptions are Humanistic Gymnasiums, the traditionally most advanced form of secondary school, with much symbolic capital that has maintained their composition of students. The partaking in ILSA has raised the number of tests, assessments and evaluations, and also brought in a different test culture (format, interval, cohorts etc.). Quality, that instantly became the sought dimension in education, has proven very hard to define in practice even for the state institution, ISB/QA, specially created for the task. Finally, the balance in the relationship between the Bavarian state party *Christlich Soziale Union (CSU)* and *Empirische Bildungsforschung (EB)*, the German equivalent of QCE has changed, that is, in the establishment of EB in Germany the dependence on local, Land or state, backing has decreased which marks a shift of relation and positions.

Key words: ILSA, Comparative Education, EB, Human Capital, Concepts, Quality, Standard, Assessment

¹ Valueated is an expression that is used in describing business and stock market transactions.

Comparative education is the Anglo-Saxon common term for the research field in focus of this study. However, since it can be understood as both quantitative and qualitative, and because of the national orientation of this study, Germany, where the term *Empirische Bildungsforschung* (EB) is used with the focal point of quantity and *Bildung* which is a wider concept than just education, I have applied the term quantitative comparative education (QCE) for international research and EB for the specific German similar research, throughout most of the study. The exceptions are limited to parts where the literature and cited authors use the term comparative education, which is the regular – but imprecise – description generally used.

Abbreviations and specific terms

<i>CERI</i>	The Center for Educational Research and Innovation at OECD
<i>EDC</i>	The education and policy committee at OECD, the former education committee
<i>ILSA</i>	International Large Scale Assessments
<i>INES</i>	Indicators of Education Systems. An OECD programme
<i>IEA</i>	International Associations for the Evaluation of Educational Achievement
<i>IBE</i>	The International Bureau of Education, created in 1925 as a private institute and became the first intergovernmental organization in the field of education 1929. Since 1969 an institute of UNESCO
<i>IIEP</i>	The International Institute for Educational Planning at UNESCO
<i>IO</i>	International Organization
<i>NPM</i>	New Public Management
<i>OECD</i>	Organization for Economic Co-operation and Development
<i>QCE</i>	Quantitative Comparative Education
<i>UNESCO</i>	The United Nations Educational, Scientific and Cultural Organization

(German -)

<i>BMBF</i>	The Federal Ministry of Education and Research
<i>Bund</i>	The Federal State of Germany
<i>DFG</i>	The Central Organization for the promotion of research at universities and other publicly financed research institutions in Germany
<i>DIPF</i>	The German Institute for Pedagogic Research
<i>EB</i>	German empirical educational research with emphasis on quantitative comparative research
<i>FDZ</i>	The research data centre at IQB
<i>IPN</i>	Leibnitz Institute for Science and Mathematics Education. Promote research in these areas
<i>ISB</i>	The Bavarian state institute that supports and advises the Bavarian Ministry for education in the development of differentiated education.
<i>IQB</i>	The Institute for the Development of Quality in Education
<i>KMK</i>	The standing conference of the ministers from the German Länder that deals with educational, scientific and cultural matters. It is not a part of the federal government and its directives has to be implemented into the Länder's individual law to become effective.
<i>Kultus Ministry</i>	The Bavarian ministry for education)
<i>MPI</i>	The Max Plank Institute, part of the Max Planck Society, a research society.
<i>Land/Länder</i>	The 16 different states in the Federal Republic of Germany
<i>TUM</i>	University of Technology in Munich
<i>ZIB</i>	The Centre for International Student Assessment, founded by BMBF and KMK

International large scale assessments (ILSA)

FIMS First International Mathematics Study

SIMS Second International Mathematics Study

TIMSS Third International Mathematics and Science Study, carried out in 1994-95. Became a recurring study, carried out every four years and the acronym now stands for Trends in International Mathematics and Science Study

IGLU German Abbreviation for PIRLS

PIRLS Progress in International Reading Literacy Study, carried out every five years

PISA Program for International Student Assessment. Started in 2000, carried out every three years

Secondary educational paths in Germany

Gymnasium The most advanced educational path of secondary school

Realschule The middle path of secondary school

Mittelschule The least advanced path of secondary school in Bavaria, former *Hauptschule*

Förderzentra Schools for children with special needs

Wirtschaftsschule Bavarian form of public school with special aim to train for business and economy

Trade unions in Bavaria

GEW Trade Union for Education and Science (*Gewerkschaft Erziehung und Wissenschaft*)

BLLV the Bavarian Teacher Union (*Bayerische Lehrer und Lehrerinnenverband*)

Specific German words

Bildung Cultivation, incorporation of culture

Testieritis The overdoing of tests in education

Drittmittel External funding

List of diagrams, schemes and tables

Diagram 1. Percentage of children with academic and nonacademic educated father, at secondary and tertiary school	25
Diagram 2. Average number of days lost due to strikes and lockouts.....	29
Diagram 3. Estimated number of births in Bavaria until 2030	29
Figure 1. Interplay of qualitative and quantitative data	18
Scheme 1. The research design	6
Scheme 2. The Quality Areas in QAs external evaluation	45
Table 1. The development of <i>Erziehungswissenschaft</i>	15
Table 2. Overview of significant events in the German educational context	22
Table 3. Percentages of students in 8 th grade in Bavaria divided on secondary educational paths 1990-2013	26
Table 4. Percentage of Bavarian students in 4 th and 5 th grade going to Gymnasium and Realschule.....	55

Table of Contents

Prologue

1. Introduction.....	2
1.1 Background.....	2
1.2 The relevance of Bavaria as the case study.....	3
1.3 KMK and Bavaria´s tripartite system	3
1.4 The study	4
1.4.1 Statistic	5
1.4.2 Design of study	5
1.4.3 Research questions	6
2. Three central concepts.....	7
2.1 Quality, standard and assessment	7
2.1.1 Quality	7
2.1.2 Standard	8
2.1.3 Assessment.....	9
3. Frame narrative and literature.....	9
3.1 The early years of a quantitative research field	10
3.1.1 Characteristic and condition	10
3.1.2 The validity of `evidence based` measurements.....	11
3.2 The shift and development of indicators	11
3.2.1 OECD, an educational actor.....	12
3.2.2 Education in OECD, from national to international.....	13
3.3 ILSA, conflicts of ideological nature	14
3.4 The introduction of ILSA in Germany, the rise of EB.....	14
4. Methodology and theoretical lens	16
4.1 Choice of actors and demarcations.....	16
4.2 Pierre Bourdieu´s field working theory	17
4.3 Critique of Pierre Bourdieu´s work.....	18
4.4 Method.....	18
4.5 Validity and reliability.....	20
4.6 Ethic consideration.....	20
5. Results and findings.....	21
5.1 The Bavarian State Ministry for Bildung, Kultus, Science and the Arts.....	22

5.1.1 Higher civil servant in the former Bavarian State Ministry of Education (D1)	22
5.1.2 The beginning	23
5.1.3 PISA 2000.....	24
5.1.4 Standards and epistemological implications.....	25
5.2 Trade Union for Education and Science (GEW) and the Bavarian Teacher Union (BLLV).....	27
5.2.1 GEW (D2), BLLV (D3, D4)	27
5.2.2 PISA and economy.....	28
5.2.3 The focus on measurement.....	30
5.2.4 Standards, evaluations and accountability	31
5.3 Institutes of quality assurance	31
5.3.1 ZIB (D5), TUM School of Education (D6).....	32
5.3.2 Quality and selection of competitive criteria	33
5.3.3 The rise of EB.....	34
5.3.4 The aftermath and consequences for practice	35
5.3.5 Communication of results	37
5.4 PISA- Scientific knowledge or fabrication.....	38
5.4.1 Experts: `Opponent` (D7), `Proponent` (D8)	38
5.4.2 Bias and stressed educational deficits	39
5.4.3 Quality and interpretation of PISA	41
5.5 State Institute for School Quality and Education Research (ISB)	43
5.5.1 QA, the Quality Agency (D9).....	43
5.5.2 How the concept of quality were introduced	44
5.5.3 The daily work at QA	44
5.5.4 Cause and effect-(too) complex matter	46
5.5.5 Quality or just monitoring	47
5.6 The Institute for Educational Quality Improvement (IQB)	48
5.6.1 IQB (D10)	49
5.6.2 Storing data for ... ?	50
5.6.3 Replaced Bildung and a `place in the sun`	51
5.7 The practitioners, school level	53
5.7.1 Teacher at a Gymnasium (D11). Changed curricula	53
5.7.2 Educational stress.....	54
5.7.3 An alien assessment regime and subjective evaluations	56

5.8 Humanistic Gymnasium.....	58
5.8.1 Principal (D12) and teacher (D13). Transformation of awareness.....	58
5.8.2 The `rush` for qualified education.....	59
5.8.3 Quality and evaluation	60
5.9 Natural Science and Technical Gymnasium	61
5.9.1 Teacher (D14) in <i>Geisteswissenschaften</i> and change in societal (value) structures.....	62
5.9.2 An increasingly diverse and complex student community.....	63
5.9.3 <i>Elternwille</i> and the notion of Market	64
5.9.4 The task of making practice work	65
6. Discussion	66
6.1 Change in the balance of power.....	66
6.2 Testing as part of <i>upgrading</i> education	67
6.3 The quality concept	67
6.4 The affect on practice.....	68
6.5 Altered power relations	69
Epilogue	
Limitations and directions for future research	71
References.....	72
Secondary literature/sources.....	89
Articles.....	89
Appendix 1.....	90
Appendix 2.....	91
Appendix 3.....	92

Prologue

In the beginning of the 21st century all of the German *Länder*² had within a few years established institutes responsible for the evaluation of the *quality* of their education systems. Since education is the domain of each Land this rapid and uniform establishment was a profound exception caused by the bad German results in the Program for International Student Assessment (PISA 2000). It was the beginning of the era of International Large Scale Assessments (ILSA) and its herald, Quantitative Comparative Education (QCE). The outcome have been an incessant production of quantitative data in which specific terms from economic theory have been used and reconstructed (cf. Ozga, Dahler-Larsen, Segerholm and Simola, 2011, pp. 1-3). The specific historical backgrounds and cultural settings of nations and regions have been marginalized in favor of a race for competitiveness. Decontextualized concepts like quality, standards and assessments have been adopted, concepts that have been advocated by International Organisations (IOs) such as the European Union (EU) and the Organisation for Economic Co-operation and Development (OECD) (cf. Varjo, Simola and Rinne, 2013, p. 51).

² Land/Länder is the term for the 16 different states in the Federal Republic of Germany

1. Introduction

The EU is a supranational organization in which summits and meetings are held and cooperation and agreements are worked out. On the 23rd and 24th of March 2000, a summit was held which would have a major impact on education, the Lisbon summit. It marked the shift towards a stronger and more unified European education policy and involved more than 10 concrete objectives with timetables (e.g. Grek and Rinne, 2011, p. 23). The president of the European Parliament, Mrs. Nicole Fontaine, outlined the main topics of the coming discussions:

The European Union is confronted with a quantum shift resulting from globalisation and the challenges of a *new knowledge-driven economy* [...]. (European Parliament, 2000)

Knowledge economy is based on the taken for granted assumption of individuals as *Human Capital*. ILSA rests implicit on this notion in way of perceiving education as not only measurable in quantifiable units, but as part of a market where knowledge becomes a capital, an asset that can be increased and used as a competitive item, by individuals and states in a competitive race. The central term Human Capital is built on ideas of economists but officially, it was Theodore W. Schultz in his presidential address to the American Economic Association 1960 who articulated the term and positioned it in the domains of education and research. The use of the term was criticized by Prof. Harry G. Shaffer on the grounds that people in education and researchers, were not merely driven by economical profits and future gains, and concluded that the term was in danger of being utilized by policy maker, legitimizing a market approach in ways of cost benefit analysis. Schultz comment was,

However, if any new knowledge were obtainable by the use of these concepts [investment in man and human capital] despite the empirical difficulties, Shaffer appears to believe that such knowledge would be grossly misused – by implication, more so than other economic knowledge – in making policy decisions. This view of the relation between economic analysis and policy seems unreal and irrelevant. (Schulz, 1971, p. 50)

1.1 Background

In the late 50s, comparative education began to establish itself as a discipline of its own in research and gained ground when IOs such as OECD set up branches of educational research (e.g. Eide, 1990, pp. 5-6; Papadoupoulos, 1994, pp. 21-24; Husén and Tuijnman, 1994, pp. 4-6). The development of comparative education was not exactly straightforward and the discipline's search for educational indicators often met critic and resistance. However, in the 90s the tide had turned and it was the decade when a series of ILSA began to make a wide and noticeable impact, especially the Third International Mathematics and Science Study (TIMSS) 1995 and the PISA 2000³. Germany had by rule not been involved in ILSA, only exceptionally as in the International Association for the Evaluation of Educational Achievements (IEA) Second International Mathematics Study (SIMS) 1980-82. This reluctance was mainly due to methodological uncertainties of validity, reliability, methods and costs (Ackeren, 2002, p. 162). Though, in 1995 this changed. After a decision in the standing *KultusMinisterKonferenz* (KMK⁴), Germany took part in two of the measurements of IEAs TIMSS 1995. The disappointing result was a surprise for many in the education sectors and experts but remained mainly outside of common mans knowledge. The outcome was the so called *Konstanzer* decision in 1997 where KMK declared, that [...] *the development of actions [or measures] to secure the quality of education is an important task* (KMK, 1997a). As a consequence Germany took part in the PISA study 2000, PISA-I⁵ and PISA-E⁶. The results published in 2001 (PISA-I) and 2002 (PISA-E), became known as

³ The PISA-study was published 2001 (PISA-I) and 2002 (PISA-E) but the work was implemented in the late 90s. However, PISA had in another form been tried before, in the 70s. See appendix 1.

⁴ KMK is the standing conference for the ministries and coworker of culture and education in the 16 German Länder.

⁵ PISA –I stands for PISA International and is a country specific study.

the infamous PISA-Shock. The results stirred up an intensive media and public debate where politicians were the targets, especially those responsible of education in the Länder. However, Bavaria was one of the rare exceptions. Their result in PISA-E, was the best in comparison and matched many of the partaking countries in PISA-I, and Bavaria remained for many years one of the most successful Länder in the German test comparisons (PISA Ergebnisse, 2013).

1.2 The relevance of Bavaria as the case study

The study's aim is to explain how three core concepts of QCE – quality, standard and assessment – are contextualized when transferred into an education system, authorized by a state's partaking in ILSA. In specific, the study tries to contribute to in-depth and definite knowledge about the affect ILSA and its endorsed epistemology has on a participating state's education system in matters of inferred core concepts. It is therefore based on a state case study, hence the study is context dependent (Bavaria) but not context *exclusive* and the selected state is part of a larger context; a federal state, Germany, that will necessarily be an intertwined part of the study. Statistics, literature and interviews often involve these two levels. Case studies have a mixed reputation but can be an important part to obtain complex knowledge – *reality* is complex as well as research of – accumulated through a step by step process in which the researcher develop the knowledge and personal skill correspondently (Flyvbjerg, 2006, pp. 3-5). However, the choice of case study must always be carefully chosen and possible to justify. Bavaria was chosen as the specific case study and the *raison d'être* was that the study's findings (if any) would gain in credibility if the object could be viewed as less probable to the objective of the study, not predisposed to validate the outcome. There were specifically four reasons for selecting Bavaria, all of which marked Bavaria and its education system education system as likely to be one of the more resistant towards external influence.

- Bavaria had been a top performer in the comparative measurements of the Länder (for example top position in the first series of the PISA studies)
- The education system of Bavaria had a reputation for excellence in Germany (cf. Rotte and Rotte, 2007, p. 293)
- The Christian Socialist Union (CSU) has governed the state continually since its foundation in 1946 (except a three year episode in the 50s) and thus, the education system has historically eluded political strife and transformations
- Bavaria was considered '*the Engine*' of German industry, that is, Bavaria is one of the most successful economies of the Länder in Germany (low unemployment rate, economical stability, so forth) with a successful transformation from an agricultural economical base into a highly industrialized region (EU.2007, no date)

In short, the choice of case study was based on the premise that the Bavarian education system had positive performance indicators, a high reputation, and was part of a state with a continuous political state order and a strong state economy in which the political and economic fields were closely related to each other. Furthermore, in Bavaria education was a strong determinant for the individual's future options visual in the state's traditional German tripartite system.

1.3 KMK and Bavaria's tripartite system

The Federal Republic of Germany consists of 16 Länder with high autonomy in cultural and education affairs, the education systems therefore differ between the Länder. KMK attended by the ministers and senators of the Länder responsible for education as well as higher education, research, and

⁶ PISA-E stands for PISA Enlarged and is an optional in-country specific study.

cultural affairs, and the federal minister of education, coordinates in some degree the education systems and policies together with the legal framework of the federal constitution (Baumert, Ardelt, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Tillmann and Weiss, 2002, pp.39-42; KMK, no date).

In Bavaria the state primary school includes grade one to four and the main form is called *Regelklassen* but there are special classes depending for example on the children's ability in the German language. There are also other school forms such as private schools with fees and so called *Förderschulen* for children with disabilities. In the 4th grade the children's grades decide which secondary school form they will attend⁷ and there are three main forms of secondary schools: *Haupt/Mittelschule* (until 9th -10th class), *Realschule* (10th class) and *Gymnasium* (12th class) (Bildung und Sport, 2011). The *Gymnasium* has traditionally been the most prestigious form and a successful final exam (*Abitur*) provides the students with an entrance pass to High schools and Universities. The *Realschule*, as second in position, also provides in the prolonging the possibility for studies in certain subjects (natural sciences) in High schools and Universities. The *Hauptschule/Mittelschule* has witnessed a drop in number of students in Bavaria depending on demography and the increasing percentage of children in *Realschule* and *Gymnasium*, caused by an educational discourse where education is expected to be the solitary pronouncer of future success or failure, in a most differentiated and competitive job market. The drop resulted in an education reform which started in 2010 where *Hauptschulen* began to be reorganized into *Mittelschulen*⁸. Officially it was declared that the outcome would be a school form with a wider range of subjects, learning possibilities, and exams with the individual in focus, as a way to improve each child possibilities (Bayerische Staatsregierung, 2010). One specific intention, often repeated, was the safeguarding of the numerous small *Hauptschulen* in Bavaria, partly by their merging. The critique targeted the reform as just a facelift to make the school form more popular among the parents with no real enhanced possibilities for the children and leaving the remaining *Hauptschule*, those unable to meet the new demands brought on by the *Mittelschule* reform, to wane.

1.4 The study

ILSA is generally perceived as evaluations of the quality of an education system or part of. There is however an alternative but less regarded dimension, how the partaking in ILSA is linked to a transferring of core concepts related to the epistemology of QCE with roots in economic theory. The study's aspiration is to contribute to in-depth knowledge of the consequences when a state or country participates in ILSA, that is, the effect on the education system in terms of the transferring of central concepts based on economic theory and how these become articulated and constructed. The study is empirical, based on official texts, statistics and literature. But a primary part of the data comes from interviews which are designed as to let central actors of the Bavarian education system – national and state education monitoring institutes, state administration, experts and school management, and teachers – speak from their own experiences and their own perceived realities.

The study rest on the assumption that the tripartite concept of quality (abstract), standards (abstract made concrete) and assessments (concrete) is positioned in locus of educational fields when, and only when, there exist a hegemonic role model of education as a market, where the individuals worth (abstract) are units able to *valuedate* (concrete). The maintenance of this stance demands the involvement in measurements because of the proclaimed need for comparing educational output in quantifiable terms to infer or increase efficiency, effectiveness and competitiveness. At the same time distrust is made official, that is, education monitoring is based on the assumption that quality control of the management of education is necessary (e.g. Schmidtlein, 2010, pp 264-265). This

⁷ This is the main principle but there are minor possibilities to alter or postpone the decision depending on reexamination, examination in 5th or 6th class special school form, age and parental decision.

⁸ The term *Mittelschulen* was in fact used before 1965 when the name of that school form was changed to *Realschulen*.

instituted requirement depends on the logic of the changed value structure where education centers (kindergarten, schools, secondary education, etc.) become semiautonomous accountable units where the accountability must be verified. The traditional (limited) trust becomes the (unlimited) distrust the lower the position of the unit, for example primary schools, placed on distance from the power center. The accountability in this form signifies the gradually unlimited responsibility of educators and thereby initiates an increasing struggle and antagonism between (related) actors in education – as principals and teachers – because of the self-interest to shield one's own group from accountability by increasing its capital. Jarl, Fredriksson and Persson (2011, pp. 429-443) gives an example of this process but draw the conclusion of a professionalization process, however that is just the outer sign – the consequence – of an innate dilemma (cf. Brüsemeister, Preuss and Wissinger, 2014, pp. 225-228). Since educational systems are national systems, or state systems part of national systems, the legitimacy of the claim that education has become a realization of an economic competitive market can be reflected in national educational agencies' and prime actors' articulation of excellence in education to secure the nation's or state's future, the market position (e.g. BMBF, 2008, pp. 5-7; KMK, 2015a). This vision, when brought into the national or state educational context, shift the perception of education from the classical diversified social institution, based on incorporating individuals into a society's different strata, to a view on education as part of an accountable production arena pushed by incentives for best practice. New Public Management (NPM) can be said to represent the embodiment of powers included in this transformation (or game) of the market approach, and change of education into accountable and monitored semiautonomous entities. Hence, the market approach is advocated by several actors who have in common that they will, or expect, future gain, may it be ideological (symbolic) or material (economic). This paradigmatic vision of education is not created on the local level but brought into the regional context by stakeholders on different levels such as IOs, national policy actors, politicians, industrial/business associations, unions and other interest organizations.

1.4.1 Statistic

The study makes use of official statistics to underline and describe changes that is of relevance for the study's aim. However, the data presented are not in line with any epistemology bias of cause and effect or position taking between qualitative and quantitative research. On the contrary, the statistics is just one piece of a complex puzzle that together with the other sources serve the purpose of a creating a less subjective description of *reality* compared to the reality of statistics, where individuals are referred to as belonging to pre- or post constructed categories (e.g. Lindblad and Popkewitz, 2011, pp. 8-9; Jenkins, 2007, pp. 59-61).

1.4.2 Design of study

This overview of the design is meant not just to present an outlay of some of the key stakeholders in German education on a federal and state level, it also illustrates the actors' positional stand and relational structure (arrows). However, there are other actors such as private corporations involved in education, selling specific knowledge and educational tools, or institutions and associations with specific interest in the educational sector (in the design called agents of transferred norms/discourses⁹). Some of these do appear in the study but the main actors are those from which the informants came, in the scheme marked with a number. In Bavaria they came from the former ministry of education¹⁰, the department of quality (QA) of the Bavarian state institute that supports and advises the Bavarian Ministry for education (ISB), and teachers and management from the school praxis. Included in the state level are also informants from the Trade Union for Education and Science (GEW) and the Bavarian Teacher Union (BLLV) as also experts – one mainly empirical educational researcher and one expert in informatics (QCE methodology) – based in Bavaria. The Centre for

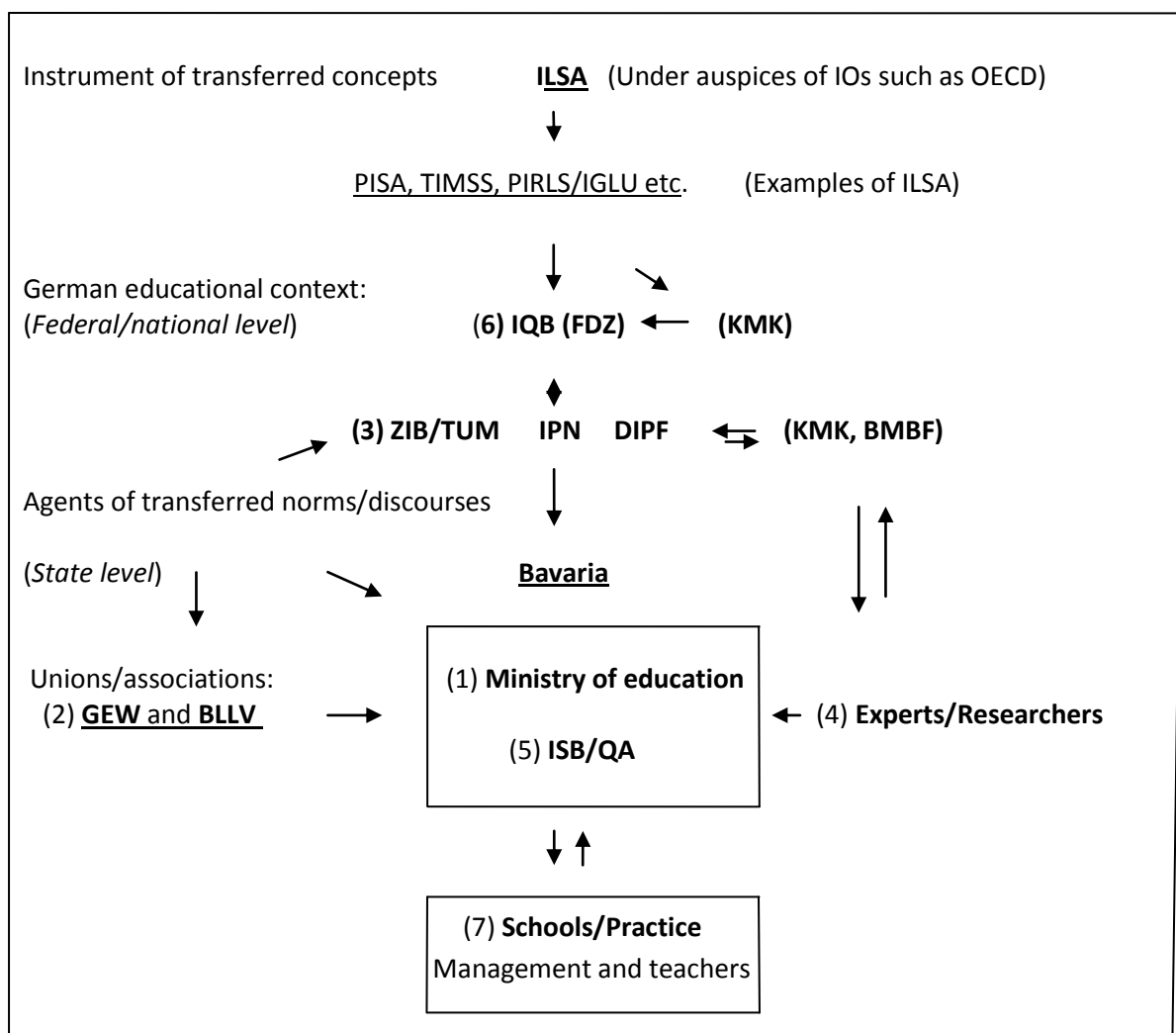
⁹ The reason is that these institutions already have transposed the methodology of ILSA.

¹⁰ Responsible for the introduction of ILSA in Bavaria.

International Student Assessment (ZIB) is located in Bavaria, Munich, but belongs to the federal level with funding from both the Federal Ministry of Education and Research (BMBF) and KMK. The Institute for the Development of Quality in Education (IQB) has a special position, funded by the KMK but also in a degree acting as a free agent with its data information center (FDZ).

The different levels in the scheme – which is not always as clear cut in practice – is meant to establish a relational structure that form the base and practical guidance of the study. ILSA are the locus of a *seismic structure* that in theory has repercussions on each of the educational sublevels. The agents of transferred norms include the Leibnitz Institute for Science and Mathematics Education (IPN) and the German Institute for Pedagogic Research (DIPF), both of which are parts of the ILSA methodological infrastructure in Germany, grounded in EB. They are also state funded, that is, not dependent on the formal organization of tertiary education.

The actors with numbers represent the informants and their specific number signifies their individual order in the study.



Scheme 1. The research design.

1.4.3 Research questions

The aim of the study is to explain the effect on an education system in terms of the transferring of central concepts of ILSA and how they become rearticulated and reconstructed. However, a prerequisite is that the methodology (theory and methods) of QCE must have become authorized, that is, internalized in the administration of the education system, the bureaucracy of the state. The

elevation of QCE will at the same time affect the different levels in the education system in matters of positions and dispositions¹¹. Hence, the central issues are

- To what extent have the status and position of EB in Germany changed?¹²
- How are the tripartite concepts of quality, standards and assessments articulated and constructed?
- What are the positional and relational changes of the different levels in the education system?
- How have the normative value system in education been effected in terms of decreasing or increasing of the different levels` forms of capital?
- What are the effects on the working practice?

2. Three central concepts

Quality, standard and assessment are three terms that are central to ILSA and QCE, terms that become concepts when materialized into a specific context, the education system. These concepts are also interrelated and rest implicit in each other's subtext. The field of QCE, as all fields, is dependent on a specific language in which certain words are fundamental, although they can be known and used in general they convey a specific value – a specific discourse – when used in a particular field (Bourdieu, 1991, pp. 164-170). These words do not just hold position because of their significance in the particular field, they instead manifest themselves as bearer of specific theoretical and imaginary values that are at the locus of the field and have implications outside the field; they are concepts and act as a form of incorporated cultural capital (cf. Bordieu, 1997, pp. 218-220).

2.1 Quality, standard and assessment

In order to make way for the study's intention, I will try to explicate a brief sketch of the background of each term, and how they can be understood when used in QCE. The term standard is of special interest for the German context and will therefore be addressed in more detail. It is a way to unveil the transferred discourse by these terms into education and the related interdependency the stated definitions brings on them. It will, hopefully, make clear the logic, the logic of practice.

2.1.1 Quality

Quality was a word used in economy to describe a product's value based on the consumer's judgment. Later, in the industrialization, with the mass production, quality became equivalent with the management and supervision of the production whereas in modern time, in high technology production, quality has become a term for development (cf. Ozga, Dahler-Larsen, Segerholm and Simola, 2011, introduction). In that sense, the meaning of quality has changed from something permanent to temporary though preserved a positive value, it has always been something desired. Harvey and Green (1993, pp. 10-11) describe the term as often understood as a relative concept in two ways: it is related to the used context and the person that use it. In the field of education, quality has become a motto that is extensively used. At the same time, the term has become synonymous with the lack of explicit criterion and distinct definition, in education as well as in educational research (c.f. Ditton, 2000, pp. 73-77). Though the term has a positive value, it is also very much related to societal value constructs, values that are time specific. Thus, the very use of the term in education would have been highly *contested*, had it been in the early 70s, however, today quality is part of the Holy Grail of education: *'Quality education is the most valuable asset for present and*

¹¹ Perceived and benefit-based group membership.

¹² Since ILSA is based on QCE, in Germany EB, the participating and promotion of ILSA must in that sense be possible to trace in a correlative promotion of EB. If opposite, the object of the study is falsified, therefore this question is central.

future generations` (OECD, no date/a, p. 3); `The core tasks of the KMK include quality development and quality assurance in higher education` (KMK, 1997b); `The National Agency for Education also ensures that Swedish education maintains a good standard of quality` (Skolverket, 2014). The term quality has been brought into the center of an educational discourse of individuality where it is utilized to describe a set of performance indicators – efficiency, effectiveness and competitiveness – reflecting its economic origin.

2.1.2 Standard

In the first decade of the twentieth century, Frederick Winslow Taylor constructed a theory on how to increase the efficiency and effectiveness of the output of industry called *Scientific Management*. One of the main components in the management idea was to set up standards – proper amount of outcome¹³ – as a way of securing the quantity of production (Taylor, 1911, pp. 31-35). Standards had been in use much earlier but the idea to tie standards – quantity and equal quality – with the output of a production chain was formulated through the idea of scientific management. Work became a series of time measured interrelated but separated components influencing not just the production of the industry, it made measureable quantifiable units the model of standards (e.g. Ozga, Dahler-Larsen, Segerholm & Simola, 2011, p. 2).

Content standards and performance standards are two definitions of standards that can be said to encapsulate the core of the transformation of education in the 21st century, from input based (content), what should be learned, to output based (performance), what has been learned. Performance standards are often referred to as normative standards, what is considered to be *normal* and based on summative assessments. These summative assessments can be of three kinds: *Maximal* which is the highest level of competence, *minimal* which is the lowest level of competence level and *rule* described as the average level of competence¹⁴ (KMK, 2004, pp. 8-9). Standards have become evaluation criteria thought to visualize the performance of students and schools, and in its prolonging the education system. However, it is not an unambiguous concept for what is performance in education? And are there really cohesive standards perfectly adaptable to different educational systems? Is it even possible to construct standards that are valid intranational? In Germany achievement in education had traditionally been judged not by *Leistung* (performance) but by *Bildung*. *Bildung* (app. cultivation) can be defined as including learning, personality, development and subject specific education in all. Educational steering by performance measurements had not been part of the German educational context. Consequently there was an extensive critique, in politic and research as well as in media when Germany in the *Konstanzer Beschluss* (Konstanzer Decision) decided to incorporate the findings of TIMSS 1995 and the (coming) OECD study PISA to make way for inter- and intranational comparisons by measurements (e.g. Overesch, 2007, p. 234; Kuhlmann, 2012, pp. 13, 21). In other words, education based on subject specific knowledge was replaced by general education, in German *Ausbildung or Basiskompetenzen* (Münch, 2009, pp. 29-31; Hartong, 2011, pp. 323-325). The outspoken prime objective was to make it possible to evaluate and compare the German Länder's different educational systems. In the course, the *Bildung* concept in education became replaced by a *Leistung* concept, so instead of *Bildungsstandards* it was *Leistungsstandards* (performance standards) which became the standard model, even if the term *Bildungsstandards* still remained (KMK, 2006). In sum, the break through of international comparative measurements with their base on and focus of, measurable units, quantity by numbers, meant that summative performance standards became the norm of standards in Germany.

The *Bildungsstandards* adopted by the KMK, take hold on general educational goals and specify the competences, the students in specific age groups, should have acquired. The *Bildungsstandards* concentrate on the core area of each subject, and describe expected

¹³ In original it says `a full day's work for a first class laborer` and `a proper days work`.

¹⁴ The specific German term is *Regelstandard*.

knowledge. [...] a mix of content- and output standards. They are based on the expected average level of the students in the end of 4th grade [...] and are thereby Rule standards. (KMK, 2004, p. 9) (Authors translation)

That is, educational standards as related to goals and targets, to age groups and subjects, and defined as summative knowledge.

2.1.3 Assessment

To gain insights into institutions as well as individual's level of skill and knowledge, *assessment* or *evaluation* can be used. However, these two words are often interrelated, intermingled and sometimes interconfused, that is, these words can be used in the same context in such an extent, that even professionals that inscribe one content actually mean the other (cf. Stake, Contreras and Arbesu, 2012, pp. 3-4; OECD, 2015, p. 1). Assessment and evaluation are common in education and frequently used in educational research. Assessment is a long-term process in which information related to a particular outcome is methodical collected, analyzed and interpreted. Evaluation is to present and analyze information regarding quantitative or/and qualitative results at a given moment (e.g. Secolsky and Denison, 2012, preface; Suskey, 2009, chap. 3). Another example of the peculiar way the both expressions are used can be found in the OECD country report series *Reviews of Evaluation and Assessment in Education* where the interpretation instead underlines the dichotomy: individual vs. collective/organisation.

This report differentiates between the terms "assessment", "appraisal" and "evaluation". The term "assessment" is used to refer to judgments on individual student performance and achievement of learning goals. It covers classroom-based assessments as well as large-scale, external tests and examinations.[...] Finally, the term "evaluation" is used to refer to judgments on the effectiveness of schools, school systems and policies. (OECD, 2011b, p 32)

However, evaluation is a fixed judgment at a given moment and assessment is a kind of recurring diagnostic that can illustrate a trend, progress or regress, but not explain it. The result of ILSA is in its perfect form considered as diagnostic information of the state and rank, and development – at least so interpreted – of an education system or part of. Nonetheless, opponents view ILSA (at best) as a kind of evaluation where results are non-comparable.

3. Frame narrative and literature

The frame narrative will give a historic exposé of the rise of a research field which laid the base for ILSA, also referred to as *empirical research*. The reason is that a study of ILSA and the transfer of the concepts quality, standard and assessment, based on the field working theory by Pierre Bourdieu required a historic reference frame that could explain how ILSA became predominate. Without this frame, it could not be understood on what premises ILSA are based, the value structure behind closely linked to a change in the societal value structure, exemplified in the study by the influence on a state's education policy construct. It is the narrative of QCE. Furthermore, the frame narrative is thought to establish the context for the embedded narrative of the study, a guidance of the reading. It is a narrative of the historicity of a specific research niche which gradually internalized an economic epistemology and developed an analogous methodology, with roots in psychometric and econometric, what we now know as (Q)CE. It will give an historic account of the progression in its early phases and its successive introduction into the German context.

3.1 The early years of a quantitative research field

An important contextual component is how ILSA come to be regarded as *the* legitimate evaluator of education systems which is the story of the positioning of comparative education¹⁵. The literature on this chain of events can be divided into written accounts by participants in the process and observers of it. Together these narratives form a story of a cumulative series of events, leading to a positional stand for QCE as *the* objective science, a tool that could be used quantitatively to examine education system.

The start of QCE is referred to as the late50s, when researchers in the field tried to work out the possibility of empirical large-scale studies (Foshay, Thorndike, Hotyat, Pidgeon and Walker, 1962, p. 7; Husén, 1973, foreword). It was the beginning of a quantitative approach and it was tied to an IO, UNESCO, which partly funded the attempt and also provided space and forums for the partakers. The link between IOs and comparative research would form a pattern throughout the years. It was research educationalists – scholars, educational psychologists, sociologists, and psychometricians – that met at the United Nations Educational, Scientific and Cultural Organization’s (UNESCO) Institute of Education in Hamburg. The main object of research was how *‘to deal specifically with problems requiring the quantitative approach, such as failure in school, examinations and evaluation’* (Husén and Postlethwaite, 1967, p 26; cf. IEA, 2014). The outcome was the so called Pilot-country study which concentrated on 13-years old students on a scale of about 10 000 total in twelve countries and tested the educational achievements in four subjects: mathematics, reading comprehension, geography and science with an additional non-verbal ability test (Foshay et al, 1962, p. 10; Härnqvist, 1975, p. 86; IEA, 2014). The feasibility study was the first ILSA and it marked the beginning of educational comparative centers, governmental and non-governmental, first of which were the International Association for the Evaluation of Educational Achievement (IEA). Husén and Postlethwaite (1967, pp. 25-26) describe that comparative research had been done before the Pilot-country study but mainly qualitative work focused on national educational fields, conducted by individual researchers or small groups thereof, occasionally instituted by IOs such as the publications from the International Bureau of Education (IBE) or *The World Surveys* by UNESCO and studies by the OECD. However the Pilot-country study and the formation of IEA had a major impact on the character of these studies: at the onset qualitative, ranging from descriptive over juxtaposing to analytic, then transformed to mainly quantitative, from exploratory to explanatory. Hence, the generally motive from start, the explorative character of these studies, was developed into a set of objectives that stressed the explanatory possibilities of cross-national comparisons (Härnqvist, 1975, p. 87).

3.1.1 Characteristic and condition

One characteristic for ILSA is the stress on what is termed explanatory factors (or visual indicators) with an emphasis on comparable numbers or quantities. In so doing, the results are always at risk of being used by actors in education in terms of oversimplified, visual charts comparing education systems. This dilemma is an integral part of ILSA, since large-scale assessments always have been labor intensive and high-cost enterprises aspects of utility is central, what motivates the use of such studies (e.g. Bottani and Tuijnman, 1994, pp. 68-69; Aljets, 2014, pp. 183-184). Conversely, the providers of the necessary and required financial support have a saying in not just the design of an assessment; they also influence the desired outcome (Tröhler, 2013, pp. 149-152; Eide, 1990, p. 28). Hence, there is a close symmetry, community of interest, between those conducting ILSA and the financiers of. This is amply exemplified in IEA: funded by the member states governments who has an interest in a comparative outcome. Margaret Browne’s analysis of the result from an IEA study, that it *‘[...] is inevitably a compromise between academic outcomes and pragmatic results desired by national governments’* (Brown, 1996, p. 194), is an overall reality for all international large-scale

¹⁵ The design of comparative education into QCE.

assessments (e.g. Kellaghan, 1996, pp. 155-156). George Papadopoulos, a former director of education within OECD, describes how the acceptance of education matters depended on the perceived relevance for the educational sectors in each country and the responsible Ministries, and the prime issue was always the financing (Papadopoulos, 1994, pp. 12-13). That is, the selling image of a programme was a key factor, and still is for such activities as ILSA which are dependent on external funding. Hence, the outcome must be perceived as policy relevant (cf. Bottani and Tuijnman, 1994, pp. 68-69). The key elements in ILSA, indicators, must therefore be based on to what extent they are policy relevant, hence subjectively chosen based on utility. Bottani and Tuijnman (1994, p. 49) describe education indicators as bearers of statistic information of education systems and by that, having a political function: they are value laden containing a value judgment. Thus, ILSA always risk being judged as biased which make it necessary to proclaim the measurements as evidence based and the possible shortcomings of measurements as belonging to the test users, not the tests. Educational researchers often explain the accuracy of comparative education as belonging to the rigor of methodological sophistication, displayed as an evidence based concept (e.g. Hegarty, 2014, pp. 47-55). However, this view shades an important tale of the development of assessment's validation.

3.1.2 The validity of 'evidence based' measurements

The apt example is the developed – and changed – definition of validity in measurements, in other words, that measurements measure what they are supposed to measure. Brennan (2006, pp. 1-3) trace the initiation of the practice of validity in measurements as beginning in the 1950s. At the start it consisted of two parts, relevance and reliability, used as a way of predicting the correctness of the results. In the mid 50s relevance was excluded and different kinds of validity were the topic: content, predictive, concurrent and construct. Then in the 60s the predictive aspect disappeared, the view changed from predictive to interpretation whereas in the beginning of the 1970s, the validation of the test itself was replaced by interpretation of the underlying data. Mid 70s, the now three aspects – content, criterion and construct – were called types of evidence instead of validity and the construct type was in center. The trinitarian model (content, criterion and construct) were in late 1980s replaced by sources of validity *evidence* including evidences based on consequences of test use (cf. Kane, 2006, pp. 18-24). The term evidence became established and the view of measurements correctness shifted from evaluating the correctness of the test, validity of content, to the correct application of the test, reliability of practice. The result was that the responsibility gradually shifted towards the user.

It was a prerequisite for the expansion of comparative education that central actors and decision makers in education perceived QCE as valid, in sense of being an unconditional, fair and impartial 'judge' of the *quality* of education systems (e.g. Mazzeo, Lazer, and Zieky, 2006, p. 681). Thus, it implies and assumes a transferred value system – an illusion – on education (as) based on comparable numbers, manifested by political statements linking economical growth to the search for best educational practice. It also affect research practice to include comparative methods to improve the probability of funding '[...] for instance, the case of the European Union where the 'comparative criterion' is [now] a requisite for financing social research' (Novoa and Yari-Mashal, 2003, p. 425).

3.2 The shift and development of indicators

Indicators are the central components of QCE and they are methodically chosen for the main purpose of each measurement. Education indicators are signifiers of special interest, '[...] *policy-relevant statistic designed to provide information about the condition, the stability or change, the functioning, or the performance of an education system or a part thereof*' (Bottani & Tuijnman, 1994, p. 48). However they are just chosen signifiers of something *thought of*, indicative of a phenomenon, not the phenomenon itself. In the 1980s the term *educational output* became established. Instead of just defining educational input, what to learn, the concerns were now focusing onto the setting of

educational standards, levels of must be learned. Strategic planning and budgeting gave way for quality control and steering at a distance, in short the predominance of output orientation (Veld, Fussel and Neave, 1996, p. 77; cf. Lindblad, 2011, p. 71). In the background a value shift had taken place where Human capital¹⁶, now reintroduced, had become a prime issue: manifested in cost-benefit analysis, teachers' salaries, value of knowledge and efficiency (Papadopolous, 1994, pp. 170-171; Lindgard and Grek, 2008, pp. 6-8). As a consequence, ways to assess students' performances came into focus, to assure that the established standards had been met, and work out valid international indicators.

The link between ILSA and IOs had always been present but it was reinforced at the Ministerial meeting in Paris 1988 where it was decided that the Center for Educational Research and Innovation (CERI) at OECD would make a feasibility study centered around an attempt to obtain a limited set of indicators of education systems, partly based on earlier attempts and experiences (IEA, no date; Eide, 1990, p. 46). The project that followed had the title *International Indicators of Education Systems* (INES). It was divided into four phases: Phase I-Exploratory (1988-89); Phase II-Development and construction of indicators (1990-91); Phase III-Shift towards regular production and use of indicators (1992-96, 1997-2001); Phase IV-Work on human capital, social capital, economic growth and sustainable development (1999-2000) (OECD-CERI, no date). In spite of the only preliminary set of indicators presented in Lugano, the outcome was far from ideal but it was published as *Education at a Glance: OECD indicators* (1991) and were so successful in spite of its shortcomings that a second study with 38 given indicators followed *Education at a glance II* (1993) (Husén and Tuijnman, 1994, p. 13; Saha and Tuijnman, 2003, p. 1127). It transformed into a series of well established OECD publications with an increasing impact. The INES project established and improved the reliability, comparability and political relevance of the indicator set, as well as developed new indicators and started up collaboration between statistical centers in UNESCO and the European Union¹⁷ (Saha and Tuijnman, 2003, p. 1128). Large-scale measurements were about to take off and that in an area which for long had been a neglected part of OECD.

3.2.1 OECD, an educational actor

The OECD convention does not state education as being the purpose or concern of the organization. However, the former deputy director of education within OECD Papadopoulos describes how matters of education corresponded well with OECDs prime and stated economic and social concerns, and therefore established itself quite early in the organization as an object of interest, although seen as secondary for a long time (Papadopoulos, 1994, p. 11). Another view on the development of educational interest in OECD is that it coincided with economical theories brought forward in late 50s, where education gradually was seen as an aspect of economy, the economy of education, exemplified in the economist Theodore Shultz theory¹⁸ on education as having an impact and influence on GDP (Eide, 1990, p. 6).

Two conferences were to have an immense impact on OECDs work on education: The Kungälv conference and the Williamsburg conference. OECDs first conference on education was organized in June 1961 in Kungälv, Sweden, by the *Office for scientific and technical personnel* with the Swedish ministry of education as collaborator, under the official title *'Ability and Educational Opportunity'* (Papadopoulos, p 23-24; Husén, 1987, p. 11; Husén, Tuijnman and Halls, 1992, pp. 36-37). The significance of the conference was that it: (1) instigated the perception of education as an international object, among member states in OECD, in sense of that nations were facing similar problems;(2) removed the older vision of education as a national or regional exclusive domain;(3)

¹⁶ The expression is foremost ascribed Theodore Schulz and used in the book "Investment in human capital: the role of education and of research" (New York, 1971). In the preface he acknowledges the influence of previous writers on the subject of capital-quality-labor.

¹⁷ Present day Statistical Office of the European Communities (EUROSTAT)

¹⁸ See also introduction.

inferred a more coherent vision on individuals as human resources that could be safeguarded through education (Papadopoulos, 1994, pp. 106-107). Economic growth was an important context of the conference and its outcome and the economic optimism of the era is mirrored in the growth rate that the very first meeting of OECD set, estimated to 50 per cent in the Gross National Product of its Member States for the decade 1960-1970 (Papadopoulos, 1994, p. 37). The linkage between increase/growth of the economy and expansion of the educational sector became obvious in the 70s, if not sooner, when reduction in economy meant hard constraints on and decrease of, the educational sector.

That education is now a central preoccupation of every nation in the world and, further, that educational plans can be carried out with maximum success only if they are made in relation to educational systems and plans in other countries. (Draft from the Williamsburg conference, chairman James A. Perkins, 1968, p. 3)

The Williamsburg conference was set up by the UNESCO institute and the department of the *International Institute for Educational Planning* (IIEP) and took place in Williamsburg 1967. The invited delegates included a wide range of experts, researchers and by then present and coming Ministers of Education from OECD countries. The momentous was that it initiated the perception of the importance of comparative and cross-national studies of education (Husén, Tuijnman & Halls, 1992, p. 37; cf. Perkins, 1968, pp. 3-12). The influence of this perception cannot be underestimated since many of the participants was or were to become, holders of important national, state and official positions in education. It was also the first education conference that underlined the necessity of regular statistics provided by experts: *‘Every educational system should regularly gather, analyze and disseminate accurate and up-to-date information [...] Trained statisticians are imperative. When they are not available, they must be borrowed from other countries, and when borrowed, native statisticians must be trained to take their place’* (Perkins, 1968, p.4).

3.2.2 Education in OECD, from national to international

The OECDs Educational Investment and Planning Program (EIP) was established in 1962 and it intended to contribute to the organizations work to assist the member states in their planning for educational development (Papadopolus, 1994, p. 47.). There had been a practice of general information within OECD – descriptive data on the member states different educational systems – in accordance with practice in economic policy, but the EIP established more solid foundations to review countries policies focused on national plan documents and therefore *‘offered real possibilities for a real examinations of the new directions in educational policies in member countries’* (Eide, 1990, pp. 18- 19). These examination teams consisted of professional and political experts. An important outcome was the establishing of educational planning groups in individual member countries who prepared the reports (Papadopolous, 1994, p. 48). Another significant result was that the economic – up to then – key aspect became less dominant, just one of a set of indicators. These reports came to include a self examination of each education systems policies in terms of objectives, planning methods, criteria, etc (ibid.). They paved the way for a change of the perspective of educational policies, considered as something internally country specific, now gradually shifting to an externally perspective, seen as important to evaluate internationally: to compare and learn from others¹⁹. The OECD became an actor in national education policies of member countries, positioned to consecrate favored policies through positive reactions (Eide, 1990, p. 19).

¹⁹ These national policy reports, first of which was Ireland, considered as a pilot study (*Ireland: Investment in Education, OECD, 1966*), changed character but the practice has remained up to date with the report series *Education at a Glance*.

3.3 ILSA, conflicts of ideological nature

One difficulty with the design of ILSA has been the practical difficulties to work out education indicators valid for more than just a single state or nation. Different education curriculum, content, criteria, (students') age groups etc. have caused recurring practical obstacles. Another recurring difficulty has been the conflicts among partaking nations, conflicts of ideological nature. Politics, as in the struggle for dominance of educational universal values, have always been at the center of ILSA design. Especially the developed economic – human capital – oriented value base has caused dissensions. An early example thereof concerning the education systems among the OECD countries, was the move in the end of 1960s until mid 1970, to change the member states education systems into a nine year common school, a similar comprehensive schooling, *'[...] a move which was explicitly supported by the OECDs examiners'* (Papadopolous, 1994, p. 97). The trouble was that in some member countries, the comprehensive schooling was strongly resisted, especially in the German speaking – Germany, Switzerland and Austria – where the traditional school was a highly selective tripartite system. In Germany the controversy was specific problematic because of its political and ideological implications: the Länder that were protagonists of comprehensive schooling was controlled by the socialist block (SPD), the antagonist Länder were controlled by the conservative block (CDU/CSU) and headed by Bavaria (Herrlitz, 2013). However, in the 80s a development were to take place which would decide the orientation and course of ILSA. It was the work on the INES project²⁰. Saha and Tuijnman (2003, p. 1127) describes how the CERI governing board in the late 1980s failed to reach a unanimous decision on the project and several nations expressed reservations. The crisis was only settled after strong political and economic pressure by the United States. Tröhler (2013, pp. 153-155) refers the initiative of both establishing and collecting descriptive statistics about educational systems in OECD, as due to pressure by the United States. The background of which is supposed to have been the then recently published report *A Nation at Risk* which combined the crises in the economy with the failures of the education system in USA. Papadopoulos (1994, p. 181) underlines the conflict between the member states positioned in two camps *'[...] on the one side, the neoconservatives, dominated by Reaganite and Thatcherite views of society, and on the other, the "egalitarians" , largely represented by the Scandinavian model of society '*. This conflict could be said to mark the ideological shift towards education indicators as primarily human capital devices which became an embedded view of the OECD ministerial meeting in Paris 1984 (cf. Heyneman, 1993, pp. 374-378). From a decade of interest in merely managing quantitative aspects in education – number of students, school facilities, numbers of teachers – the work was now directed towards an interest in cost-effectiveness and qualitative aspects (Husén and Tuijnman, 1994, p. 11). The background was the public sectors adoption of private corporate management's terminology (NPM) which coincided with the takeover of conservative, right-wing governments in many member countries, with the outcome of a higher pressure on public sectors and reinforced market practices, first implemented in the 60s theories of economic education (e.g. Kellaghan, pp. 144-145; Husen & Tuijnman, 1994, pp. 11-12; Eide, 1990, pp. 47-48).

3.4 The introduction of ILSA in Germany, the rise of EB

For an observer of reforms in the German education system, there has been a remarkable expansion and alteration in the last two decades²¹ (e.g. Klieme, Avenarius, Blum, Döbrich, Grubel, Prenzel, Reiss, Riquartz, Ross, Tenorth und Vollmer, 2007, pp. 11-14). The TIMSS 1995 report marked the beginning of a changed view of German education but its results did not provoke any public debate as opposed to the announcement of the PISA 2000 outcome. It caused a massive response, as witnessed in the media through chronicles, panel discussions and readers' reviews and the entire German education system was questioned (Spiegel, 2001; DIPF, 2002). The critique can be divided in two parts or levels.

²⁰ See 3.2.

²¹ The background of which is partly laid out in the study's first part. A more complete version would take a PhD thesis.

One part concerned the effectiveness and efficiency of the knowledge transfer in German education, for example with the stress on the input, describing what ought to be learned instead of what must be learned, the output. It was stated that the German education system was out of date and too diversified, or to less diversified and not adapted to the individual's ability or talent. The list can be made much longer but a second and more fundamental dimension that included many of the topics in the debates, were the status, standing and content of *Erziehungswissenschaft* in education, the traditional education research field, and its definition or redefinition. *Erziehung* can be translated as both education and pedagogy but include also the fostering of values such as responsibility, awareness and self-knowledge. In that sense it is closely linked to the *Humboldtian*²² concept of *Bildung* but with emphasize on a transferred normative behavior (e.g. Josephson, Karlsohn and Östling, 2014, pp. 1-3.). The changed view on education resulted in practices of output steering, introduction of tests and evaluations, quality monitoring by establishment of standards; in sum the introduction of a performance culture (e.g. Hartong, 2011, pp. 61-70; Klieme et al., 2007, pp. 9-12). Furthermore, a part of educational research, (empirical) comparative education, *Empirisches Bildungsforschung (EB)*, began to establish itself as the claimant of educational facts on which the overall performance monitoring should be built (cf. Zlatkin-Troitschanskaia and Gräsel, 2011, pp. 9-11). The demand on research to solve the deficits in education affected *Erziehungswissenschaft* in particular since it was a diverse multidiscipline which included empirical qualitative studies to discourse analysis, with emphasis on methodology and theory, the results of which could not, and did not, claim general validity (e.g. Hornbostel and Keiner, 2002, pp. 634-638).

The results of the TIMSS 1995 survey and the PISA 2000 study channeled a strong pressure on research to be, or become, productive in terms of producing knowhow that could be of practical use. In this temporary vacuum, EB entered and positioned itself as the modernizer of *Erziehungswissenschaft* and it was the start of an empirical analytical drive in the research of *Erziehungswissenschaft* (e.g. Weishaupt, 2008, p. 5; Zlatkin-Troitschanskaia and Gräsel, 2011, pp. 9-11). Instead of describing, the emphasis shifted towards explaining. It was a shift that had been under way, characterized by influences from natural science, that now became reinforced (e.g. Raithel, Dollinger and Hörmann, 2009, pp. 180-182). *Erziehungswissenschaft* gradually divided into an empirical and a theoretical part which underlined the decline already present (table 1). Even though the gradually decreasing number of professors was relatively moderate, the declining trend becomes more obvious when the co-workers (W.Mitarbeiter) is divided into a central group of assistants and docents (Ass/Dozent), and other minor co-workers (Böhm-Kasper, Weishaupt, 2008, pp. 20-21).

Year	Erziehungswissenschaft (EW)	Psychologie	EW. Co-workers divided	
	Professors/W Mitarbeiter	Professors/W. Mitarbeiter	Ass/Dozent	W. Mitarbeiter
1998	966 / 1924	517 / 1746	330	1594
2000	979 / 1874	524 / 1751	277	1597
2005	861 / 1972	579 / 2000	190	1782
2006	843 / 1957	570 / 2044	158	1799
2007	875 / 2133	541 / 2133	121	2012
2008	905 / 2350	535 / 2370	86	2264

Table 1. The development of *Erziehungswissenschaft*. Source: Statistisches Bundesamt, Fachserie 11, Reihe 4.4, verschiedene Jahrgänge; in Muders and Weishaupt, 2012 pp. 158-160. (Modified by the author.)

²² It refers to Wilhelm von Humboldt, the Prussian government functionary and diplomat who, according to earlier research, reformed higher education in terms of strengthening the role of professionals as also education with special emphasize on the individual's development and autonomy in a social and societal context. However, this view is contested and of late birth according to recent studies (cf. Josephson, Karlsohn and Östling, 2014, pp. 11-14.).

Muders and Weishaupt (2012, p. 166) underline the difficulties of recruiting doctoral scholars for Erziehungswissenschaft depending on the decreased number of qualified posts, such as assistants and docents scholars. Whereas Erziehungswissenschaft witnessed an increasingly difficult – and questioned – position, EB gained ground. Aljets (2014, pp. 29-32) has gathered data that show that in 2005 the number of professorship in EB was 27 and in 2012 75, that is, from about 3.5 % of all professorship in Erziehungswissenschaft to about 8.5 %. Hence, an ever increasing part of professorships in Erziehungswissenschaft became oriented towards EB. The trend is also visual in the funding of research where for example BMBF and the German Science Community (DFG), an important third-part financier of science, have supported and implemented special programs, development projects, for EB (cf. Zugenrucker and Weishaupt, 2008, p. 32; Zlatkin-Troitschanskaia and Gräsel, 2011, pp. 9-11). The impact on and increased focus of empiricism in German research even got an epithet, *Die Empirische Wende* (The Empirical Change). The proponents of it argued for a more utilitarian approach of research to enhance German education quality and competitive chances (Klieme, 2006, pp. 765-768; Buchhaas-Birkholz, 2009, pp. 27-28), though the opponents underlined the *perpetum mobile* of such research in which conclusions/results lead to questions destined for empirical (measurable) research; the steering of research into EB (e.g. Bellmann, 2006, pp 487-488; Dietrich, Heinrich and Thieme, 2011, pp. 9-11).

4. Methodology and theoretical lens

The frame narrative of the study is about a quantitative research field QCE that has been able to develop, establish and position itself, as well as influence education policy and education practice. The outcome has been the institutionalization of ILSA by an epistemology based on a normative supremacy with origin in economic management (concepts and values), experimental psychology (psychometric) and mathematics (mathematical analysis, linear algebra, stochastic analysis, differential equations and measure-theoretic probability theory). The consequence has been the establishment of value laden but vague concepts in education as quality, standard and assessment. The study uses a German state as a case study to show how these emotive concepts empirically form a relational web of pattern in which actors' positions, positional changes, as also forms of capital, are affected. All these aspects are part of the French sociologist Pierre Bourdieu's field working theory of reflexive sociology, chosen as the theoretical lens in the study.

4.1 Choice of actors and demarcations

The subjects of the study are actors on a local level, Bavaria, that were involved in the implementation, the practice and the monitoring process of the three concepts: quality, standards and assessments. One actor, IQB, is situated on a national level but is included partly due to the central position it can be said to hold in this practice, but also the significance of this actor as a representative of the structural steering from a national level to the local level (associated with IZB). Still, there are other actors involved in the process but the focus of this study was the local level. Another decision was the demarcation in time after 2000, the year the first PISA study was presented, though some of the presented data is in time situated before 2000 in order to exploit the background and give a more detailed interpretation of events decisive for the understanding of the study's objective.

4.2 Pierre Bourdieu's field working theory

Pierre Bourdieu's field working theory²³ in sociology entails a bridge between structural (positivist) and phenomenological (post-positivist) interpretations of the world, between vision of objects and subjects. Bourdieu envisioned the social world as accumulated history, hence, a narrative cannot be *mechanistic*, cannot be reduced to a theory where people are exchangeable (Bourdieu, 1997, pp. 217-218). It is based on the perception that society consists of a struggle between and within different fields where individuals are predestined by habitus towards their particular field. Though it does not mean that the individual will make this transition, she may not succeed, but she is always influenced in choice and possibilities. The power of the fields are the amount of (historically) accumulated distinctive capital each field hold, and capital is a central term and explanative factor of the individuals habitus and by that, the chance to succeed. Bourdieu separates four forms of capital: economic (dominant), cultural, symbolic and social (Bourdieu, 1986, pp. 241-258; Bourdieu, 1990a, p. 55). The working theory is relational, the fields stand in a set of relational correlations towards each other but these are not fixed or permanent, as the relations shift in the constant struggle for dominance or status quo, the fields' amount and predominant type of power change (Bourdieu and Wacquant, 1992, pp. 224-235). The struggle is also a part of the dominant's efforts to inscribe a just legitimization of their dominance – in the eyes of the dominator and dominated – through the transferring of a subjective value system preordained to uphold the vision of being objective (Bourdieu, 1996, pp. 264-265). The relational analysis is at the center of this model (Bourdieu and Wacquant, 1992, pp. 224-235). Thus, there is a hierarchy within each field as between fields (dominant and dominated subfields) and the corresponding levels between each field share intrinsic values. It is often referred to as Bourdieu's field theory but Bourdieu never labeled it²⁴, more described it as a useful working model of reflexive sociology (cf. Bourdieu, 1990b, pp. 12-21).

Researchers have often selected specific parts, or terms, of field theory to be the main principle. Hilgers and Mangez (2015, pp. 3-4) describe for example, that an essential part of field theory is not so much the properties of an object ` [...] as the network of (cor)relations that is woven between them and other neighboring formations that is the focus of analysis`, it is not the individual object but the system of relations that are central to field theory. This is a slight different interpretation then given by Bourdieu in his empirical field studies (e.g. Bourdieu, 1996, pp. 10-19) where relations are part of individual attributions, an asset or limitation, be it purely individual or individual schools, groups, etc. That is, there is no single basic principle to grasp the working theory by Bourdieu, conceptual parts cannot be understood or interpreted explicitly, they are always conditioned. `Such notions as habitus, field and capital can be defined but only within the theoretical system they constitute, not in isolation`, relations are equally conditioned and acquire meaning only within a system of relations (Bourdieu and Wacquant, 1992, p. 96). The stress on the fallibility of science dependent on either theoretically prepositioned and preconceptualized concepts, or empirical postpositioned and postconceptualized, derives from Bourdieu's conviction on empirical based understanding – research based on statistic, interviews etc. – decoded by a theoretical broad reference frame which combine the object and subject. In short neither perspective can in isolation unravel what a combined perspective can, the complex reality where the obvious and the concealed are homologous (Bourdieu and Wacquant, 1992, p. 96, fn. 48). Both empiricism and theory is conditional for the model (or approach). Bourdieu opposed the split or antinomy of methodologies by opposing epistemologies in sociological research which according to the working method of field theory, derives from the different scientific fields' guardian of territory, the (old and new) dominants of dominated subfields in academic research (Bourdieu, 1984a, foreword).

²³ The term working theory derives from the outspoken ambivalence of Bourdieu to state his working method as a distinct theory.

²⁴ Underlined by his resentment for *Grand Theories*.

4.3 Critique of Pierre Bourdieu's work

Bourdieu's field working theory has been influential in sociology and implemented in several studies up to date, though a particular criticism of its diffuse borders – where to draw the line – is mirrored in the multiply ways it has been adopted (e.g. Lee, 2013; Couldry, 2003), partly due to an interconnected desire of the practitioner to identify some of Bourdieu's concepts as the basis for the entire working theory (e.g. Cheal, 2005, pp. 154-162). Another critique has targeted the language in Bourdieu's work as inaccessible and abundant which makes the reading and understanding difficult, although this is a fact, Bourdieu recurrently argued for the necessity of a specific and exact language, devoid of *common language*, to describe the complex reality. The consequence has been a plethora of different understandings and misunderstandings. Bourdieu's work has also been targeted for being specific French context based and not replicable, and his favored motive of excluding dichotomies, as social value constructions, is not always implemented in his writings (Swartz, 1997, pp. 52-64). One major point is that the field working theory includes the use of both qualitative and quantitative data – to acquire and obtain as much as time and skill allows – in order to explain the relational pattern of the object. The scientific high-valued criteria of rigor can therefore be hard to claim, implicitly underlined in Bourdieu's own studies of phenomena in society (as religion, education, science) where the empirical theoretical findings are brought forward in an excessive language which can give a vague and indistinct impression. However, Bourdieu's complex use of specific language and expressions, refers to the imperative need to understand each property and its foundation, based on a precise interpretation of a vast amount of collected data²⁵ and deter the reader from making own interpretations (e.g. Bourdieu, 1984, pp. 1-4). The most problematic aspect in using the field working model on the study's aim is that Bourdieu's work was concentrated on individual and collective 'utility', the re- or devaluing of capital which can be seen as in accordance with human capital and knowledge economic theories. However, where these portray a singular dimension in which the value of the individual can only be monetarily estimated, Bourdieu's concept of capital is multiple – economic, cultural, social, symbolic – and not only a measure of monetary value but class (or category) dependent, regarding both the meanings and applications of the different forms of capital (Swartz, 1997, pp. 75-82; Weininger, 2005, pp. 92-95).

4.4 Method

The method in the study was foremost qualitative, based on interviews, but official descriptive statistics was an important source of guidance – acting as a marker of significant events, for example the positional changes between and within fields. In that way the study could be said to follow the methodology of Bourdieu by combining qualitative and (to a degree) quantitative sources, even if an *orthodox* pupil no doubt would have collected data for a component or multiple component analysis.

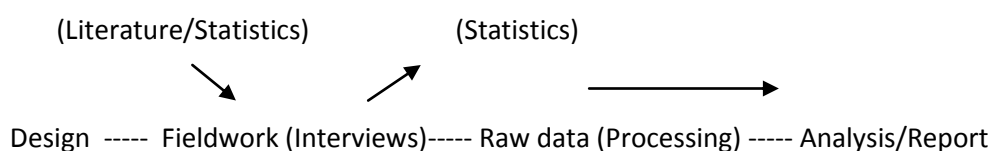


Figure 1. Interplay of qualitative and quantitative data in the study. (Miles and Huberman, 1994, p. 41). (Modified by the author.)

The model attempts to explain the role of descriptive data (statistics) as an integrated part of the study until the analysis, where it played a minor but complementary role. The main body of statistics used in the study comes from BMBF²⁶, Statistisches Bundesamt²⁷ and IQB, working on the German

²⁵ The collection of data in Bourdieu's work often involved sizeable project groups of colleagues, students, administration and external agencies.

²⁶ The Federal Ministry of Education and Research.

federal level, respectively the Bavarian Education Ministry and the State Office for Statistics from the state level. The exceptions (and minor part) are statistics done by individual researcher or group of. The former statistics were more comprehensive, containing a variety of aspects while the latter was concentrated on a single or a few dimensions. One more important difference is that the latter was sometime funded by actors with special interest in the outcome as the study by Klaus Klemm (2015), funded by the German Telecom Foundation.

However, the major part of the data comes from the interviews, conducted at or close to the informant's workplace. Official representatives from the contacted institutes were always sought but when this was not possible the network, built up during the preparations for the study, formed a second opportunity. The dialogue with the representatives were more formal than with the network but since the meetings were of some length – between 35 to 90 minutes – it meant that personal opinions were always given a certain space. One contributing factor was probably also that the informants were quite relaxed in the meetings, more curious and interested than reserved. They were all contacted by mail which included the same explanation of the study, its background, focus and the areas of specific interest²⁸. The total amount of informants were fourteen: one former chef of the education ministry, three teacher unions representatives from two different unions, one researcher from ZIB, one researcher from TUM, two `experts` of PISA – that is, one reputable researcher from LMU²⁹ and one expert in informatics³⁰, one researcher from the ISB/QA, one researcher at the IQB and three teachers and one principal from the secondary school form Gymnasium. Each interview began with an introduction and background of the study, after which the informant was assured anonymity. Only at one occasion was a promised interview cancelled and it concerned a statistician for one of the teacher unions.

The questionnaires that were used in the interviews were semi structured although the emphasis did differ depending on the informants – previous or current – position. All interviews were recorded and processed in three steps: first listened through, secondly transcribed (from German to English) and third, the transcriptions were used to analyze the data by focusing – selecting and reducing – that is, to simplify and mark the statements that were valid for the study's objective which is a precarious operation for an inexperienced researcher (cf. Miles and Huberman, 1994, pp. 9-12). There is always the subjective influence by the researcher, but the guidance of the research questions and the use of a reflexive sociology – to question the researcher's interpretations and influence – have (hopefully) maintained the study's equilibrium. Another critical part of an explanative qualitative study is its description of claimed causality, events linked into a timeframe meant to explain the outcome. At worst the relationship is invented, at best interpretable. The frame narrative of this study is a try to as clear as possible, let the reader understand the researcher's interpretation of causality – the underlying assumptions about the chain of events portrayed in the study, how certain events were linked.

The crucial part was the analysis in which the collected data were organized and systematized in order to be interpreted, that is, to give the findings meaning and content. In other words, the raw data were reduced and framed by the premise of the study to a higher abstract level (cf. Friedrich, 2008; Robson, 2005, pp. 269-277). In this part the data was carefully compared to the concept and parts of the working theory by Bourdieu. In the analysis some features of the data was rather obvious whereas the relevance of other phenomena took time and a considerable amount of work to disclose.

²⁷ The Federal Statistical Office.

²⁸ See appendix 2.

²⁹ Ludwig Maximilian University.

³⁰ A prominent figure in the early PISA debate in Germany.

4.5 Validity and reliability

The study is empirical and context based, it includes both quantitative sources, statistics, as well as qualitative, interviews. Hence, multiple methods are used which make data and methodological triangulation of the study possible (cf. Torrance, 2012, pp. 113-114). Multiple methods can be a tool to extricate and condense a mass of divergent data, together with theory, into a logic entirety, a way to present a reality. However, it can also become the misrepresentation of if not rigorous conducted, instead of an entirety it becomes diversity (e.g. Bryman, 2007, pp. 20-22). The data collection by these methods is presented so as to ensure the validity of instruments. The inferences by the data are developed step by step and as completely presented as time and space allowed, making a scholarly judgment possible, being it representative or misrepresentative. Validity also concerns the selection of informants; that the sample corresponds to the objective of the study. The informants in this study can be said to consist of two blocks: one chosen first-hand by the researcher and one made possible by the built up network on location. The validity of the latter can be questioned but in a case study on *foreign soil*, this is a stipulated condition. It has the advantage of creating a preset relationship of trust and make room for a less formal discourse.

In sociology, the reliability of research depends on a high degree on the instruments (methods) used, that is, that they are used – and well suited – to extricate the objective of study. The researcher has to explicate and motivate the use of methods: that they are in distinct accordance with the aim of the study. However, a degree of subjectivity will always be present, and at worst predestine the outcome, therefore the researcher must not only claim but be able to inscribe the analytical practice of the methods (e.g. Brym and Lie, 2007, chp. 2). In the present case, the recorded interview data is the most valuable source to ensure the consistency with the outcome of the analysis.

4.6 Ethic consideration

The study was conducted in accordance with the Swedish Research Council's four basic principles of research ethics: (1) the requirement of full participant information regarding the study's objectives, the participants' role and the right to discontinue, (2) the requirement of consent, (3) the requirement of confidentiality and (4) the requirement of research data available only for research (Forskningssetiska Rådet, 1990, pp. 5-14). One important dimension of the study was that it was carried out in another state which meant that the researcher was dependent on the benevolent help from researchers who were well acquainted with the German and specific Bavarian context. The research ethics were therefore always in the foreground and worked as a set of rules of conduct.

5. Results and findings

The presented actors may give an impression that they all were official representatives of institutions and groups. This is true in sense of some – BLLV, ISB/QA, ZIB, IQB – where an official request meant that the researcher was asked to get in touch with, or was approached by, a selected person. There are, however, individuals included in the study belonging to professions important for the study but not acting as official representatives of these. They are part of the network of informants, former described, which arose in connection with the preparation of the study. In the presentation informants belonging to the same or similar, category have been merged. The idea was to facilitate a comparison. For the reader's convenience, the section begins with a chronology of events and dates, important for the understanding of the German educational context.

1995: German takes part in IEAs survey the Third International Mathematics and Science Study (TIMSS). It is the first time since the 60s (IEAs FIMS³¹ survey) that Germany is included in an International Large-Scale Assessment. Development of educational standards in the subjects mathematics, German and first foreign language.

1997, June: KMK presents and discuss the, in some parts, bad results of the TIMSS study and the OECD study Education at a Glance. The *Konstanzer Beschluss* (Konstanz Decision) in which KMK declares to develop measurements aimed at securing the quality of education and also to engage in international and national comparisons of educational achievements. The participants of KMK, including union representatives, declare that an intense quality debate is vital (KMK, 1997b).

1997, October: KMK decides to develop and initiate standards for educational certificates on a Länder equal basis to secure the quality on external and internal evaluations, grounded on an output orientation (KMK, 2015a)

2000: Germany takes part in PISA-I and the enlarged PISA-E for comparison of the Länder.

2001/2002: The PISA test results are published and the so called *PISA-Shock* follows depending on the bad test results. At the same time KMK announce a *Handlungskataloge* (Catalogue of Action) containing seven priority areas of which number five reads '*Securing Quality through obligatory Standards and Evaluations*' (Tillman, Dederig, Kneuper, Kuhlmann and Nessel, 2008, pp. 90-91).

2003: Establishment of the Bavarian agency for quality within ISB, the acronym of ISB thereby changes interpretation into State Institute for School Quality and Educational Research.

2004, December: The decision of the implementation of equal educational standards in all Bundesländer by KMK. An openly declared paradigm shift takes place oriented towards output, reporting and system monitoring (Bildungsstandards der Kultusministerkonferenz, 2004).

2004, December: The Institute for Educational Quality Improvement (IQB) is created and funded by KMK and the BMBF. IQB will be responsible for the implementation of the *Bildungsstandards* and the supervision of the German school performance.

2006-07: The Federal Reform which mark a split³² between the KMK and the Federal authorities. KMK vote to become less dependent on the Federal authorities in matters of education. However, the latter manage to bring in the article '*Feststellung der Leistungsfähigkeit des Bildungswesen im Internationalen Vergleich*' which refer to the judgment on the level of skill/knowledge in the German educational system (or systems) in international comparison. This key area becomes the shared responsibility of both KMK and BMBF (KMK, 2007; BMBF, 2007).

2010: The Centre for International Student Assessment (ZIB) is founded by BMBF and KMK. It consists of the School of Education of the Technical University Munich (TUM), the Leibniz Institute for

³¹ First International Mathematics Study.

³² An historic ongoing struggle between KMK and Federal authorities.

Educational Research and Educational Information (DIPF, Frankfurt) and the Leibniz Institute for Science and Mathematics Education (IPN, Kiel). The Research Data Centre (FDZ), which is led by IQB is also attached. The research area of ZIB is International Large-Scale Assessments (ILSA) of which the largest is PISA and the research is carried out by the [DIPE](#), the [IPN](#) and the [TUM](#) (ZIB, 2012a).

1995	1997-June	1997-October	2000	2001/2002
KMK-decision to take part in TIMSS. Limited Standards work.	<i>Konstanzer Beschluss</i> -KMK decides to take part in ILASs	Presentation of TIMSS. KMK decide to work on and enlarge Standards	Germany takes part in PISA-I and PISA-E	Presentation of PISA results. <i>Pisa-Schock</i> . Catalogue of Action.
2003	2004-December	2004-December	2006/07	2010
Agency for Quality within ISB. Changed interpretation.	Implementation of equal educational standards in all Bundesländer by KMK	IQB is founded by KMK and BMBF, to monitor the quality of education in the Bundesländer	The Federal Reform. KMK and BMBF commit to joint work in the area ILSAs	The ZIB is founded. Cooperates with IQB.

Table 2. Overview of significant events in the German educational context.

5.1 The Bavarian State Ministry for Bildung, Kultus, Science and the Arts³³

Bavaria is the largest and oldest state in Germany. With over 1000 years of history, Bavaria's cultural heritage is very much alive today and continues to define the region and its inhabitants. The face of Bavaria is unmistakable and at the same time inextricably bound in German and European culture and history. Each state in Germany is responsible for forming and implementing its own educational policy. Annual comparative studies regularly confirm that Bavaria boasts one of the top performing educational systems in Germany (Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, no date/a).

The official presentation of Bavaria and its education systems reflects a self-image of a state with long standing traditions and an excellent education system, an image that is most common mirrored in the views of Germans from other Länder. The State Ministry for Education has undergone numerous organizational and name changes and was in time of the first presented PISA studies called Ministry of Education and Kultus, responsible for and in charge of the implementation of PISA in Bayern.

5.1.1 Higher civil servant in the former Bavarian State Ministry of Education (D1)

The informant (D1) held office in the former Bavarian State Ministry of Education and Kultus for a long period, and was a senior civil servant with a high position and life history in education, from Gymnasium over teacher profession unto early attachment to the ministry. The informant studied at a Gymnasium at a time when only a small fraction, about 1/10, of the students entered that school form depending on achievements and qualifying notes. It was an elite institution in which students' amount of (economic and symbolic) capital, made a difference, that is, those with the proper habitus, *'[...] the systems of disposition to a certain practice'* (Bourdieu and Wacquant, 1992, p. 77). The

³³ The ministry responsible for school administration and school inspection.

individual's that know the game³⁴, had the best possibilities. Being a part of an exclusive minority in education indicate to have, or have assimilated, the habitus of the field, the inferred notions of that specific field's collective ideas and associations of what it means to belong to it. The informant's educational trail that provided a successful bureaucratic carrier with the outcome of a high post in administration, has in itself, by the educational field's notions of valued cultural capital, incorporated in the individual's sense a high esteem of the practiced *tools* used in the field to differentiate its collective. These are in the educational field tests, assessments and cherished competitive assets – certificates, degrees, diplomas – imposed on and by the members of the field. The bureaucratic field is close bound to and dependent on the educational field in ways of reproduction. It is the education system in form of the primary, and to a degree secondary, school that mediate the reproduction of a state and in so, the bureaucratic field with its control marked in regulations and prescriptions (Bourdieu, 2005, pp. 111-118). In a sense, bureaucratic administration is founded on principles brought in by education and become the refinement and justification of it.

5.1.2 The beginning

In 1995 Germany took part in two sections (excluding the part for primary school) of the IEAs TIMSS survey, the first major ILSA in Germany since the 60s. However, the decision to take part is partly clouded and difficult to find in spite of being such a subversive event. That it was a big surprise was underlined by D1 *'In 1995 Germany decided to take part in TIMSS study and that was a big astonishment [...] a surprise for the whole teacher body and most of the education ministry'*. There was no real anticipation or preparation in the Bavarian state administration that Germany was just about to take part in an ILSA, no indication of what was to come. TIMSS 1995 was conducted by the *Max-Planck-Institut für Bildungsforschung* (MPIB), the *Institut für die Pädagogik der Naturwissenschaften* (IPN) and the Humboldt University in Berlin. The survey showed that German students were not among the best achievers in mathematics and physics. *'(D1) The result for Germany was bad, the first shock [and it was] the reason behind the Konstanzer Decision [by KMK] to compare, but foremost to compare nationally, between the Länder [...]'*. The aim was to secure system relevant steering knowledge by comparisons of the German Länder (KMK, 1997a; Tillmann et al, 2008, pp. 16-19), it had until then not been politically possible. The Konstanzer Decision was a landmark in that respect, enforced partly by the political pressure caused by the bad outcome. The two main political blocks in Germany by that time was *Sozialistische Partei Deutschland* (SPD) with its allies, and *Christlich Demokratische Union Deutschland* (CDU) with its Bavarian ally CSU. These two fractions in KMK often formed two opposite blocks. KMKs history is somewhat marked by a recurrent struggle between these blocks, and the Konstanzer Decision is therefore quite unique with the SPD block finally agreeing (after an intense and intern struggle according to D1) to the testing and comparison of the Länder. However, the aim to take part in ILSAs did not include the PISA study since it was not set. The KMK established a chief office commission of *Quality assurance in schools* to work out a frame of collective preconditions of the Länder's educational systems on the basis of the data from TIMSS 1995 and OECD³⁵ (KMK, 1997b). In the chair of the commission was a delegate from Hamburg, a representative of the SPD block. D1, who also at that time a member of the commission, portrayed the delegate as a fierce proponent for empirical research. Then, in 1998, OECD made an offer; Germany was invited to take part in the PISA 2000 study.

(D1) [...] and then suddenly OECD came with an offer to include (Germany) in the PISA-study. We [KMK and Federal authorities] decided to take part in the PISA study but we wanted an extended version, in order to fulfill the obligations of the Konstanzer Decision, and PISA-E could accomplish just that.

³⁴ The metaphor of *'game'* in a field should be understood as following regularities but not explicit and codified rules. It underlines and accentuates the necessity of habitus for the feel of the game (cf. Bourdieu and Wacquant, 1992, p. 98).

³⁵ Education at a Glance.

There were no real interest in the international part (PISA-I), it was the Länder comparison (PISA-E) that made the KMK decide to take part in the PISA study. Prof. Jürgen Baumert became head of the study, by then a director of the MPI, one of the early centers of EB in Germany, established 1963, and one of three institutes that had been involved in TIMSS 1995.

In the field theory the domain of a field embodies the principles by which certain ideals, ideas and values are incorporated and expressed by the collective, the *doxa* of a field. The boundaries of a field are marked by the transition into other or changed sets of principles, and different fields are standing opposite each others in a homologues hierarchy of dominated and dominant. Every field is an arena where the individuals compete to position themselves. The *rules* of the game are always the underlying principles which weight can shift depending on the temporarily outcome of the continues struggle among the collective and the status of the field's believed value (cf. Bourdieu, 1990a, pp. 88-89). The result of TIMSS 1995 meant that the up-to-then dominant research field in education, *Erziehungswissenschaft*, became questioned by the political field, by other research fields as also from within the field itself. It was a start of transformation of positions in which EB, by then a subfield of *Erziehungswissenschaft*, proclaimed itself³⁶ as the *interpreter* of objective (educational) truth. The outer signs of this were for example the chosen chair of the commission for *Quality assurance in schools* (an intransigent herald³⁷ of EB), the head of the PISA study and the employed institutes.

The motives behind Bavaria's and Germany's involvement in the PISA study were partly a desire to compare its education system, to know its standing as a highly competitive education system, in other words to confirm its superiority.

(D1) A lasting abstinence after about 25 years without international comparisons, a wish to compare [...] It was a general view that education in Germany were of high international standards.

Even so, on account of the TIMSS 1995 results³⁸, Bavaria was one of the first Länder that introduced new tests and measurements in different school subjects, formally to secure the quality but the act also held a legitimizing aspect: legitimizing the measurements tools and the education politic in the state. *(D1) In Bavaria we acted on the consequences of TIMSS. In 1998, that is before PISA, we introduced performance tests and comparisons, and they were extended.* The undertaking of tests and assessments were though not decentralized to regions (in Bavaria) or individual schools. Following the Bavarian tradition of a strict and centrally controlled education system, they were state ordered and state managed.

5.1.3 PISA 2000

In December 2001 the results of PISA-I were presented in the Medias and were met with astonishment and surprise. Even if the number of articles by the press were substantial, the peak was reached in June 2002 when the result of the Länder comparison study, PISA-E, was presented (Tillman et al., 2008, p. 72). Beside newspaper articles the result was published, discussed and debated in an endless number of the Länder's official documents, meeting papers and so forth.

(D1) In Germany it was a media circus, partly not understandable. In Germany 480 articles were written in the first week, in Italy 27! [...] The attention was incredible, Bavaria came out as the best, close to and comparative with many nations results [in PISA-I] we [in the ministry] was called upon the evening before and nothing were allowed to be disclosed [...] and it [Bavaria's result] was not such a great surprise.

³⁶ Elevated by the comparative educational methods used in the TIMSS 1995 study and Education at a Glance.

³⁷ As in the originally sense, a messenger of proclamation from the dominator, the ruling or noblemen, or in this example the hegemonic value construct.

³⁸ The sampling came from 15 of the 16 Länder and the outlay of the results were not Länderspecific.

The image of Bavaria's education system seemed to be confirmed but there were problem areas that the PISA-E supplementary questionnaire showed: the high number of school leavers without certificates, gender specific related results and socioeconomic aspects. *(D1) The result showed gender specific, ethnical and social inequalities, quite a substantial at-risk-group.* These results were met with surprise, that is, the significance of them in Bavaria. In Germany as a whole, the disparities in educational possibilities for children with different socioeconomic background were already documented.

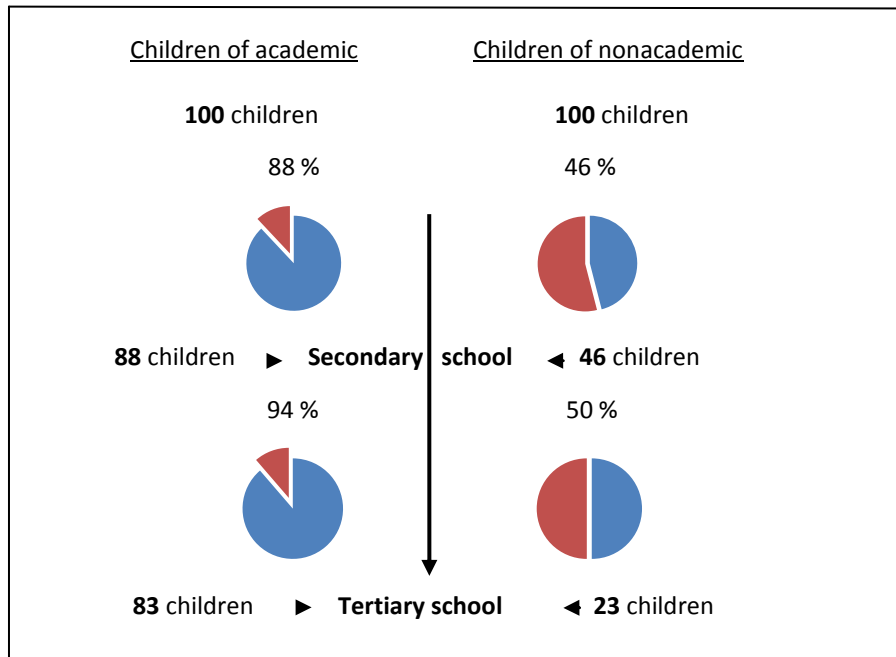


Diagram 1. Percentage of children with academic and nonacademic educated father, at secondary and tertiary school form in Germany 2005. (From: Isserstedt, Middendorff, Fabian and Wolter, 2007, p. 111.)

However, the result now stressed the severity of these issues in Bavaria and as a consequence the extent of the inequalities in the education system became official and public knowledge (e.g. Stanat, Artelt, Baumert, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Schümer, Tillmann and Weiss, 2002, pp. 19-22). Rotte and Rotte (2007, pp. 297-298, 301) showed that Bavaria was the German state in which access to higher education had the highest correlation with the student's family's socio economic background. With the PISA study the interest for education in general was elevated and the schools, and their test results became the focus of many educational actors and parents.

5.1.4 Standards and epistemological implications

Standards were to some extent already established in German secondary school in the mid 90s, but the boost came with the outcome of TIMSS 1995 and PISA 2000. The development of common standards for all Länder was a *de facto* accomplishment which depended highly on the pressure towards a more standardized and comparable education system in Germany.

(D1) We developed standards, something we did not have, but since 2004 we have them in German, Mathematics, Science, Biology, foreign language and they describe what competence a [...] student must have as also [the performance levels] of different school certificates, which all Länder have made an obligation to follow [...].

The circumstances demanded action and a set of standards were introduced, and the applied tests were worked out on so called *Regelstandard*, based on the average level, such as the average competence level of the pupil in 4th grade ought to reach (e.g. KMK, 2004, p. 9). However, KMKs

decision to use Regelstandards did not follow the advice of the expertise involved in the preparatory work who had recommended standards based on minimal levels (Klieme et al., 2007, p. 27).

Standards and tests resulted in an increased awareness of the hierarchical education system but instead of questioning the selectiveness, parents now started to demand access to *best practice*. The Bavarian reform of secondary school was a consequence thereof. Hauptschule had always been the least prestigious school form, seen as inferior to *Realschule* and especially *Gymnasium*³⁹. In table 3 the starting point is around the year 2000, even if the increase of the number of students at Gymnasium and Realschule had been a trend before, it now became manifest.

Year	Gymnasium	Realschule	Mittel/Hauptschule	Wirtschaftsschule	Förderzentren
1990	26,8	27,6	38,1	3,2	3,6
1995	26,4	27,2	38,3	3,5	3,9
2000	26,5	27,3	37,8	3,3	4,4
2005	29,7	26,2	35,9	3,4	4,1
2010	32,3	29,9	30,1	3,3	3,6
2013	32,3	33,0	27,2	3,1	3,6

Table 3. Percentages of students in 8th grade in Bavaria divided on secondary educational paths 1990-2013. (Source: Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst. 2014a. Reihe A, Bildungsstatistik Heft 59. Tab I 2. Relative Schulbesuch in der Jahrgangstufe 8 in Bayern.) (Modified by the author.)

Education moved towards a marketized arena where actors (pro)claimed the equal *privilege* to compete by access to best practice; the inflicted epistemology set the *modus operandi*. At the same time, the cultural reproduction transmission was confronted – in terms of the link between knowledge, power, socialization and education – and even if the reform of Hauptschule into Mittelschulen can be seen as part of a genuine organizational reply on a new demand, it can also be viewed as a way to secure the existent reproduction. Hence, the increase of students at Realschule 2010-2013 *vis-à-vis* the standstill at Gymnasium is one sign of adjustment where a demand on reforms could be interpreted as averted into a school form less attended by the categories with most capital. The public demand on equal terms was thus channeled and the societal structure sheltered, with the outer signs of democracy (cf. Swartz, 2000, p. 209).

ILSA (re)introduction in German context did not just create a consecration of the face value and orientation towards practical solutions (standards, assessments etc.), it also meant a theoretical commitment in terms of acceptability, translation and familiarity with the imposed discourses.

(D1) Germany had the last 20 years [only] taken part in minor international studies noticed by few, they were regional and nobody believed in empirical research [...] the empiric first of all had to set itself. That education could be measured was unbelievable for the Professors [...] the Germans had ignored the American educational research for 20 years. In America they had different kind of tests and the results were followed up and then they had these anchor items since the 70s in USA. [In that way] they could follow the development over time.

One mayor cause of dissensions was that German education had, more or less, been oriented towards the classical Bildungsconcept where education was seen as a fostering of individual and society values where one part was the development of the innate skill, the profound change now was the development of educational basics in terms of prescribed topics of the ILSA.

³⁹ Gymnasium could be translated into grammar school and Realschule to junior high school.

(D1) For example, PISA does not test German [language] according to the traditional Humboldt view [...] instead the students are examined if they can read functionally, examines literacy not literary education. You must remember that PISA measures a segment and that this segment is Anglo-Saxon embedded, utilitarian, the competence of what the students must have in the working life. It's good but that is not the only thing that is conveyed in education d in education.

ILSA did not just come as a limited test, it was a breach of the old in favor of an intensified individualistic economic concept marked by a new technocratic language, alien even for the administration in the state.

(D1) The specific language of researchers in the psychometric field contains and includes quite different terms, concepts and meanings than in traditional pedagogy and administration. We had to learn it all anew, it was not easy, a total different scientific language and a scientific knowledge that we had to make use of in administration practice. That was the most difficult part [...].

5.2 Trade Union for Education and Science (GEW) and the Bavarian Teacher Union (BLLV)

Two unions in education that can be said homologues, not just of member groups included but also on the basis of their structural organisation and background, are GEW and BLLV. However, the dissimilarity of organizational level – GEW is nationwide, BLLV Bavarian specific – and a changed age structure of members against a continuous stable ground, are markers of difference that direct the focus of GEW to reestablishing (internal focus) and BLLV to growth (external focus).

GEW is an independent education sector union affiliated to the German Trade Union Confederation (DGB), which organize educators and teachers in schools, universities, early childhood education, vocational training and adult education. In Bavaria it has about 10 000 members, in Germany all together about 250 000 (GEW-Bayern, 2015). It is the largest organization in the educational sector in Germany, about 70 % of which are female. The age of the members has gradually become higher, *‘The age structure of our membership has shifted dramatically in the last decades. We have grown older’* (GEW, 2015). Each of the 16 Bundesländer has their own regional GEW executive committees responsible for much of the daily work in each Land. Every 4th year a convent is hold with delegates from all 16 Bundesländer. BLLV is an independent organisation with about 55 000 members. It is the largest and oldest teacher and educator union in Bavaria, counting 1861 as the founding year. It is represented in every region of Bavaria and has a strong local attachment with representatives organized in a hierarchic frame. Every 4th year delegates from all over Bavaria meet at the so called *Landesdelegiertenversammlung* where the governing body is elected. (BLLV, 2015a).

5.2.1 GEW (D2), BLLV (D3, D4)

The disparity in terms of GEW being the nationally largest education union though small in Bavaria in contrast to BLLVs long standing position as *the* education union in the state, expose a significant difference of positions in the state. A state in which traditions and continuity are highly regarded as mediated by the hegemonic state party CSU which base its position and legitimacy on the maintaining of traditions and typical Bavarian values, or cultural capital (cf. CSU, no date). *‘(D2) [...] there is a very strong constellation in the education sector that represents the teachers and takes part in the educational politic [BLLV] and GEW is in Bavaria, compared to the other Bundesländer, very weak’.*

One example of the BLLVs influence on educational politic according to D2, was the unaltered tripartite education system of the state, said to be a consequence of BLLVs stand on the issue, unlike

GEW that wanted a prolonged time in primary school⁴⁰. PISA 2000 seemed to confirm the CSUs vision of an excellent education system which had consequences for the unions work because why change a winning concept?

(D3) It was quite bad for us [BLLV] that wanted to take part in the education politic since the responsible in charge could always ask why, we were already the best, as also when we asked for reforms such as decentralization and projects. (D4) It is more difficult now [...] they say that we need the three different school forms from 5th grade and that is decided [...] and when other Länder introduce reforms [of the tripartite system] and show improvements then they have no interest, they say it is not relevant for us. In Bavaria the notion of the individual's right to develop its talent and personality is very strong also among the parents, it's a very fixed talent concept.

The notion of the right to develop one's abilities is not specific for Bavaria, it is for example inscribed in a KMK decision (2009) for all the German Länder. However, this declared and decided individual right in line with the *doxa* of ability, genetically transferred and inherent, was said to be a cause of professional distress among teachers responsible for setting the marks in 4th grade, the marks which determine the students' path of secondary education. D3 particularly pointed at the limits of marks in pinpointed school subjects as set by ILSA (mathematics, science, literacy) to show the full extent of the individual's ability, and how it was used by the educational authorities to justify the education policy (cf. Schneider, Preckel and Stumpf, 2014, p. 6).

5.2.2 PISA and economy

D2 and D3 underlined that the PISA test construct caused Bavarian students problems since the traditional tests did not include such forms of testing (multiple answers etc.). The validity was questioned on grounds of students and school forms not included, for example the good results of *Sachsen* were a relative high proportion of students in special schools had been excluded. Furthermore, PISA had played a decisive role in pin-pointing socioeconomic aspects but no real change or guided management had followed.

(D4) For example the social links or the magnitude of social selection, the CSU never wanted it, never wanted it to become public. It was only done when Bavaria took part in the study and then it was rapidly pronounced as an issue in all the Länder. (D3) [...] The link between educational success and the income and educational level of parents, is still exceptional compared to other Länder. (D2) [...] also that the possibility of education in Bavaria compared to other Länder has the strongest link to the parents' income. This is in the [PISA] study and something to work on but what is said has had no effect.

The close relations between the political and the economic fields in Bavaria are part of the problematic socioeconomic differences in education, in other words, diversity as a necessity. 'The focus of the Bavarian school system lies on a custom-fit education to support individual talents, skills and the plans of children and teenagers for their future', 'The dual vocational training system for young specialists is a guarantor of high quality standards for Bavarian employees.', 'The Bavarian school and training system provides the requirements for Bavaria's companies to succeed globally.' These passages are all excerpts from the brochure *'Bavaria. A guide for Investor'* by the Bavarian Ministry of Economic Affairs and Media, Energy and Technology (2014). The education system in the state portrayed as a developer of future possibilities, a signatory of a promise to secure the needed skills for the economy where dedication and loyalty is inscribed in the workforce.

⁴⁰ Primary school consists of the first 4 grades.

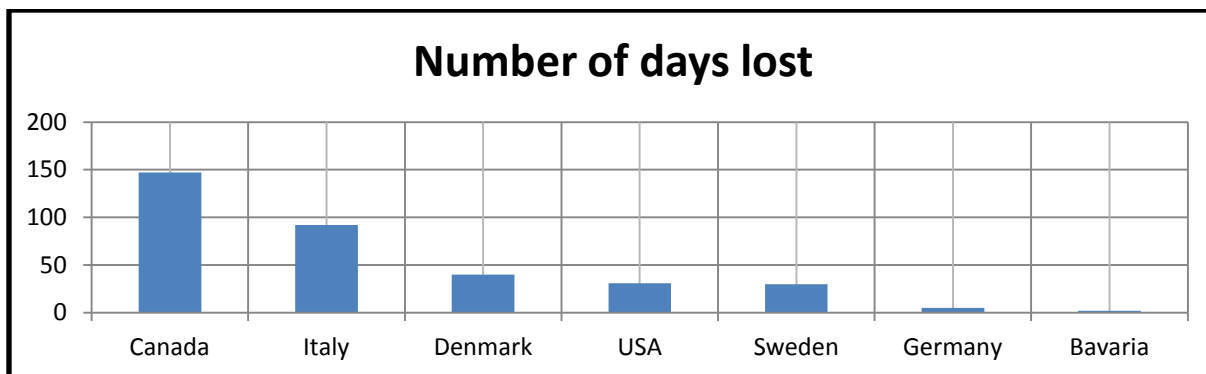


Diagram 2. Average number of days lost due to strikes and lockouts between 2001-2010 per 1000 employees (Bavaria. A guide for Investor, 2014, p. 83.) (Modified by the author.)

The education ministry's official claim on improving the educational equality, based on the findings in PISA, must be seen as part of the main policy objective of economic development in which stratification of education, in terms of securing the subdivisions of classes within, is taken for granted, a created (illusory) reality. An education system where privileges – referring to the socioeconomic differences – is countered, has the implication of adjusting the equity and thereby the human capital, in sense of limiting the provision of skilled adjustable `units`. ILSA, as OECDs PISA study, contains an implicit promise to secure an education system with the knowledge it needs to uphold the competitiveness of the economic performance of the state (cf, Hartung, 2011. pp. 90-91). If a state, like Bavaria, already is proven competitive in regard to other (Länder), what justifies then costly adjustments of equality improvements? Which is costly in a double sense, that is, the actions are economic costly as also `costly` in respect to the deflation of a political ideology based on distinction (cf. Rotte and Rotte, 2005, p 8).

However, the demography of Bavaria is a forbearer of change. Four out of the seven government districts are estimated to have had a decrease in population in 2020 (ISB, 2012, p. 6). Although the number of births is estimated to show a small increase in 2020, the long term trend shows a decline.

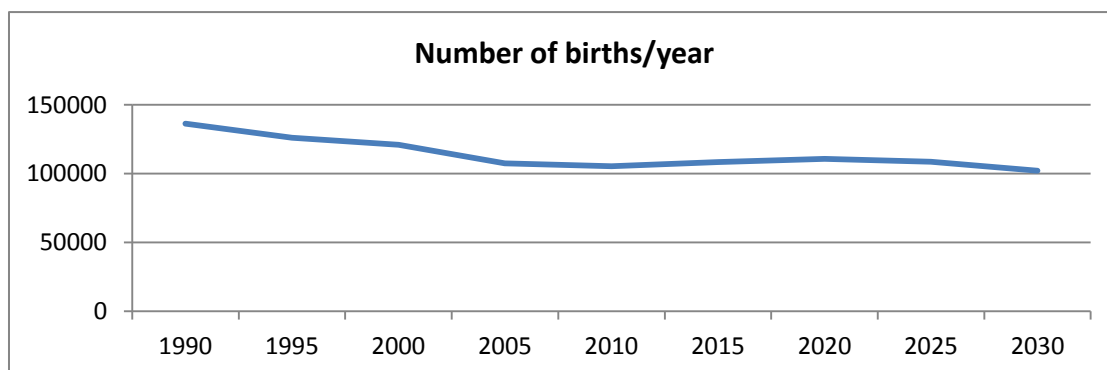


Diagram 3. Estimated number of births in Bavaria until 2030. (Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, 2014a, p. 25.) (Modified by the author).

In 2011 2,47 million of Bavaria's 12,56 million residents were immigrants⁴¹, 2/3 of whom were born outside Germany but in the age cohort 15 years, 9/10 were born in Germany (Bayerisches Landesamt für Statistik, 2012). Hence, 20 % of Bavaria's population were immigrants in 2011 and the steadfast former attitude of CSU on heritage and culture, has weakened (to some extent), probably coupled to the demographic trend, an increasing percentage of immigrants of the Bavarian population, that is, future voters. The sinking number of the work force also contributes to create a demand in search of

⁴¹ Person that has come to Germany after 1950 and is born in Germany with other nationality, including those who have at least one parent in the home that has immigrated, or was born in Germany as immigrant.

more people into the work force. It is visual in the Bavarian Ministry of the Interior for Building and Transport's (no date) declaration of the need to incorporate and promote older, part-timers, unemployed and immigrants into the work force, to secure the economy.

5.2.3 The focus on measurement

Before Bayern's entry in ILSA, the testing consisted mainly of examinations of a general nature at the end of each school form, something that has changed. The subjects and competence levels which are now tested are in accordance with the worked out concepts of quality assurance (ISB, no date), with the outcome of a more uniform approach, to the detriment of diversity.

(D2) It really took off with PISA [...] it wasn't a theme [before] because you were content and then this trend. You started to measure in numbers and it was made politic, extern and intern evaluations and also comparisons in Gymnasium [...]. (D3) More and more testing and it's in an idea that you improve the more you are tested [...] and our conception of teaching in BLLV is in line with process and reflection, and sometimes in PISA the result have their own [intrinsic] value and that is the only you look at and in that way you lose the multi causality. (D4) [The examinations] were never used to compare schools, to say this is a better school and this is not as good, and in relation to other Länder, it wasn't interesting and was never carried out.

In the schools the orientation towards output in terms of test results and the importance of, had given rise to a new phenomena, `teaching-for-the-test`.

(D3) [...] a lot [of teachers] prepare the students for tests so they will show good results, and in that way demonstrate the story of testing in Bavaria, on what basis these comparisons of schools are done, that it is completely oriented towards [viewable] results, numbers and degrees. You lose the oversight of how achievements are reached [...] is the total educational system caricaturized [...] do you learn something from these tests?

Testing has taken its own logic, within five months (February until June 2015), Bavarian students had to take part in five mayor assessments: VERA, PISA, TIMSS, IQB Länder comparison and the (local) orientation works (BLLV, 2015b). Even though tests always were a part of education it is a prerequisite that this *rite* be perceived as objective and thus universal (Bourdieu, 2000, pp. 306-307). The teacher unions (as all unions) are actors that base their symbolic and societal capital on collectivity and group membership, more so than other groups⁴². In an education system based on (proclaimed) democratic values, the legitimacy of the teacher unions depends on the maxim of impartiality, the *esprit de corps* of teachers. It is at the root of the capitals this group possesses: academic, social and symbolic. If the corner stone's of the reflecting image of teachers' make belief, may it be in the doxa of talent genetically transferred or development through training and guidance, the role of ethical and just gatekeepers of knowledge are put into question, it erodes not just the represented (teachers), but also the representatives (teachers' unions). That is the effect the testing discourse of ILSA have, because the test results are not (meant to be) put into practical use, it is just a symbolic act. In other words, excessive measurements will result in the decapitalization of the dominated positions in education. In the dominant positions however, the multiplied testing can be used as proof of action and decisiveness, the realizations and implementations of grand statements. Hence, it will serve an administrative and political function in justifying an educational policy based on its policy rhetoric and economics (cf. Brüsemeister et al., 2014, pp. 232-233).

⁴² Political parties are also collective organisations whose strengths depend on number. However, these are specified by their ideology whereas unions have the number – the amount of affiliated member from specific groups – as their specificity.

5.2.4 Standards, evaluations and accountability

Assessments, measurements and thereby the necessary development of standards were soon to become the visual token of success or failure, but what were the gains?

(D4) [...] with the measurements, the development that followed were the very forceful educational standards [...] and the measurements [proved] if you had reached them or not [...] it was a very strong notion of 'new steering' explaining why the measurements were done but the individual school, in practice, did not benefit.

One major obstacle for schools in Bavaria can be portrayed as the centralized steering of education. The schools are denied real influence and are not delegated the authority to make any major changes, thereby hampering the possibilities to address their special needs and problems. The Agency of Quality at ISB, the Bavarian state institute for school quality and educational research, manage the state's external evaluations of quality in schools and make recommendations for the schools which are followed up within a stipulated time (ISB, no date/a). The schools are in that respect accountable for the improvements but there are no real sanctions in practice.

(D4) The ISB organize teams that visit and view the schools. They have worked out manuals with questions and so forth. They also attend lessons and discuss with the principals, from time to time also with some parents. Then they make a report with two or three recommendations, the implementation of which has annoyed [because] after some days they just ask the schools what they have done and then nothing. (D3) [...] you can always learn about the results but not how to deal with them and in this I can see a system critical aspect. There is a high degree of school accountability but even though you got that, the resources are centrally controlled and rarely handed out.

In education standards means that some specific subjects (or areas) are mapped out as particularly important, selected as control knowledge, *Steuerungswissen*. ILSA are build on this ontological consecration of system monitoring where governments and expertise are united in their belief in a developed reliable control instrument which will improve the education system (cf. Tillmann, 2008, et al. pp. 16-17). More specifically, by standards the outcome – seen as the product – of an education system can be value dated and schools can be made accountable. It is furthermore the measuring stick of individuals and schools (if it is) set by and adjusted to, the specific circumstances of the educational segment or part, on which it is applied. However, standards that are set in the upper hierarchical levels (the dominant position) where the outcome is general, not adjusted to the specific circumstances, involve a process that will alienate and divide. In schools of the lower strata with few resources, the possibility to make improvements will not be realistic if not extra resources are supplied. In contrast to schools which have better resources based on a 'better' clientele, students' with socioeconomic advantages. The personal identification of the educational ethic's equality goal, teachers' professionalism, dedication so forth, will strengthen and confirm the self image and identity of the latter while the reverse occurs in the former case (Bourdieu and Passeron, 1977, pp. 178, 198). Standards, how they are established, interpreted and conducted and adjusted or not, in short who benefits and takes advantage of, is in this sense a way to shade the hierarchy and distribution of capital within a field (Gorder, 2000, pp. 330-331).

5.3 Institutes of quality assurance

The reintroduction in Germany of ILSA resulted in the establishment of local and national institutes for the monitoring and safeguarding of quality in education. The Centre for International Student Assessment (ZIB) is a national institute founded by BMBF and KMK in October 2010. It is located in Munich but part of a tripartite cooperation: The School of Education at the *Technische Universität München* (TUM), the Leibniz Institute for Educational Research and Educational Information (DIPF, Frankfurt) and the Leibniz Institute for Science and Mathematics Education (IPN, Kiel). The goal was to increase efficiency and promote synergy effects by combining the competence of the three

partaking institutes and the FDZ, the Research Data Centre (ZIB, 2012a). ZIB is part of the Institute for Educational Quality Improvement (IQB) which ZIB cooperates closely with (KMK, 2010a; ZIB, 2012b). ZIB's area of research is ILSA and the largest project is PISA (ZIB, 2012c). ZIB was the national project manager of PISA 2015 but it was already in 2012 responsible for the national PISA study. The aim for ZIB was from start to secure the quality of the PISA studies, promote research based on these studies as also research needed (KMK, 2010a). Professor Doctor Manfred Prenzel was the first chairman of ZIB and Professor Doctor Kristina Reiss the second (present) chair of the board (TUM, 2015). Prenzel had been a former director of IPN and involved in the PISA studies in Germany since the start. The story of ZIB is connected to the demand of new and improved centers for teacher training. In 2009 TUM School of Education was created, responsible for the education and post training of teachers. Professor Doctor Manfred Prenzel was elected to the professorship and also became the first dean with Professor Doctor Kristina Reiss as the prodean.

This background shows the close (inter-) relationship between TUM School of Education and ZIB, something that is characteristic for the limited numbers of members in the community of EB involved in the PISA studies. Several of whom has gained high or leading posts in institutes involved in the conduct and follow-up of ILSA in Germany (cf. Hartong, 2011, p. 151). Gläser et al. (2014, pp. 283-289) state that before the *Empirische Wende*, the number of researchers involved in quantitative empirical research were very limited depending on meager funding, a direct consequence of the limited contribution of practical value the field produced. The result was that there were only three research institutes – MPIB, DIPF, IPN – that specialized in QCE, all funded by the different Länder⁴³. When the demand for empirical comparable assessments once again increased, in the aftermath of the unification of Germany, only these three institutions could carry out and perform large scale measurements. Since QCE are not only costly but also highly complex and demand specialized knowledge, the position of the limited numbers of this research community were all the more safeguarded and unchallenged. Other educational research institutes and researchers came to adopt the empirical quantitative approach, a prerequisite since this was in increasing demand and could provide funding, compared to the decreased demand and funding of qualitative, or classical, educational research. Research institutes or part of, even had to give up previous research in order to stay in the field and adjust in line with QCE (ibid., p. 284).

5.3.1 ZIB (D5), TUM School of Education (D6)

The establishment of research institutes (like ZIB), programs for teacher education (like TUM School of Education) and a focused interest on teaching practice were all consequences of the PISA result. Whereas other nations created national project managers to handle measurements such as PISA, Germany created scientific research institutes with the aim to handle and analyze national and international measurements as also to develop research in education, and in that way develop special knowledge and skill.

(D5) [Other countries] establish national project managers but with no scientific [base] and the data must have a decisive quality if they are to be significant and used as benchmarks [...] I think Germany was quite clever because from the start, we let it be a part of science and research, as in ZIB. We now have 3 very strong established research institutes and IQB is responsible for storing data from and research of, such studies for others to use. (D6) TUM has thirteen faculties and the latest/last was the School of Education which include and unite, educational science and didactic [and] the main chair was that of EB. Manfred Prenzel was the first holder of the chair and this was important for the development and establishment of ZIB in Munich [...].

⁴³ The funding model provided a 'shield' against influence from other actors in education, being situated outside of the 'University sphere'.

In 2006 KMK and BMBF decided to take over the comparison of the Länder from PISA-E, that is, not to take part in PISA-E. Instead the IQB in Berlin have since been in charge of the Länder comparison with IZB doing the groundwork. The organizing, data collection and processing is run by IEAs Data Processing Research Center (IPC) in Hamburg and the final stage – the analysis – by IQB.

(D5) [...] well, there is no more PISA-E [...] and at the last Länder comparison [...] Bavaria was not the best, instead it was Sachsen with better results than Bavaria, and that is politically quite interesting because now you will see what Bavaria will do with this outcome. Before they could say that okay we have a higher selectivity but that is the reason we are better, that is really not true anymore.

There is a corresponding similarity between the response by the state (the political field) of Bavaria on the one hand and representatives of EB (the QCE field) on the other. Although diverse in articulation, the core has been the *ressentiment* of change. The official response to the Länder comparisons in Bavaria – ever since the first PISA study – has been to proclaim the educational system's par excellence. This view has been contested and later measurements, such as the Länder comparison by IQB 2012, has marked a rise of other Länder's results and the gradual decrease of Bavaria's (IQB, no date/a). The response has not been in accordance, instead the official version, as presented by the Bavarian education minister Ludwig Spaenle and (on the national level) the political coalition CDU/CSU, have been to portray education as constantly improving (Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, no date/b; CDU/CSU, 2013). EB, as represented by institutions such as ZIB and IQB, has been able to advance their claim on interpretative prerogative based on specific knowledge and skill to such extent as to have become the judge of who is allowed access to measurement data, and by that, who is allowed to validate measurements. However, these positions by the political dominant in Bavaria, CSU, and the research dominant, EB institutions, mean that a struggle is developing in which the interpretation of measurements data is at the locus. The legitimacy of education policy and by that the politic, is at stake. Hence, EB could be said in a process of increasing its autonomy versus the holder and source of symbolic capital (the political field) on which QCE was dependent (Süddeutsche Zeitung, 2013). This misrecognition of *genealogy* – promoted, guaranteed and closely linked to the political field – bears the marker of a *de facto* change in the relation between the 'servant and the master'.

5.3.2 Quality and selection of competitive criteria

(D5) [...] the central themes are international school performance studies, such as PISA, and the weight on a certain visibility, that is, Germany as scientifically positioned and positioned in research [...] and what you can make of the results, for example how you can translate the findings in our PISA rapport into school quality.

School quality is materialized by the results of certain selected and measured parameters; it stands for the visualization of a school's competitiveness. In the aftermath of ILSA school competitions became established. One most prominent is *Der Deutsche Schulpreis* (The German School Price), a competition⁴⁴ established 2006 with a price of 100 000 Euro. It was established in the frame of *Die Deutsche Schulakademie* (the German School Academy) which is said to be both an agency for quality, platform for networks, post educational guidance and a spokesman for successful school practices (Der Deutsche Schulpreis, no date/a; *ibid*, no date/b). The sponsors are two private foundations, the Robert Bosch foundation and the Heidehof foundation, and it is promoted by the two media partners' Stern⁴⁵ und ARD. *(D6) We have in Germany a school price with a quite high reputation, der Deutsche Schulpreis, with six worked out criteria that I agree about [...] and you also try to apply these School quality criteria in [other] evaluations. In Bavaria all schools are evaluated in fixed intervals and*

⁴⁴ In which the winner is committed to participate in the promotion of the contest.

⁴⁵ Stern is a media corporation owned by Bertelsmann, a multinational mass media corporation with specific interest in education (Hartong, 2011, pp. 252-265).

then these issues are studied. Quality can serve different purposes, how it is defined determines the precise, or vague, demarcations of its specificity and utility. In broadest sense, it can position a nation's education system – standing and status – and by that, quality can serve as a competitive marker of high standard. In the narrow sense, quality can be used to position a small part of an education system as an individual school, a class, a group or even a single student. The German standards, the interpreter of quality, were a cooperative and beneficial work by DIPF, IPN and involved experts (Hartong, 2011, pp 170-172). BMBF implemented the work in cooperation with KMK and the mission was given to DIPF and Professor Eckhard Klieme⁴⁶, to select an interdisciplinary expert group⁴⁷. The aim was to *'[...] develop and implement compulsory Bildungsstandards on which to base the quality development'* (Klieme et al., 2007, p. 14). The recommendations of the so called *Klieme Expertise* included that the Bildungsstandards would be adopted in accordance to each educational context, each of the Länder's education system, thereby avoiding any pitfall between theory and practice (ibid. pp. 9-10). It was most likely a necessary precondition to justify and promote acceptance of quality standards, as seen in the extensive declarations in the report (ibid, pp. 55-70) but at the same time the vagueness and differentiated contexts in which standards were to be applied – the 16 Länder's educational systems – meant that other actors could (and would) influence the specific quality areas of which education would be assessed. Not just the ILSA predetermined prime school subjects specified in the report, such as mathematics, science and reading literacy (until then a term unknown in German education), but also subareas listed in the German school price which in turn would direct the practice of work for educational actors such as administrations, schools and teachers.

5.3.3 The rise of EB

QCE was a minor research field in Germany until the late 90s when ILSA made it the research field in demand (cf. Fickermann and Mauritzen, 2014, pp 11-13). Since then, the *production* of empirical research data has increased each year.

(D5) [...] well, EB did exist but not that explicit, not to such an extent [...] educational research – it really has to do with TIMSS, PISA, PIRLS – became more expressively focused on data, less on theories, from level of consideration, level of philosophy [...] became aware of the possibilities to base knowledge of our education system on data. (D6) [...] you can tell from the number of chairs [of EB] and number of published research that the acceptance has widened [...] there are the departments, the German Science Community [DFG and] there are several foundations ready to invest in education also in [EB] research, we have no problems to recruit. (D5) [...] Yes, it is quite interesting, I think. They say it [EB] is at the front, ever more important. You can see that the number of professors in the field is increasing but there is also the discussion of the amount of data and the need to evaluate it, otherwise there will be an overproduction [for no use]. I think that is the most important issue for the future, how do I [translate] these data into practical use, into practice?

To a large extent, the strong position of QCE is not just based on the field's ability to identify problem areas or deficits in educational structures, it is founded in the perception of QCE as the redeemer of practice, how to solve identified shortcomings, in line with the traditions of natural science. *'(D6) [...]*

⁴⁶ Professor Eckhard Klieme the head of the commission, started with an educational focal point in psychology and mathematic and became 2000 docent in Erziehungswissenschaft. From 1983 until 1997 he worked in the Institute for Test Development and Talent Research (ITB) established 1971 by the German National Academic Foundation which was founded by public and private donors, with the aim to promote especially talented students. Klieme was 1998-2001 at the Max Plank Institute for Educational Research and has been both a coworker and manager of the department of Educational Quality and Evaluation at the DIPF. He has also been involved in the PISA studies and TALIS, another OECD study.

⁴⁷ With representatives of general *Erziehungswissenschaft* and EB, Teaching (and Learning research, Education Law and Historic), Systematic *Erziehungswissenschaft*, Pedagogical Psychological methodology as well as Didactics of Mathematics and Foreign Languages.

and EB of course use scientific methods and try in that way to promote the understanding of complex events [...], ` [...] these verifications of ideas and theories can be further developed thanks to EB [but] all the [EBs] limitations, I think that is sometimes not enough considered [...].` In Germany the controversy between the quantitative and the qualitative educational field, has been apparent ever since the rapid growth of EB. Statements by individual actors in education, researchers and published articles and reader's letters (in paper and magazines) display this conflict.

(D5) Yes there is [this conflict]. I think this is also for the benefit that these controversies exist but I often think that it is not necessary in ways of just being two different tools and the important part is always the questions asked in the research. I think that quantitative tries to expose the overall patterns [...]. (D6) [...] I mean we thought for a long time that we didn't need it [EB] and the focus was on philosophical grounded theories, besides we also thought we had a super school system [...] yes it was only with the ILSA that you understood that we needed to compare, and to compare you need data, and you need people skilled in this kind of research, able to handle the data.

The Research Center of Educational Research⁴⁸ at the MPIB was responsible for the first PISA study, MPIB was established in 1963 in line with a recommendation by the OECD. It was to become one of the centers in Germany for EB. The director at the Research Center of Educational Research was Professor Baumert, also involved in the TIMSS 1995 survey, and it is a name that is bound to show up in discussions with EB representatives. Baumert was one of the few in German research who had the knowledge in theoretical philosophy combined with insights in the quantitative field. Professor Baumert advocated the enlargement of the study, established a theoretical foundation with key terms and included a questionnaire in spite of the limited timeframe – caused by the late decision of KMK to take part in the study. Not least, Baumert also underlined the necessity to repeat these kinds of studies for the benefit of knowledge and education, and managed to extract a literary serial from the results. It was published in nine main volumes: *`D6 [...] an incredible achievement, without it we wouldn't be where we are today`*⁴⁹. The narrative of ILSA in Germany is closely related to the person Baumert in a way that more accurate can be described as a certain reverence based on the notion of prime knowledge in both theory, based in philosophy, and practice, based in EB. This dual competence or high symbolic capital in two educational fields mattered for the impact of the first studies (TIMSS 1995 and PISA 2000). The small community of EB institutes in Germany were all present in the TIMSS 1995 and PISA 2000: MPIB, DIPF, Humboldt University and IPN (e.g. TIMSS, 1997; KMK, 2003). The lack of other educational research fields in the work of these (early) ILSA can be perceived as logic and obvious, `common sense`, because of the specific knowledge needed to design the tests and interpret the results. However, it bears the inscription of *territorial* safeguarding and misrecognition (cf. Webb, Schirato and Danaher, 2002, pp. 78-79). Insights and knowledge of other approaches were present by individual researcher but not representatives of other educational research traditions. What followed was the promotion of the idea on education measurability (Hartong, 2011, pp. 28-29). Instead of a cooperative alliance involving a wide spectrum of educational research fields to work out best guiding principles for education and policy decisions, the outcome was a shift of positions, relations and capital, both symbolic and economic.

5.3.4 The aftermath and consequences for practice

The results of TIMSS 1995 and PISA 2000 unleashed an extensive mass of critique targeting actors in the educational field, not least the practitioners, the teachers. This heterogeneous group became accused of a variety of shortages and educational reforms followed in quick succession. One example was the establishment of new educational centers for teachers as TUM School of Education. Though, after some time the German results in ILSA improved and normalization took place, the norms of

⁴⁸ The Center of Educational Research existed from 1996 until July 2010, when Prof. Baumert became Director emeritus.

⁴⁹ D1 commented it as *` [...] too much for anyone to read.`*

ILSA. The education debate now focuses on state issues. Among the educational reforms undertaken, the Gymnasium reform in Bavaria was especially targeted by D5 as an example of a hasty and ill-considered decision partly due to the Bologna reform, partly to the steering of education in line with economic theory. The follow up of early results showed that German students were older than their European counterparts and also that they were examined later. These areas were seen as obstacles for the economy which became the main argument to shorten the time in Gymnasium with one year, from 9 to 8 years⁵⁰.

(D5) [...] okay, then they shortened the time span of Gymnasium, but the mistake was that no one really considered the consequences for the Gymnasium, they just took the Curriculum and deleted some parts, the 8 year were excessively compressed and then they were astonished that [the students and teachers] were totally stressed, with the only outcome to change back again [...] I think that this school form was partly sacrificed. It was in accordance to the economical orientation, for the renewable of resources, that is, students seen as renewable raw materials [...] I don't know if it is possible to see in the data [...] but the sad thing is that it will just add to the educational panic. (D6) [...] the term educational panic is quite associated with it, that is, many parents seems to think, partly because of studies such as PISA, that children must go to the Gymnasium, otherwise they don't stand a chance [...].`

It was the economic *impetus* that became inscribed in the educational context, the idea of education as the single most prominent promoter of future competitiveness. The implemented reforms also targeted the stated deficits in reading and reading literacy. Study materials have been renewed as also programs directed to promote the knowledge in mathematics and science, the prime subjects of PISA and TIMSS, `D6 [...] it came in science [and mathematic] SINUS, it was a large scale attempt that tried to improve the Mathematic pedagogy`. SINUS was the first joint venture based on the findings in TIMSS 1995 in which State and Länder, in the Bund-Länder Kommission (BLK, 2001), coordinated a program in line with the outcome of the measurement. It was headed by Professor Baumert.

ILSA brought in not only an increased intensity of and an emphasis on, comparisons alongside the designation of areas of concern, it had wider implications; the inclusion of a foreign language with terms and concepts former not relevant or even used in the German educational context such as *Lifelong learning, Innovations, Performances, Evaluation and Competence* (e.g. Gräsel, 2011, pp 320-328; Spiel, Lüftenegger, Wagner, Schober and Finsterwald, 2011, pp. 305-319). However, the foremost prominent term was quality, so in a short time quality steering became the fashion. All the German Länder established some kind of *Quality Institutes* as also programs for the development of school and education by data, for example MARKUS/Rheinland-Pfalz, QUASUM/Brandenburg, QUISS/Nordrhein-Westfalen, SINUS/Bavaria (e.g. Lücken, Thonke, Pohlmann, Hofmann, Golecki, Rosendahl, Benzing and Poerschke, 2014, pp. 127-128). On another level, the introduction and *tributal* manifestation of the quality concept, changed and sharpened the struggle over authority. EB gained theoretical authority in form of quantitatively data backed research as a form of rational science construct, with pin pointed school subjects. The authority of practice – for example, which reforms that were to be implemented – remained by the political dominant, as seen in the shorten of the Gymnasium (and its reconstruction) based more on political legitimacy than rationality (cf. Tillmann et al., 2008, pp. 397-400). While the traditional authority of professionals became reduced, the politics claimed success, for example the permeability of the Bavarian education system that officially was said to have been much improved (Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, 2015a). However, D6 point to the fact that very few children outside the Gymnasium track go on to study at University. Contrary to the official version and proved high percentage of successful examined students at college, a published study by the Berlin Social Science Center (WZB) and the German Institute for Economic Research (DIW), show that the number of

⁵⁰ Another cause was the unification of Germany where the new Länder (former parts of DDR) already had an 8-year gymnasium.

students that do reach college is only 36 %, placing Bayern next to last in comparison of all the German Länder (Augsburger Allgemein, 2015).

The measurements and monitoring have for many schools had the effect of increasingly strained working conditions, both financially and in terms of staff (more to do). One further aspect for practice has been the intensified educational research in schools which have created a counter effect, a resistance.

(D5) My research area is school absenteeism and it is an area that many are interested in but my experience is that when I try to cooperate with schools – we of course have a dialogue but – they don't let me in. They don't want me there to study how they deal with this problem even if it doesn't cost them [...] I try to explain the approach of the research, they really don't want it [...] it is just perceived as a threat to them when someone from science come [...] I don't know how to solve this, I think it is a task that will take a long time before the relations of trust between schools and science are [restored]. Especially the teachers are skeptical and sometimes even aggressive, because of PISA [they have] the feeling of being controlled and monitored, and they are blamed when something don't work out so well. It is sad because I can see in the PISA 2012 that we really have improved a lot and we [researchers] have said that it is really thanks to the teachers.

Schools and teachers are targeted by the monitoring of their and their students' performances, by parents' anxiety over the children's future and by the management of schools – anxious to show good results. But this is only the logical outcome of the disparity between the QCE diagnostically measurements and the official (state) administration's follow up in form of a normative quality monitoring. Normative as in *best practice*, what for the moment is judged best, but with the absence of practical guidance of the monitored which means that every school has to work out action plans targeting their specific problem areas. Hence, the monitoring is based on the same evaluative and normative concept as ILSA, being descriptive but not explanative, thus avoiding responsibility for actions (cf. Maritzen and Tränkmann, 2014, pp. 42-43). When research become an obstacle for the practice in terms of workload, extra time spent and other negative effects with no constructive positive outcome in sight, practice exclude research. It can be termed a (predetermined) struggle between the repositioned (EB) and the dispositioned (practitioners).

5.3.5 Communication of results

The results of the research by ZIB are brought out and presented in varied forums and forms. One important dimension is the political networking, the base and safe keeper (in terms of support and funding) of continued research.

(D5) That is done foremost by professor Eckhard Klieme from the German Institute for International Educational Research [DIPF] in Frankfurt [who is] very engaged also in the networking with the politic [...] for example representatives of ZIB goes to the KMK and BMBF and have discussions of the results, to explain, take part in podium discussions and answers questions from schools, that is, visiting schools when they ask for it.

The data from EB research by IZB must be submitted to IQB in Berlin which stores it. Researchers can then apply for the use of the data but the request must include an extensive outlay of the purpose and content of the research. D5 underlined that transparency was a cornerstone in IZBs work. The public presentation of results from measurements data is presented in national reports, brochures and meetings with press and media, small concentrated summaries are made and many official presentations. To contribute to the transmission of information and good practices is another aspect of IZBs work, for example to let know Germany's positive trend—and developed methodology—in education.

(D5) [...] earlier this year, I was in Stockholm at Skolverket and at the Goethe institute [...] while naturally the Sweden's story ever since PISA 2000 is a bit different compared to

Germany's [...] it was quite interesting and an important task too for IZB that it becomes international understood that Germany base our research on evidence so to say, that others are aware of the development in [Germanys] educational system.

QCE requires a continuous, reinforcing net working, especially on the crucial political level. Social research always depends on political backing since it is research on the society and ILSA is research of education, the foundation of society. Hence, the implications are that QCE is dependent on political backing but also needs to be valid in the eyes of the practitioners, those who transfer and implement the legitimated dominant culture (Jenkins, 2007, p. 153). If this balance is upset, the importance of political support becomes even more crucial, but at the same time, the condition of politics is the confidence of the masses, the different *categories* in society as teachers. The only way to counter a loss of confidence is to portray the expertise as unique, *gifted* with a very specific skill necessary for the development of society. QCE must state its form of symbolic capital – specific knowledge of numbers – as fundamental for nations and states economical competitiveness (e.g. Hartong, 2011, pp. 71-73). However, the disadvantage of information (formatted) capital is that it vanish if not renewed, numbers is a perishable commodity and soon yesterdays news. Information capital must constantly be renewed and perceived as renewable – ILSAs are recurrent studies – and reliable in sense of the scientific theoretical foundation. (cf. Zlatkin-Troitschanskaia and Gräsel, 2011, pp. 11-17).

5.4 PISA- Scientific knowledge or fabrication

In Germany in the aftermath of PISA 2000, a struggle of interpretation started. Critical interpreters were at first few and scattered but for each successive PISA study that followed, the critique intensified (e.g. Jahnke and Meyerhöfer, 2006; Hopmann, Brienek and Retzl, 2007; Wuttke, 2009; Meyer, H-D and Zahedi, K, 2014). The experts, academic researchers or academics professionals in the field of education, mathematics and informatics, became divided about the measurement methods, validity and the interpretation of the results. The official representatives and researchers involved in the PISA studies rejected for a long time promptly the presented criticism without any lengthy comments with some rare exceptions (e.g. Prenzel, 2006; Hopmann and Brienek, 2007, pp. 13-14), but a shift of position has occurred, representatives have engaged in answering the critique, often in the form of a collective (and fierce) response towards individual researchers (e.g. Prenzel, Baumert und Klieme, 2008; Klieme, Köller and Prenzel, 2013).

5.4.1 Experts: `Opponent` (D7), `Proponent` (D8)

In Germany there is a wide gap between educational researchers that are positive versus those that are negative towards ILSA, the intensity of measurements in education and the interpretation thereof. *`The strife really surprise me, the opposite camps don't listen to each other, I don't understand it`* as a prominent researcher in Erziehungswissenschaft declared⁵¹. D7 is a stern opponent towards the constructs and interpretations of the PISA studies ever since the beginning and has, as an expert in the field of informatics, published the critic in detail. D8 is a professor in pedagogy and an educational researcher with a foremost positive view of EB but has underlined the need for a multifaceted and broad perspective, where research must be conducted on depth. TIMSS 1995 is said to have been the first study that demonstrated weaknesses in the education system.

(D8) Yes [TIMSS 1995 was acknowledged] but not publicly and also, for example in the vocational education it was already in the 90s known that there were a functional analphabetism, that young people had difficulties in reading and writing but it was not believed. It was only due to the PISA study, a measurement managed by the state, that it was made public knowledge [...] it was also dramatized, excessively fearfully perceived. We already knew some parts but this, that you discussed it in politic, that you developed programs for

⁵¹ Private meeting in Münster, March 2014.

example for whole day schooling [comprehensive school] and early childhood educational training [...] it was now promoted.

For D7 the first encounter of ILSA was the TIMSS 1995, especially after a lecture for teachers in which the debate afterwards circled around the result. Then in 2001-02 the PISA (I) study was published and became the onset of a whole set of critical questions.

(D7) It was in 2001-2003 and PISA was the major theme [...] I got a report of PISA and the reading amazed me a lot, which is why I downloaded the original data and presentation and made my own evaluation. [TIMSS and PISA] are built on the same kind of statistics, teachers were surprised of the outcome but especially how PISA was orchestrated [...] how PISA as a major achievement was published, for several weeks it was announced, it was well done and also how it was revealed a couple of days before the official release date to increase the intention, orchestrated with maximum success.

TIMSS 1995 was similar to PISA in construct and setup, but PISA brought in the subject of reading literacy, a concept unknown in Germany (GHR, 2005). The PISA studies all had their special themes or critical points which enhanced the interest even further. However, the interest in the measurements as also their impact in media, has declined even though the results are still broadly commented and interpreted: *'You get used to it and at the beginning it was a shock because the education system was believed to be better than the results showed (D8)', [...] and now with the study every third year, they arouse less and less attention (D7)'*.

Ever since the TIMSS 1995 there had been an awareness of particular deficits in education. The first PISA studies highlighted some of these special problem areas that researchers in education, for example in the multifaceted Erziehungswissenschaft, already had mapped out. In that sense the findings were welcomed by a variety of actors in education such as researchers, unions and teachers, and legitimized the use of measurements as the sanguine way to improve conditions in education for boarder groups and – not least – deliver solutions on German educational problems. The political field received an educational diagnosis, descriptive data, whose interpretation and translation into action plans, they were to manage. The result was a massive funding of EB research. This research field was in a rare position to control and run the game at the start since they in practice already had a hegemonic position in terms of total control of the skill and knowledge in demand, interpreted in a specific technocratic language. The attendance of this line of research meant that in 2012 the *Society for Empirical Research in Education (GEBF)* was founded with the specific goal to single out EB from the German Society for Erziehungswissenschaft (DGfE) and mark out a territory of its own. DIPF and IPN were represented in the board with Professor Olaf Köller (IPN) in the chair (GEBF, no date).

The doubts of the German – and the Länder's – education system meant that past educational research came in disrepute, in particular such research that went under the label Erziehungswissenschaft where an internal debate started about its legitimacy and content (Fatke and Olgers, 2014, pp. 7-13). In Bavaria the top results just confirmed what was *'well known'* and as a consequence the media circus were not as intensive as in other Länder. The concern for Bavarian politicians was instead the education systems verified short comings to handle low performers such as immigrants and groups of lesser socio economic level.

5.4.2 Bias and stressed educational deficits

The performance standards, the quality standards for each educational path and subject, are not fully implemented in all the Länder, partly because of the autonomy of each Land, though the aim is to reach common standards. The length of time consumed for the insertion of measurement models and standards differs in the Länder which can be interpreted as depending on the degree of correspondence between the measures and the political discourse. Bavaria with its conservative, right-wing ideology, were especially well adapted to a testing regime, *'(D8) But Bayern you know,*

have always focused on this aspect of license [educational measurable performances], it has always been very much in focus, more than other values I would say`.

The promoted performance culture in Germany has resulted in the development of new test- and evaluation programs, adopted in the Länder, in accordance with the economization of education, endorsed and sponsored by educational actors such as economic interest groups, employer organizations, bodies producing education/school material but also teacher unions (e.g. BDA⁵², 1998; DEGÖB⁵³, 2015; Deutsche Lehrerverband, 2000; Piopiunik and Wössmann, 2011, pp 40-41). This may not have been an issue in itself had it not been for the position of the German PISA experts major role and dominant position in the development of these programs.

(D7) The German PISA reports starts with dry statistics and then give some concrete recommendations. I do question if the recommendations really are possible to set in practice. Some are most surely quite good [but] these people are responsible for the PISA studies and the publication of it, and they want to achieve some specific ideas and these ideas are always `found` and confirmed [in the material they themselves have developed] and because of that you just can't judge them as positive or bad [...]. And a funny aspect is, well now when the study show improved results, that is, the instigated measures are said to work for example in science, then one of the front persons for PISA in Germany – who is a big name in science didactic – happen to be the one who have developed a special program for the evaluation. [In other words] he evaluates himself.

The examples of these close bounds are numerous (e.g. Hartong, 2011, pp. 143-144). One example is the development of SINUS, a program for Increasing the Efficacy of Mathematics and Natural Sciences Education (BLK, 1997), a program widely in use. It was one of the first testing programs developed after TIMSS 1995. Professor Baumert and professor Prenzel where among the key people and coordinators of SINUS, and IPN was one of the institutes to manage the development of the program (BLK, 2001). TALIS, Teaching And Learning International Survey, and DESI, German English Student Performance⁵⁴, are two other examples of tests, both headed by DIPF and Professor Klieme. Another significant example of the outcome of the promotion of longitudinal and quantitative research is the National Educational Panel Study (NEPS), a major longitudinal study based on EB whose aim is to analyze and monitor educational processes in Germany from early childhood to late adulthood. It consists of an interdisciplinary consortium of research institutes and researchers, coordinated at the University of Bamberg (Bavaria), located at the Institute for Longitudinal Educational Research⁵⁵. The project had six starting cohorts with about 100 000 participants and was funded 2009-2013 with about 85 Mio. Euro by BMBF which can give an idea of EBs current position.

The first PISA study highlighted a specific problem for Germany and Bayern that contributed to the partaking in ILSA, the predicaments with the large group of students that passes the education systems without exam. Even if vocational training is a quite extensive part of German education, built as a parallel track to the grammar school, it is mostly out of reach for the group of low performance students. Bavaria boasts of having special transition tracks for this modern *pariah* of education.

(D8) These [transitional tracks] have a large part of the students, this is most critical and of course it cost a lot of money, and they do not have the same value as the traditional vocational training or education, just short-term valued certificates. But politics is once again about that some [students] will make it.

⁵² BDA (Bundesvereinigung der Deutschen Arbeitgeberverbände), the Association of German Employer.

⁵³ DEGÖB (Deutsche Gesellschaft für Ökonomische Bildung), the German Community for Economic Bildung. See DEGÖB, 2015 – list of publication.

⁵⁴ DESI (Deutsch Englisch Schülerleistungen International) is a specific German measurement.

⁵⁵ INBIL (Institut für Bildungswissenschaftliche Längsschnittforschung).

In Bavaria the *doxa* of a highly successful differentiated educational system is inscribed in the official rhetoric of the state's governing political body, the CSU. The *commodification* of education (education as capital) and the participation in ILSA has brought up the issue of the body of students that exit the system without necessary qualification to meet the demands on the workforce. This is just one of the problems facing the education system: decreasing demography, increasing percentage of young people with immigrant background, relative constant number of school leavers without qualifications, regional disparity within Bavaria and increasing competition of educational centers and schools. In spite of the verified findings of such casual factors as socio economic disadvantages and language deficits, the official reply has not included any structural relationships. However, nor has the individual's own responsibility been stated, which often is pin pointed as the dilemma for low performance in education and economic competitiveness, in line with the ideology of the governing party. If it had, the predicament would be the questioning of the *doxa*. Blaming, more or less explicit, the individual would be to incur a debate of the verified findings of inequality in the education system that would involve the structures of the educational system with its tripartite system, and implicit its reproduction of core ideological values. Instead the dominant has avoided the strife and focused on schemata to strike back and prevent the eroding of their legitimacy. It includes public statements about grandiose plans like in the CSUs government plan for 2013-2018:

In Bavaria we have almost defeated the unemployment for young people. In the future it will disappear. Every school leaver will have access to a qualified vocational education path. We will have none educated without an exam. (CSU, 2013)

The factual measures taken has been marginal as transitional paths to lower the number of less qualified or those without examines, though the number of enrolled dogmatic slogans has been plenty: *Qualität und Gerechtigkeit* (Quality and Justice/Equality), *Investitionen in Bildung* (Investments in Bildung), *Fördern und Fordern: Wissensvermittlung and Persönlichkeitsbildung* (Promote and Demand: Knowledge transfer and Personality development) and *Kein Abschluss ohne Anschluss* (No school leaving without connection) (Bayerische Landtag, 2013).

5.4.3 Quality and interpretation of PISA

Comparable education statistics were for long an unwanted theme in Germany, to compare the different Länder's educational systems was out of question but the outcome of TIMSS 1995 paved the way for change.

(D7) With TIMSS it wasn't possible to compare the Länder, a topic that for decade's was taboo [...] everything about testing intelligence was most unwanted [...] the outcome to instigate quality assurances is everywhere now, quality institutes as IQB in Berlin are the direct results [and] the first director of IQB came from the circle of PISA experts. (D8) Now, some people believe that performance is the same as education [but] education is so much more, not just specific subjects. I am very much in favor of these measurements, that I want to say. I think that these educational measurements are necessary also to receive important feedback [...] you need them in different subjects, in order to develop standards of quality, that is, minimal standards not maximal standard, which all should achieve. [But] it would be detrimental for educational research if all were to be focused on measures of performances [...].

The supremacy of comparable measurement that for long was an unflinching concept for quality assurance, soon reached a stage where the inherent promises could not be met in practice. Even projects as the DFGs to establish new education research centers in EB were less than successful when out of 39 outlined proposals only 3 research units in the end were accepted (Niessen, 2011, pp. 162-166). However, what has been achieved is a developed infrastructure of performance monitoring by the establishing of (EB) institutions and quality institutes on state and federal levels (cf. Zlatkin-Troitschanskaia and Gräsel, 2011, pp. 9-11). The specificity of the early ILSA were their ability to build up and expand the necessary political support, partly due to the interpretation of the

furthered data. According to D8 the former resistance in Germany to compare Länder was based on political issues. The SPD controlled Länder did not wish to make public the cost of their more equality education policy compared to the CDU/CSU controlled Länder. *(D8) [...] the SPD had a different aim, social equality, and they did not want it to be known that it also cost [...] and [the union] GEW was particularly against any measurements. In the first PISA study they called for a boycott in Berlin`.*

Yet, a specific part of the interpretation of the PISA 2000 data had the effect that the political left in general, and especially SPD and the union GEW in particular, began to shift position and became more lenient towards ILSA. The turning point was the issue of comprehensive schooling in form of all day schooling, a perennial topic in Germany depending on the more customary educational tradition of half-days schooling⁵⁶ (e.g. KMK 2002, pp. 14-15; BMB, no date). The results of PISA 2000 were said to support comprehensive schooling which the political left favored: *(D7) And then Schleicher from OECD interpreted the results as pro-all-day schools and that is why the whole left spectrum became pro-PISA`. Whole-day school became the miracle cure in education.*

(D8) The debate of all-day schooling then developed [...] here it was always half day education, in the afternoon many children were left to themselves [and] home works were not made and so on. With all days schooling it could be possible to make a better [educational ground].

SPD saw comprehensive schooling as an issue of equality but for CDU/CSU it was an intrusion of the families' sphere (cf. Tillmann et al. 2008, pp. 184, 187). This strong traditional view were also a determinant in the outcome of the all-day schooling debate following PISA 2000 where the conservative block managed to interpret – or change the interpretation of – the results as favoring their cause. *(D7) [...] you can always find something that supports your cause, and for the discussion in Germany, I don't know how they [CDU/CSU] managed but they said that the best results were in the Länder with conservative Systems.*

Evidence based education monitoring, the seal of ILSA, need longitudinal data without which it cannot be performed, not cross sectional. That is, comparable data based on recurrent quantitative tests. Since tests are the prerequisite of such a construction, the number and diversity of tests involved evolve. This is evident in education today where the number of developed and used tests, has grown in the last 15 years. In Germany the word *'Tester'* or *'Testeritis'* is used to describe the lot of tests and measurements in school and education.

(D8) [The testing] occurs maybe a little too often, a kind of evaluational tiredness grows, we already have it. You notice it in the schools, sadness, frustration, resistance and defense, I notice these issues too [...] the OECD do not change it [PISA study] and why should they, it is a splendid business model, they know it's read [...] but basic, every four year would be better. [...] the negative is maybe that in school they do feel checked and controlled, this could be improved by fewer measurements but you have to measure.

Hence, the measurements are developed by certain *specialists* and conducted and /or monitored through quality institutes. The developed data is then stored (and guarded) by centralized data centers as FDZ at IQB, who allows access to researchers with the *correct* research design. It can be described in terms of an establishment of a dense job market, in number of new established carrier paths for experts, with recurrent disadvantages for the practice: increased workload and responsibility for teachers and administration on the local school level without benefits (e.g. Bernt and Remme, 2014). The measurement practice has affected the school practice and the recurrent extra work required of them. The developed monitoring cult has meant in practice the establishment of organizations with a top-down design, controlled by politic and performed by experts (Hartong, pp 178-179).

⁵⁶ See appendix 3.

5.5 State Institute for School Quality and Education Research (ISB)⁵⁷

ISB is a subordinate authority of the *Bavarian Ministry for Bildung and Kultus, Science and Art*. It was created in 1984 when the Institute for State School Pedagogy and the State Institute for Educational Research and Planning were merged into *Staatsinstitut für Schulpädagogik und Bildungsforschung*. In 2003 the name was changed to *Staatsinstitut für Schulqualität und Bildungsforschung (ISB)*. The work of ISB is to support and advice the Bavarian Ministry in the development and promotion of the differentiated educational system (ISB, no date/a)

The Quality Agency (QA)⁵⁸ is a branch of ISB established 2003 and can be described as a state level equivalent to IQB, where the work is regulated by political decisions. It is responsible for matters of quality assurances in the education system, evaluations and monitoring. The base is said to be *'scientifically established knowledge'*, in the way that QA *'collects data and evaluates them with methods of empirical school- and educational research'* (EB), and provide the actors in education – teachers, school administrations and (the state) school inspection – with feedback of the Bavarian schools Quality (ISB, no date/c). The aim is to further a systematic and data based development of schools and education/teaching. The QA has four focus areas: (1) identification of quality-relevant characteristics, (2) education reporting, (3) the Länder's comparison work (VERA) and (4) the external and internal evaluations in schools. The agency is because of that organized in four departments: QA 1-Education reports and monitoring, QA 2-VERA, QA 3- characteristics of school quality and their empirical recording and QA 4- extern and intern evaluation of the individual school. (ISB, no date/b). The establishment of QA, as the setting up of education quality agencies in the Länder, was a direct outcome of the altered view on education that followed the German partaking in ILSA. It was a change into output steering and the setting of educational standards (e.g. Erhard, 2010, p. 7; Berner, Oelkers and Russer, 2008, p. 210; Köller and Pant, 2010, pp. 55-57). The main principle of quality assurances – in comparable numbers or written testimonials – as *the* crucial guiding principle had hitherto not been a main issue even if such work had been the task of state institutes since the 90s (Schneider, 2006, pp. 13-15).

5.5.1 QA, the Quality Agency (D9)⁵⁹

(D9) [Comparison measurements] were there before but nobody wanted to be a part of because you thought it wasn't necessary. [Measurements] had occurred but the participation was very low. It was then, the predecessor to this ministry, they were called School Development and Bildungs Research, yes, they were, I think in 1997 or so., they were in Berlin and was informed of the TIMSS [1995] survey and that is the first time it was stated [the results of TIMSS]. I was in another department but I was with them and when you went home, you knew something were going on, it would come to us [ILSA]. That was the start.

The result of PISA 2000 in Bavaria did not lead to a negative outburst, but it did point at the inequality in the state's education system: *'(D9) That is, we have a relative consistent and stable [education] system but the PISA Shock hit Bavaria as well. We all thought [that] we are the crown of creation, well we weren't and since then [...].'* Even if the first results were good – especially in comparison to other Länder – there were an instant discussion at the education Ministry, according to D9, how to proceed and secure the Bavarian top position and deal with the less positive sign of high inequality.

⁵⁷ *Staatsinstitut für Schulqualität und Bildungsforschung*. Notice the term *Bildung*.

⁵⁸ *Qualitätsagentur*.

⁵⁹ D9 worked at the department QA-4 (Q4), intern and extern evaluations within QA. In the text both abbreviations are used.

5.5.2 How the concept of quality were introduced

The establishment of a special Quality Agency within ISB was a result of the participation in ILSA.

(D9) At that time there were diverse working groups with the task to decide the central aspects and I was here in the working group for Evaluation [...] well then it was around 2000 I think, we had a major Congress in Bayern [Podium, 2000]. The main interest was to further the development of schools, to make known the concepts of quality assurance and quality development. And then, well, there were also other levels in the work in the Kultus ministry when the first PISA results were published and they were stated in a memorandum. There was a Bavarian quality memorandum in which the focus areas were set and to the measures belonged the establishment of an Agency of Quality, that is, we became reorganized`.

The mentioned Podium 2000 occurred in April 2000. It was the Bavarian state ministry for Education that held a statewide school congress in Augsburg. The motto was `School innovation 2000 – Schools on the path` and over 100 schools participated in workshops, and demonstrated their specific initiatives and projects in line with the motto (Podium, 2000). In the opening session the education minister Monika Hohlmeier called for a new Bildungsdialogue. Hohlmeier stressed the schools` self responsibility and creativity based on the so called `12 Augsburger Theses for internal school development` which included 12 areas for development. Number 3 deals with the introduction of defined standards and TQM (Total Quality Management). The latter is a term from business management literature established in the 50s and 60s, based on the principles of scientific management where each unit of an organization is tested vis-à-vis a set of criteria, quality criteria (cf. Sirvanci, 2004, p 382-383). At the Augsburg congress the `Internal school development innovation price` was presented, a contest for Bavarian schools at different levels, instigated by a cooperation between the state and the foundation `Stiftungspakt Bayern`, in line with an increased interest by private foundations to take part in the `development` of the education sector (Stiftungspakt Bayern, 2015).

5.5.3 The daily work at QA

An extensive part of the work in QA is to ensure the quality of their own work.

(D9) Okay, we [QA] have basically 4 departments [Q1-Q4] but 2 main areas in our department. One is Bildungsmonitoring, the other is Teachers and then we divide the task of the external evaluation. My group deals with the conceptual basics, that is [...] which areas do in fact decide the quality and what is quality, and how do you measure it? That is our tasks and the other group[s work] is then teachers, [those that] we have here from different school forms, those that have the task to educate the evaluation teams, the evaluators. They have the responsibility to secure that the work [of the teams] is conform and standardized. They check the reached values [...] that is really the base to secure our own internal quality so we can be sure of the content of the work is made in the way we want it.

Overall, the quality work in the different Länder is quite the same, divided into a (small) number of chosen quality areas but the QA in Bavaria has a slight different approach: `(D9) [...] about the details, the structure of the different [quality] aspects, it`s maybe a bit different [...] we have four basic areas; (1) frame conditions and (2) results, and between (3) the process quality of the school and (4) process quality of the teaching.` These are in turn divided into 16 subareas with different number of criteria. The concentration of important school quality as stated in selected criteria or indicators, were supposed to make improvements for practice manageable but instead the link between have become questioned.

(D9) You can ask yourself if an indicator that is meant to measure quality really is appropriate as a practical tool for action. That is, does it bring a quality improvement to work by the indicators or is it two different things? It is one thing to admit flaws, another to take action

under the specific circumstances. The Kultus ministry has held us back so to say in order to avoid these misunderstandings and not overwork the schools with a giant list of indicators.

4 areas of quality

Framework conditions (descriptive)	Process quality of school (evaluative) (13 criteria)	Process quality of teaching and education (evaluative) (10 criteria)	Results of school work (descriptive)
--	---	--	--

16 subareas

Location of school	Management of school -Supportive -Goal orientation -Efficiency of organization	Course -Efficiency of used learn time -Efficiency of behavioral regulation	Characteristics of teaching
School community	The work of the collegiate -Openness towards the schools external field -Coordination of the collegiate work	Presentation -Structure of the presentation -Clarity of presentation	Level of learning outcome
Structure of staff	Development of school -Openness for change -The systematic of quality development -The systematic of monitoring	Design -Individual support -Promotion of self managed learning -Promotion of learn motivation -Securing of successful learning -Promotion of over-all competences	Satisfaction (with the schools work)
Materiel and financial resources	School culture -Attention of parties involved –Promotion of interests -Intensity of participation -Promotion of the school identification -Promotion of integration/inclusion	Teaching conditions -Teaching conditions for promoting the learning	
Organizational specificities -Offered specializations -Career options			

Scheme 2. The Quality Areas in QAs external evaluation (Bayerisches Staatsministerium für Unterricht und Kultus, 2010, p. 13). (Translated by the author.)

The first literature review of the term quality in Germany was made by Professor Hartmut Ditton in 2000, a researcher with a base in psychology and working with evaluation and quality in education as

also empirical methods. According to D9, this review together with the theoretical work of Professor Andreas Helmke became the manual and guidance for the conceptualizing of quality work in Bavaria and in many other Länder. Professor Helmke⁶⁰ was a representative of the empirical part of Erziehungswissenschaft, early positioned in the process of quality management of education with pedagogic and development psychology as the main domain. The Q4s quality evaluations of schools are of two overlapping kinds, external and internal (Lankes and Huber, 2014, pp. 5-7): The external evaluation report defines the schools quality work, strong and weak aspects, while the internal evaluation – measures to improve weak aspects – is done by the individual school. But it is not a straightforward task for what is quality in a school? The term has to be made manageable which means (1) to define the main or most important areas and (2) to rank them. The next step (3) is to define the most crucial criterion/criteria in each area which means to (subjectively) choose among a variety of issues and also to (4) define what quality in each of the chosen criteria is. It is a time consuming and indefinite work with many turns.

(D9) [...] we have the criteria of quality and then the process [...] we have tried to define each criterion of quality [...]. Then we have said that for these criteria to become distinct, become outspoken, the requirements must be specified. Then for each criterion we have three to four demands and we have carefully defined the key concepts, I will show you another example. In the criteria of teaching for example we have the efficiency of the spent time of learning and then this term is first defined. Then here comes the demand that the teaching was conducted without interceptions and then it says what interceptions are [...] we have tried to find indicators but the work is now being restructured.

5.5.4 Cause and effect-(too) complex matter

(D9) Effects, we have three areas when we deal with it. One is naturally to consider what the schools do after the external report, how they get on with it, what are the outcomes, target areas, processes. The second is that we try to establish if schools with better evaluation results really are better, is there a correspondence, is there an overlapping between students results and the school's quality criteria? That is also very difficult [to establish] because we have [chosen] such a number of different criteria [...]. We have good empiricists in our group that try to develop the knowledge of these matters. You must realize all the factors to analyze, trying to establish the most important results and criteria. We are in progress but have not been successful yet.

In Bavaria the first attempts to work with school quality started back in 1995 and continued until 1998-99, according to D9. It was the schools that managed the first attempts, guided by central authorities under the command of a very interested and devoted Kultus minister. The early start in Bavaria with school quality work – which brought in the term quality in education – and the good results in the first PISA measurements, created a working environment that was characterized by trust and a feeling of already having a good quality, sanctioning the state conservative policy.

(D9) That is also an advantage by us, this with the continual and conservative Bavaria that keeps its course. I mean that there are no signs that you have to change over and over again, yes, like in other Länder, for example in Sachsen with recurrent cuts in education. We are, thank God, quite self-reliant. [...] the [first PISA] results reassured us.

It has been twelve years since the QA was established in order to watch over the quality in schools, so what are the results? Have the agency been able to establish an order of cause and effect and what education quality work is? The recurring problems have been to establish and select quality criteria, to judge and rank them as also how to transfer knowledge of what the concept means.

⁶⁰ Codeveloper of a number of large scale tests as the German VERA tests and responsible for one of the first inner German mathematical test, the so called MARKUS (*Matematikunterricht in Rheinland-Pfalz*).

For example the evaluators' own (mis)understanding of how to use the chosen criteria of quality. Even if the structuring of each quality criterion is part of a whole, the QA never meant it to be used as a total summary concept. Instead each piece of information were to be judged in light of the specific criterion it concerned. *(D9) [...] and that is not the case, when just three from six [criteria], you can't say that only half are met, you can't interpret it that way and that is difficult to grasp`.*

The term quality is but one example of a field specific language, but one that is dependent on the adoption by a state, the political legitimation and confirmation (cf. Bourdieu, 1991, pp. 45-46). That is, all fields have specific languages; it is a central part of each field and a marker of its borders. However, the specificity depends on acceptance founded on political consent, without it the field cannot claim autonomy based on specialized knowledge. The significance of a term such as quality is that it can be *`understood`* because of its generalizability, the term is well known even if not (if ever and if possible) well defined. The field of QCE is a field where numbers are the prime locus and the effect of such numbers depend on the use of a universal – and positively – associated terminology because of the wide impact they accomplish when used. Or in other words, these words have a direct impact on a chain of actors, for example media and (through media the) public, which affect the political power, dependent on public support. The problem is though the weakness in the terms vagueness or impreciseness, which at the same time is unconditional when a mass effect is sought. The outcome is the difficulty with which these concepts can be implemented in the actual work in practice, exemplified in the QA external evolutionary work which repeatedly has been reversed. Furthermore, the report of the external evaluation work does not instruct the schools what actions to take, it merely give recommendations for the individual school to decide on, *(D9) [...] we have ten years of experience [and] it was a kind of special critique from the schools, they said that, well the evaluations do not contribute because nothing happens afterwards [no sanctions]`. In this respect Bavaria differs from a number of Länder which do have different kinds of approval systems a school must pass if deficits in quality criteria have been stated.*

5.5.5 Quality or just monitoring

The safeguarding of quality is an important part of the Bavarian school policy. The state wide curriculum and the central final exams are expressions of the efforts to ensure an equivalent and high level of quality in all the schools in the Free State. (Bayerisches Staatsministerium für Unterricht und Kultus, 2010, p. 7)

The public schools in Bavaria are evaluated within a five to six years interval and the results of the external evaluation by the QA (Q4) are followed up by the state school inspection. When the system of evaluations of schools was introduced, it met with considerable resistance according to D9 and partly refers this with a mix-up with another form of monitoring that already were established in the school system, the control of teachers every four to five years by the school inspection.

(D9) [...] it came as another control [of teachers] but it isn't, we drafted it differently. We make an evaluation of the school and the school processes, not of the teachers' ability or competence, which is the work of the school inspection. That made our life difficult because many said that here comes still another one to check us [...] the term inspection is problematic so that is why our work is continuously called external evaluation [...].

The quality work by QA (Q4) has no end, that is, schools are told to establish a process which is supposed to become a heuristic working model that has no finish, in time as well as in achieved levels, there are always something more to be done. This puts a pressure on the school management, the teachers and students while every domain can be targeted for improvement or development. This transfer of responsibility is in line with NPM, responsibility transferred to the level with no or just exceptional influence on the directives that guide the quality. In contrast to the level in control of the know-how, EB, with a (now) well established organizational structure – institutes, channels of influence, status of importance etc. – that control the interpretation of its research, of the quality.

However the interpretation of the term quality will always be inexplicit and broad, in line with QCE because quantitative educational research cannot specify specific conditions, it can only make broad recommendations, based as it is on *en masse*, something that has its corresponding impact on those who participate in the game.

(D9) About the conditions for our work: The Kultus ministry [have stated that] we are not allowed to publish results and that was a disappointment for us because such documentation is part of our work, which is one part of it. [Secondly] it would also justify our work [...]. The third, I mean the interesting part of the PISA results were that we could compare ourselves, that it made research possible, that it made comparison possible. Because of [our restrictions] these aspects are lost.

The rule of the game is set by the Kultus ministry which does not want comparisons within the Bavarian education, partly because published results could be misinterpreted. Education policy has become a most delicate matter, ` (D9) *Therefore no ranking, not under any circumstances*`. This reluctance is a visual trend in German education policy that includes the prohibition to use test results for comparison of Länder. Since 2009 the educational standards of the KMK form the basis for comparisons, not PISA-E. Instead the German Länder comparisons (*Ländervergleich*) were to be based on the work of IQB, which work out and monitor that the tests meet the requirements of the standards. Hence, the same community of research that was responsible for the (enlarged) PISA-E has become in charge of the Länder comparison. The test results are archived by the research data centre (FDZ) at IQB and besides the official published Länder comparisons by IQB (IQB, no date/b), only approved research projects can obtain authorized access to measurements data, though with the limitation not to publish comparisons between the Länder. Breaking the rule means to risk a penalty⁶¹.

5.6 The Institute for Educational Quality Improvement (IQB)⁶²

The Institute for Educational Quality Improvement (IQB) is an academic institute that supports the 16 states (Länder) of the Federal Republic of Germany in improving and assuring the quality of its educational system. The IQB's work is founded on the educational standards adopted by [KMK] of the Länder in the Federal Republic of Germany [...]. It helps the states put the standards, which are intended to serve as the main framework guiding all actors in the education system, into practice (*implementation*). The IQB is also one of the leading institutes actively engaged in empirical educational research in Germany [EB]. (*IQB, no date/c*).

IQB is an institute that is supposed to support Germany's Länder in their work to continuously develop and assure the quality of the school system. It started as a pilot project financed by the KMK, ` *The Institute for quality development in education (IQB) is a scientific institution of all 16 States of the Federal Republic of Germany at the Humboldt University of Berlin and stands for the joint activities of the countries in the area of quality assurance and development in the secondary school system* (KMK, 2015b). It was founded because the introduction of educational standards by the KMK in 2003/04 created a need to, ` [...] *specify and develop the 2003 and 2004 by the KMK adopted transnational educational standards, and on their basis develop tasks to monitor and safeguard the achievement of the, in the standards, formulated competence expectations* ` (ibid). It is about the practice work of standards, the validation of test tasks as well as development of tasks for teaching, in short, the implementation of a norm.

In Germany, the numerical increase of international and national as well as *Länder* specific tests has been substantial⁶³ and the start of the expansion was the announcement of the TIMSS 1995 survey

⁶¹ The penalty can – theoretically- – be high fines and even imprisonment

⁶² *Institut zur Qualitätsentwicklung im Bildungswesen.*

and PISA 2000 study. Furthermore, in all Länder educational monitoring institutes have become established, that is, quality institutions that monitor the (said) education quality. Taken together, the amount of generated (quantitative) data is massive and increasing, most of which⁶⁴ is documented and archived by the research data centre FDZ⁶⁵ at IQB. The purpose has been to create opportunities to re- and secondary analysis by researchers but the actual probability of access to the data is a controversial topic. Proponents of IQBs control say that many applications do not meet the required criteria whereas opponents say it is being misused, that is, even approved research projects are not allowed to use the data for any comparison of the Länder (cf. KMK, 2012/2014).

5.6.1 IQB (D10)

IQB was from the start 2004 financed by the KMK but organized as a so called AN-Institut at the Humboldt University in Berlin (Humboldt-Universität zu Berlin, 2004). AN-Institut is a German designation which means that a particular line of research or training becomes linked to a University and managed by one of its members. For IQB Olaf Köller, professor in EB, became the first director of the institute (2004-2009). Professor Köller had a background in psychology, had been a co-worker at the IPN and co-worker and research director at the Max-Planck Institute under Professor Baumert who also became executive board member of IQB. In the plenary session of KMKs meeting in Munich 2010, the chair, also the Bavarian state minister, Ludwig Spanlae thanked the former director of IQB, professor Köller, for having laid the foundation for IQBs work, *‘In the federal system of the Federal Republic of Germany, the IQB has performed indispensable tasks for the Länder in the field of quality development in education. The Institute has received a high national and international scientific reputation in a short time’* (KMK, 2010b).

D10 explained that an AN-Institut is of mutual benefit for both parts, the Humboldt University can gain from the strong market value of IQB and receive positions for PhDs as well as entrance to the specific research field of IQB, the know-how of EB. *‘(D10) [...] IQB is an AN-institute, that is, it does not belong to the structure here, Humboldt University, instead we have the facilities that we use but [the] management, organization and financing [is] completely divided from the Humboldt University. [Though] it was different, it has changed since early 2014’*. What happened was that the Berlin court of Auditors in 2012 declared that the agreement between the Humboldt University and the IQB were insufficient and had to be changed and clarified because it caused the University unnecessary risks and costs, and thereby the State of Berlin. For example, co-workers of IQB were on the payroll of the Humboldt University, the over-head cost for research project funded externally did not apply for IQB projects and premises for IQB rented by the Land Berlin was paid by the Humboldt University (Abgeordnetenhaus Berlin, 2013; Der Tagesspiegel, 2012). IQB have grown both in number of assigned tasks as well as number of employees, reflected in the increase of its budget, from about 2,5 million euro in 2004 to about 5 million euro 2014.

(D10) At the beginning [IQB] was very small, and in between it has [...] we have an extensive list of co-workers. Yes we began with one director, that was Olaf Köller, he is now the director at IPN, and in those days we had four to five PhD students and about four colleagues, that is, teachers from the Länder that coordinated the project and some post docs [...] it is much bigger now but we also have many more tasks.

The main task for IQB was from the start to transfer the educational standards into practical tests as valid and reliable as possible, and thereby implementing a norm setting procedure: *‘(D 10) [Our work was] to transform the educational standards into tests and then we began with a pilot study to check the tests, and then we conducted norm setting studies, and in 2009 it was time for the first Länder*

⁶³Example of ILSA: PISA/TIMSS/PIRLS/DESI; National tests: VERA/Ländervergleich; Länder specific tests in Bayern: Orientierungsarbeiten/Jahrgangsstufenarbeiten/Schulaufgaben etc.

⁶⁴ This includes all international tests data as well as national but not all state specific tests data.

⁶⁵ Forschungsdatenzentrum.

comparison study. The *Länder* comparison study is thought to verify that the competence levels in the German education system are in accordance with the adopted Bildungsstandards. IQB have for example developed the competence levels for subjects and school grades in the VERgleichsArbeiten (VERA/Comparative Work). This test was first launched in 2004 by seven Länder as a primary school comparison study. The centre for empirical pedagogical research at the University of Koblenz-Landau was responsible for the study and one of the main characters involved was Professor Andreas Helmke, one of the persons behind the manual and guidance for the conceptualizing of quality work, and also the developer of the first German large-scale assessment study, the MARKUS study (VERA, no date). VERA is now conducted in grade 3 (VERA 3) and grade 8 (VERA 8). The difference between the Länder comparison and VERA is that the first monitor and promote the standards while the second is said to promote the individual school and its education (KMK, 2015c; KMK, 2015d). Another difference is that in the Länder comparison just a sample of schools is included while in the VERA tests all 3rd and 8th grades in all schools are included (IQB, no date/d). The schools receive their results but are not obliged to make them public, the reason is said to be the potential (public) pressure that could be an obstacle to the intended development (KMK, 2015e).

The tests and the *Testeris* (the excessive testing) have been severely critiqued. VERA in specific have been in focus but the targeted weaknesses concerns other tests as well: poorly managed regarding the integration and lack of follow-up, deficits in design concerning too complicated test language, the construction of test questions, the costs for the partaking schools and municipalities, the frequency, the increase of workload for teachers and the lack of significance for the praxis (cf. Wittmann, 2010; Zweites positionspapier des EMSE-Netzwerkes, 2008; WAZ, 2015). It could be said that a lot of the criticism deals with a gap between two un-compatible spaces; EB, *developer* of collective tests (mass tests) and the practitioners, developer of individuals (students). In that sense, the principal cause for the critique and struggle is EBs implied and embedded position as *the* problem and efficiency solving device in education, in praxis. The increasing critique of EBs role in education has had the effect of a gradually shift of position and attitude by EB representatives. For example One of the two directors of IQB, Professor Hans Anand Pant, has implied that QCE is better suited, *'[...] rather for the detection and description of context conditions that influence the effects of measures and programs, and less for strict proof what work best [in education]* (Pant, 2014, abstract). In the light of how critique was dealt with previously – broad sweeping statements or just ignored – it gives the impression of a change of course (e.g. Hopmann, Brinek and Retzel, 2007, pp. 13-15).

5.6.2 Storing data for ... ?

The Research Data Centre (FDZ) at the IQB) archives and documents data sets resulting from national and international assessment studies (such as DESI, PIRLS, PISA). Moreover, the FDZ makes these data sets available for re- and secondary analysis. Members of the scientific community can apply for access to the data sets archived at the FDZ. (IQB, no date/e).

FDZ is at the centre of an ever increasing amount of (EB) data.

(D10) [...] you can say that [research] become listed and kept here. They are free at disposals for researcher who apply for them by explaining the scientific purpose of the study, which then normally is examined and the wanted data are given [...]. The data are free at disposal with one restriction.

This special restriction concern data that can be used to compare Länder which demand an extra authorization.

(D10) We are not allowed to publish them because it would make it impossible to validate trends in these recurring tests. They must be kept secret in order to show trends. If the tasks were known it would be possible for the schools to work on the type of questions in order to improve their results.

One aspect of this is that the same was (and is) said about the PISA study, that the objective of the study made it impossible to grant full access to the complete outlay of the tests, which made it impossible to check or reexamine the validity and reliability of the tests⁶⁶ (cf. OECD, no date/b). The accumulated data at FDZ is huge and ever growing, and can in that way be seen as a successful indicator for educational research (in particular EB) accessed or not, but the transformation of the data into manageable practice guidance for education is still a deficit after more than 10 years.

(D10) Yes, I think there is obvious a huge amount of quantitative data that show certain levels of competences or describes features of students' indicators by teachers [...]. The problem is that it is not possible to draw any conclusions of best practice from these results, that is, no consequences of how change should come about, what works best [...] all in all, no causality [...] that is hard to achieve in these major school studies, that is the problem with these longitudinal studies [...] in that way there is a certain frustration partly also in the educational administration of the political level because they want guidelines for plans of actions.

At the same time, as the practice of empirical studies is making its way into every research field, the hostesses of this educational research, schools, are burdened.

(D10) The other aspect is that schools are indeed really burdened with this whole data collecting. There are schools that are targeted with a mass of research, constantly more with researchers coming with requirements, performance tests and wanting to perform surveys and interviews with teachers, and this comes on top of the ordinary school day, that is, the staff must manage more and sometimes the question for them is why? It's the same procedure again and for what good? There is no practical outcome for us`.

One discernible effect is that the number of schools that do not want any research is increasing and among researchers this problem is often seen as a cause of the hierarchical *edictum* from the start, the demand and decision for empirical research, in which the professionals and teachers never was involved in (e.g. Tenorth, 2011). Furthermore, that decision was brought out as in line with a legitimate public claim for improvements of the competence levels, to make German students more successful competitors in the race for excellence. A broad range of interested and influential actors – such as experts, politicians and foundations – excelled in demands for reforms in education and acted as dedicated but impartial facilitator of what people (the silent majority) wanted and the *nation* demanded. The normal and slow path of negotiations between the different levels was circumvented by the strong pressure to act fast, to achieve rapid change. The contradiction was that even if the public were used as the base for action, their counterparts in schools, the lower level of the education field, were not really involved in the process (cf. Champagne, 2005, pp. 118-119).

5.6.3 Replaced Bildung and a `place in the sun`

One of the major issues in the educational debate in Germany has been how the former honored concept of Bildung has changed.

(D10) Yes, that is really debated [...] the classical pedagogical goals are based on a strong tradition in Germany. On the other side, this development of standards do affect this, there are some interesting discussions. Personally, I come from the EB where I belong, but I would say that the standards are competence oriented [...] it is about building special competences [...] an overall knowledge possible to use in new contexts [...] creating a flexible individual knowledge.

Creativity and problem-solving ability (competence) based on real life is a concept from the OECD and (is said to be) a part of the testing of ILSA, such as PISA; *This fifth volume of PISA 2012 [...] provides the rationale for assessing problem-solving skills and describes performance within and*

⁶⁶ The first test, PISA 2000, published technical report was extensive but this changed already in the following study, PISA 2003.

across the 44 countries and economies [...]. The volume also explores the role of education in fostering problem-solving skills (OECD, no date/c). Since standards are the yardstick that enables such comparative measurements, they have in Germany been targeted for breaking down the former ideal of Bildung, for inferring a cost-benefit perspective and increasing the work load of the teachers (cf. Hessische Kultusministerium, 2006, pp. 6-8). `(D10) Yes, well I think standards and standardization are now days regarded as something negative, in each case among teachers`.

In the beginning of the 90s the field of EB, as in educational performance measurement, was a peripheral, circumscribed and restricted activity on the outskirts of Erziehungswissenschaft (Aljets, 2014, p. 25). In just a decade the scene was totally transformed. The small and non-university institutions that together formed Germany's know-how of QCE were at the center of a massive funding and promotion activity where EB centers and chairs were created and established. One outcome was that the small circles of established researchers in EB 1990 a decade later were holding numerous and key positions in the field of educational research⁶⁷.

(D10) Yes, well it is interesting, it is above all the question how come? I mean it is correct, you have these excellent personalities, those that are very linked and connected to EB, [Professor] Baumert of course, and the question is if this trend [with EB] has developed, spread, or if it has become a sort of elite education? [...] I think that the research field has widened, become much stronger also at universities where centers are created. [...] I think it's just natural that the major studies, and the control of these, such as PISA, PIRLS and so on, are strongly linked to certain institutes and persons [at these institutes]. [...] I think that this field of EB, the know-how, have widened and spread to other disciplines, also to the didactics, subject specific didactics [...]. It [EB] plays a completely different role in education now days [...].

From the start of Germany's (abrupt) reentrance in the partaking of ILSA, the methods and approaches were met with critique but any form of constructive debate was absent. However, this has changed in the last few years. The numbers of actors involved in the criticism have grown, and representatives of EB have become much more cautious about what the major educational/school studies can achieve. Nevertheless, the pressure to demonstrate efficiency makes it necessary to adapt, `(D10) Erziehungswissenschaft and also comparative Erziehungswissenschaft orients in a greater extent towards qualitative empirical research`. Not only natural science as mathematic but also the Humanities have been affected, the *empirization* has become the standard yard stick and evaluation the working form. Those research fields that do not adapt risk their autonomy.

On the base of the growing empirical tendency in [German] science and an ever increasing politico-societal demand for legitimization, the Literature didactic have decided to self evaluate the efficiency of its working forms of mediation in use, in order not to be externally evaluated. (Nickel-Bacon, 2006, p. 95).

One educational researcher that from the start was critical was Professor Tomas Jahnke, chair of the math didactics at the University of Potsdam. Over the years, Jahnke has written several critical essays. One especially straight forward and critical of EB practice in math didactics, called *zur Kritik empirischer Unvernunft*, was written in 2009 but not published until 2015, an incident that bears the mark of a changed *Zeitgeist* in which EBs positional stand becomes possible to challenge (Jahnke, 2015, pp. 121-137).

⁶⁷ And also numerous awards.

5.7 The practitioners, school level

There are two growing problem areas in the Länder concerning teachers: an increased medium age and a growing lack of educated teachers in certain subjects and school arts. The age structure of teachers shows that 45-55 % of the teachers are more than 50 years of age. In the so called *MINT* subjects – Mathematic, Informatics, Natural science and Technology – the shortage of teachers in secondary and high school is already prevalent (Klemm, 2014, pp 2-4: KMK, 2015f, pp. 4-5). In Bavaria the prognosis for the future differs depending on school form. Primary school has already a deficit of teachers as also *Mittelschule*, *Förderschule* (special schools) and the vocational trait. The deficit is generally increasing for the latter, while there will be a surplus of educated teachers for primary school in the coming years. The exceptions are the *Gymnasium* and the *Realschule* with a surplus of educated teachers predicted to increase (Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, 2015b).

In the educational debate, representatives and individuals from different levels express and promote their opinions, though one level is on the whole invisible: the practitioners. Instead the dominant perspective is to problematize teachers' skill and professional knowledge, often pin-pointed as the cause of deficiencies in education. Statements are often heard from central actors on adjusting this predicament by improving the standard of teachers. The origin of this education 'miracle cure' – the improvement of teachers' competencies – is often referred to as the OECD report '*Attracting, Developing and Retaining Effective Teachers-Final Report: Teachers Matter*' (OECD, 2005; cf. Stifterverband für die Deutsche Wissenschaft, no date). The unions give a slight different picture of the problems in schools. VBE⁶⁸, an umbrella organization for unions in education, state in their *Mainauer Erklärung* that Bund and Länder must cooperate and provide sufficient means for education. They target the large size of classes, the high amount of compulsory teaching hours, the decreased amount and possibilities of diversified and special education (VBE, 2015). Furthermore, they are specifically critical of the amount of assessments.

The VBE does not identify any political will to really invest in a genuine education republic. The constant measurement in education through performance tests and work of comparisons [VERA] has not in a single case improved the conditions for learning and teaching at schools. (ibid.)

5.7.1 Teacher at a Gymnasium (D11). Changed curricula

Since the turn of the century, the demand for higher education has increased remarkably in Germany, official forums like KMK and BMBF have announced the importance of education for Germany's possibilities to compete in a global market. Foundations and private stakeholders as well as different interest and parental organizations, have increasingly put pressure on politicians to enhance the quantity and quality of higher education. This concept or idea is reflected in the increasing number of students in Gymnasium, the traditional highest level of secondary school in Germany. In fifteen of the sixteen Länder⁶⁹, the number of students in Gymnasium has increased, for some the increase has been dramatic. In Sachsen-Anhalt, the amount of students in Gymnasium was 23,9 % in 2000/01, ten years later it was 40,4 %, in Niedersachsen it was 20 % 2000/01 up to 35,8 % in 2010/11. The most populated Länder show a lesser but still clear growth in those years. For Germany overall the rise was about 4 %. Bavaria was one of the Länder with least increase, from 31,6 % to 32,4 % 2010/11 (Statistisches Bundesamt, 2012, p. 12).

(D11) I am a teacher of mathematics, physics and Italian at a gymnasium [...] my former principal wanted me to have Italian because we, by then, took part in a school experiment with other European Gymnasiums [...]. Tests in Bavarian education has a long standing

⁶⁸ Together with its sister organizations GÖD-APS in Austria and LCH in Switzerland.

⁶⁹ The exception is Mecklenburg-Vorpommern.

tradition but since TIMSS 1995 the number and different forms have increased significantly: TIMSS [since 1995], PISA [2000], PIRLS [2001], Orientierungsarbeiten [2003/2004], VERA [2007/08], Ländervergleich [2008/09]. [...] my impression was that the new curriculum that were developed after the first PISA test, when the Gymnasium were shortened to 8 instead of 9 years in Bavaria and it was said [the content] had not changed, my impression was that it was very much oriented at increasing the results in PISA. For example in 5th grade you suddenly were supposed to start the teaching on diagrams, that wasn't part of the former curricula but it was a part of the PISA test [...]

That ILSA have changed the curricula's of Germany's Länder is not a debatable subject, it has. On an overall level in form of the introduction of specific measurable standards, the output steering and the emphasizing of the term quality but also in the form of introduced term like competence (cf. KMK, 2006; Klieme and Tippelt, 2008, pp. 7-8; Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2008, pp. 7-9). The run for higher and most qualified education affect the teaching, with a more diversified and heterogenic audience, the teacher must adjust the lessons and level of teaching. Although the educational transition – from primary to secondary school – is decided by the teachers in 4th grade, the parents' wish have become an influential factor. Hence the changed rules for the transit in Bavaria (in 2009) with a strengthened parental position (Bayerische Staatsregierung, 2010). `(D11) [...] yes, I have the impression that primary school teachers let the children pass rather than to face problems with the parents. [...] we receive these children and for the moment I have a fifth class with 30 students, and I would say that 12 of them fit in a Gymnasium`. With an increased amount of students, the numbers who have difficulties in the course of their studies multiply.

(D11) [Difficulties in education] has more to do with where the students come from, what you have. I mean the people consider and take into account which Gymnasium the children should apply for. Some say not on that [Gymnasium] because of the surroundings which mean that the child will sit together with `that` kind of children and it will lower the learning outcome. It is a bit socially characterized [...].

5.7.2 Educational stress

The first selection of school art after primary school [4th grade] is no final decision on the academic career of the child. The Bavarian educational system opens up an individual education path for each student [...]. Each student is given the ability to adjust the education to new realities and objectives periodically. In that way, inadequate- or excessive demands are avoided. All the schools in Bavaria offer several ways to achieve graduation. As a general rule: when a graduation is reached, the path is open to the next higher school goal. (Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, no date/c). (Authors translation).

The possibility for a student in Germany to change educational path – decided in the 4th grade⁷⁰, `der Übergang` – has been much criticized, that is, the possibility have existed but in reality the possibilities have been small. In 1998 the specifications regarding accumulation and transferring of merits into higher education were regulated in law and in 1999, the Bund and the Länder signed the Bologna declaration, thereby expressing the intention to incur a merit based system similar to the ECTS⁷¹. It was done in order to enhance the mobility of students and promote the possibility of lifelong learning (HRK, 2003; Sekretariat, KMK, 2008). In Germany research has ever since been trying to establish the progress of permeability but it is still a deficit, though with the exception for high-performance students. The number of students that change their educational path to a lower level is still twice as many as those who are able to change to a higher level. Bavaria is one of the Länder with the lowest permeability and high inequality concerning integration of students with special

⁷⁰ The common practice in Germany has been that primary school consists of grade 1-4' but may differ depending on the Land and school form.

⁷¹ European Credit Transfer and Accumulation System.

needs, but in the top when it comes to promotion of performance⁷² (Bertelsmann Stiftung, 2014, pp. 9-11, 36).

(D11) I hear it mainly from primary school teachers in 4th grade, because then it's time for the transition [choosing educational path]. In that respect it was different, before we had this central test for the transition and it was the same for all. That was much better for the teachers than nowadays when the teachers decide the grades over the year [...] of course there are parents that pressure the teachers because they want their children to move on to a Gymnasium, unconditionally.

Year	Percentage of students in 4 th and 5 th grade going to Gymnasium	Percentage of students in 4 th and 5 th grade going to Realschule
2004/05	36.5 %	25.8 %
2008/09	39.6 %	28.3 %
2012/13	40.5 %	30.3 %

Table 4. Percentage of Bavarian students in 4th and 5th grade going to Gymnasium, Realschule and Mittelschule. (Source: Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst. 2014a.)

Already in 2008/2009 about 4 out of 10 students went to the Gymnasium but four years later, around 20 % had switched to another secondary school form (see table 3, page 26). Gymnasium was as late as in the mid 60s considered as an elite education with just about 15 % of the students from 4th grade, *(D11) Obviously the number has increased compared to my days [...] and you hear all the time that the Mittelschule, the former Hauptschule, gets emptier and emptier [...]*. The reform in Bavaria of Hauptschule into Mittelschule was stated as not only being a change of names but a real change into a more qualified secondary school form with more and diversified traits. However, the critique has been substantial and especially targeted the undersized number of teachers relative to the amount and groups of pupils and needs. Mittelschule had for example the highest percentage of students with migrant background, almost 25 %, in 2011 (BLLV, no date).

The prevailing notion of the importance of qualified education does not only affect parents, the children are also stressed to compete successfully irrespective of background, socio economic aspects and habitus.

(D11) I think it's always sad that parents unconditionally want their children to go to Gymnasium, no matter if they can handle it, and then they don't and maybe fail again [...] then they have lost so much time and really feel frustrated and have hours of afternoon teaching. Instead they would have attended a Realschule in the first place, with the option to change path later.

Gymnasium that once was – and for many parents still are – a symbolic marker, an entrance pass to an exclusive selected group, have deteriorated or devalued in pace with the increased number of students. Since about 40 % of the students in 4th grade are going to a Gymnasium the exclusivity is no more in contrast to the study requirements that remain high, resulting in many drop outs.

⁷² Based on the performance of mathematic in 9th grade compared to the standards.

(D11) But many of the parents say they want the best for their children and that has to be Gymnasium, they are quite persistent and don't listen to any advice, the primary school teachers complain about that. It's a shame, The parents think about Gymnasium as it was in their time when a change of path was very difficult [...] a primary school teacher told me that nine out of ten [that change path] goes from Gymnasium to Realschule, only one of ten from Realschule to Gymnasium.

The pressure of success in education has intensified in an already competitive educational discourse where the job market is the prime beneficiary and locus (cf. The Bavarian Ministry of Economic Affairs and Media, Energy and Technology, 2014, pp. 80-86). In the last two decades the massive increased proportion of students in the more qualified track in secondary education follow a global trend (OECD, 2014, pp. 3, 13). Though at the same time, the student corpus consists of an increasingly heterogeneous collection of individuals of different ages and with different backgrounds, Instead of the former saying *'this is not for us'* the maxim now is *'for all that can compete'*. Parents have gained an increased influence on the choice of school and data is used to compare schools (cf. Bildungsweb, 2015). However, some children are better equipped to compete successfully depending on the familiarity with the educational context and former training, that is, training as a *mechanical* preparation but also as an implicit inferred know-how of the discourse: to know how to act and what to do, and when. Students from socioeconomic higher strata where parents and/or surroundings have symbolic, social and economic capital will always be better predisposed for educational success. But not only the – trained and transferred – *properties* of the individual increase the chances; the *feel for the game* does also include networks and relations to ease the way. Many of the new students in higher education will not have these advantage, the outcome is a differentiation within secondary education with a hierarchy of schools where only exceptionally *'gifted'* students from lower strata will compete successfully, which will serve the legitimacy of status quo in two interconnected ways: first by the *illusio* that ascension is possible and second, by passing the eye of the needle, the correct habitus have been incorporated; hence, these individuals will be the system's main conservators (Bourdieu, 1990a, pp. 66-68, 111-113; Bourdieu, 1996, pp. 116-120).

5.7.3 An alien assessment regime and subjective evaluations

Countries that partake in ILSA have agreed to an educational assessment regime based on QCE. This regimentation is a necessity; it enables the comparison of educational systems. In that way, ILSA has brought in and changed the assessment discourse in Germany. Before the TIMSS survey 1995, large scale testing of specific age cohorts in school was (more or less) nonexistent in Germany (e.g. Aljets, 2014, pp. 22-25). Though it was not only the amount of students involved in the assessments that ILSA changed, also the testing methods were unfamiliar to the German and Bavarian educational system.

(D11) Between Germany and Bavaria there is a major difference. Well, I was in a US University and taught mathematics. [In secondary education] they can choose mathematic⁷³ or, if they prefer, some other subject [...] and then they come to a University and want to study biology for example, then they have to study mathematic and at the University at a level equivalent to 8th grade in Bavaria [...] and they have a lot of these aptness tests [...]. I had to do these multiple choice tests for mathematic, like the Kangaroo test. Of course these take no time to correct but you need another kind of strategy than in the normal school tests [in Germany].

The Kangaroo test was an early premonition of what was to come. Originally from France, this multiple test in mathematics has developed into a worldwide contest. In June 1994, in Strasbourg, at the European Council, the General Assembly of the delegates of 10 European countries decided the creation of the "Kangaroo without Frontiers" Association (Kangourou sans Frontières, no date). The test was one of the earliest promoters of multiple tests, adopted by ILSA. Germany took part in 1995

⁷³ The possibility to choose subjects and levels is restricted in secondary education in Germany.

for the first time with 3 schools, in 2015 the number was over 10 000 (Känguru der Mathematik, no date). The German students' lack of experience of multiple choice tests was one of the main points of the early criticism of PISA and was said to explain part of the bad results, as also the way the questions were formulated, '(D11) Yes, I can see that the formulation does effect the results. Depending on how I formulate a question, I will receive different answers'.

In the wake of assessments evaluations followed as part of the monitoring of the educational systems in the German Länder and the qualitative assurance of (cf. KMK, 2015g). Monitoring in education is perceived as a legitimate and *prima facie* linear operation that can be accomplished through periodic evaluations of quality. In reality, it has proved difficult – if not impossible – to achieve. The main predicament is how to define quality, but also how evaluations can be performed in an objective and accurate mode (cf. Bernt, Remme and Langer, 2014, pp. 267-274). Without either dimension, the evaluation will not have the consent, assent or acceptance of the practitioners and the outcome will only be '*Monitoring at a Glance*'.

(D11) Yes, I have been part of two evaluations. Should I be honest? I was on two different schools [...] My old principal could really be convincing even if there wasn't much behind and he succeeded to get the evaluation results he wanted [...] he managed to disguise things very well and in that way the evaluation did not benefit or improve the school [...] the evaluation became a stepping stone for the principal to reach a higher level and he was successful. [...] the internal model [for school evaluation], what is discussed is quite questionable. They want to establish a hierarchical structure in the schools. Today we have the principal and then the rest who are almost on the same level [...]. Now the idea is to institute a middle level for the most skilled teachers who are supposed to guide or advice the rest of the teachers. I am most critical since the outcome will be a competitive situation for the teachers and less team work because everyone will climb on the other to reach a higher position.

The inflicted concept of lead teachers or first teachers to officially reward and promote best practice and role models, can also be seen as one part of a concept to divide groups and reduce teachers' autonomy by building levels and promote competitiveness (cf. Save Our Schools NZ, 2014; Sundell, 2014).

(D11) And I think this idea of competition is harmful, I don't think it endorses better education. That is, I think from the students' perspective it would be better if this hierarchical structure don't come, in spite of being a reality in business, in the economy [...] cooperation is so much more important, crucial that it exist between colleagues so that no one is excluded.

However, leveling and competition works clockwise; leveling promotes competition, competition promotes leveling. In other words, creating structures of monitoring, hierarchical structures where each level has its value (capital) decide the division of monitor or monitored. To monitor is to coordinate the behavior in such a way as to benefit the appropriate or desired behavior. Hence, evaluations in education give rise to leveling; students are subordinated to the assessment practices of practitioners, practitioners are subordinated to internal school evaluations, schools are subordinated to external evaluations, state quality institutes are subordinated to central evaluations bodies – may it be federal or private. The monitoring process involves a struggle of positions where judgments of evaluations tend to be more explicit at lower levels and less explicit the higher level it concerns. The process is part of a differentiation where those with more valued capital – both in terms of what kind of capital and its exchange rate – can gain ground and autonomy, for example elite educational sites and schools with a special kind of clientele (cf. Bourdieu and Wacquant, 1992, pp. 98-109; Bourdieu, 1984b, p. 125).

5.8 Humanistic Gymnasium

In Germany, the Humanistic Gymnasium is traditionally the most prestigious form of Gymnasium where the classical languages Greek and Latin are valued. The emphasis is on European cultural values as expressed in classic literature.

The Humanistic Gymnasium opens up the possibility for the students with the profile subject Greek in particular, to acquire comprehensive knowledge about the foundations and Bildung of our European culture and cultural competence, in its original language (Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, 2013a).

5.8.1 Principal (D12) and teacher (D13). Transformation of awareness

The `rush` for education has not affected the Humanistic Gymnasiums to the same extent as the other forms of Gymnasium but the partaking in ILSA has had its consequences.

(D12) Well it has affected a lot, quite surely. In specific I think the main effect is basically a transformation of awareness. That in the heads of colleagues, parents, and students the idea had to develop that it [the partaking in ILSA] was a necessity, to orientate towards international standards because comparison, not just within Germany but also European and globally, was essential. It meant that even colleagues had to be persuaded. [...]. [The] biggest effect was in the field of test culture, I think, the test format. In that respect it really is an improvement [...] and here the normal procedure in the testing of students [is that] we didn't use that kind of test format. They also have to write essays and these you can't evaluate [that] through those kind of test measurements methods. [...]. We have though benefitted a lot from those kinds of tests concerning the test culture [...].

The field of education has always had problem areas, the focus of which has alternated depending on the Zeitgeist. The distinctive feature of ILSA was that shortcomings were defined by a perception that education, *in itself*, was a problem. This all-encompassing (non)definition became a predicament; ILSA did not work as an explanatory device that could guide action. Instead testing multiplied, testing inscribed by QCE as the sought and dictated prescription, and by that the introduction of an *alien* test discourse (cf. Boller, 2009, pp. 31-34). Even if approved, change of practice takes time because acceptance and habits are not easily transformed, not least when fundamental elements are concerned. Grading is an essential part of a teacher's work, primarily based on (traditional) assessment tests which mean that the introduction and implementation of an alien test culture takes time, or becomes relegated to a complementary and subsidiary role, depending on the recipient's amount of capital.

(D12) Yes and it took some time before everyone understood that these kinds of tests were something different, that they tested something else compared to how the normal tests we had in Bavaria, with our schoolwork or home work. For example, we have an English test in grade 10 [...] it is objective and we fulfill all test criteria [...] but it involves a different set of competence than in our classical comprehensive test and it is important to look at as a part of a complementary element. It is also true that the levels of requirement in these [new] tests are frequently lower than those we usually have when we test school assignments. In the latter case, we ask the students for a limited part of the new material or subject [...] and then we can demand more from the students than a test based on a whole years learning [...].

The divisional concept, dividing a complex task in smaller units, is the prime value of standardization, in education it is the *prima facie*, the `first step`. Hence, the traditional educational valued ability to inscribe the complexity of the whole in the parts is reverse to the competence stated by ILSA, where the parts are the whole. The result is that the complexity is, literally, lost. The former is relational knowledge, the interdependence of parts, the latter structural knowledge, the independency of parts, in line with QCE where specific parts are transformed into variables, said to be corresponding and quantifiable (e.g. Hörmann, 2005, pp. 41-43). These `evidence` based assessments, in the

aftermath of TIMSS 1995 and PISA 2000, was one of the most significant conversions of German education according to official sources (Erdsiek-Rave, 2006, pp. 12-13). However, even if reforms such as the initiating of educational standards and national and state tests, in the securing of `evidence`, did affect the procedures – or rituals as in a repeated collective pattern of behavior – in education, the `new` did not always replace the `old`.

(D12) And that it [the new test format] is additional and complementary. Well it's often the case that the parents say, my child was great on the [new] test, and it had a one or at least a two, how come it had a three or four on the [traditional] school work? Then you also need to explain that [the old and new] are two different test cultures and test formats and so on. You still really need to explain that and spend a lot of work to persuade, particularly the parents.

(D13) For the parents it is quite new, really. Also this difference, that it is a different test format, that a different set of competences are sought. It's hard to grip for the parents because as a rule, most of them once went to [a Humanistic] Gymnasium themselves and this shift is really hard to understand for them. [...] the tasks are basically oriented towards competencies, referred to as real-life, reality [...].

Measurements of the competence defined as the ability (creativity) to solve and apply knowledge to `real life problems`, is a part of the OECD concept and methodology of ILSA (OECD, no date/d). Another interpretation is that it is part of the economizing of education; part of the human capital theology where a price tag can be put on applicable knowledge (e.g. Krautz, 2007, pp. 81-87). In secondary education, Gymnasium is the form that holds the highest position and promises of future success. Though, with the increased amount of students, the struggle of positions between the various forms of Gymnasiums has intensified but the struggle for the Humanistic path is not a struggle of *quantity* (number of students) but *quality* or safe guarding (cf. Bourdieu, 1996, pp. 84-90). The Humanistic Gymnasiums with their illusory discourse of upholding important cultural values – the knowledge of old languages and valued ideas – and thus to hold on to traditions, are well equipped to defend their dominant positional stand, exemplified in the socio economic heritage of their students and the relative ease new procedures are incorporated and subordinated.

(D13) Though it would have been a loss, sorry, a loss if we only had to work with these [new] kind of tests, this test format. That would really have been an unambiguous loss. I teach Literature in German where the students write themselves and that [skill] is not possible to test with [the new] kind of test format. [...] instead that test format can be used to expand and complement such a complex and partial creative task.

5.8.2 The `rush` for qualified education

In order to secure the human resources of the knowledge society, it is regarded as necessary to increase the number of students entitled to higher education and University graduates. (Statistischen Bundesamt, 2014, p. 32).

The cost in Germany for education, research and science in 2000 was 115.3 bn euro, in 2012 it was 247.4 bn euro, or 9.3 % of gdp (Avenarius, Ditton, Döbert, Klemm, Klieme, Rürup, Tenorth, Weishaupt and Weiss, 2003, p. 38; Statistisches Bundesamt, 2014, p. 6). This steep raise was accompanied by the increased number of students in secondary school where Gymnasium was the most preferred educational path. In the school year of 2012/13 in the Länder Mecklenburg-Vorpommern, Sachsen-Anhalt and in the city states Berlin und Hamburg, more than 50 percent of the beginners in 5th grade went to a Gymnasium. The educational track Hauptschule showed at the same time a declining trend (BMBF, 2014, p. 73). It was visual in Bavaria as well with 1/3 of the secondary students in Gymnasium 2013 and an ever lower part at the Mittelschulen⁷⁴, ` (D13) It started in the 70s [...] for Bavaria there are several contributing factors, specific for the Land.`

⁷⁴ See table 5.

(D12) Well, the main reason is of course the tripartite system in Bavaria. That the cause of, the road to, well, how would you describe it, the allowance for higher education, the entrance pass not only to Fachhochschule⁷⁵ but also regular higher education, that there are more ways even if many still believe that going to the Realschule excludes higher education. That has really changed in the last few years, it's really a very positive development, even if you start in the Mittelschule or Hauptschule you have the chance to go on studying, it is not only dependent on the Abitur⁷⁶.

In 2002 24 % of the graduated in Germany could go on to higher education, in 2012 it was 35 %. The educational track in secondary education that accounted for the bulk of this increase was the Gymnasium (Statistischen Bundesamt, 2014, p. 31). Even though the increase 2000-12 in Bavaria was 7 %, the percentage of students in Bavaria 2012 that took the final exam, the entrance pass to higher education, was only 26.7 %, the lowest percentage of all the 16 Länder (Statistischen Bundesamt, 2014, p. 32; Bayerisches Landesamt für Statistik, 2014). However, the transitional proportion of students to Gymnasium, about 40 % since 2010, is a quota that has met objections from teachers' Trade Unions as too high to guarantee the quality of the education (Bildungsklick, 2015).

(D12) It's in the last 15 years, it's really, how can you say, a democratization process in Gymnasium has taken place. Before, the Gymnasium was really just for the intelligent, that is, the best 10 %, children of academics that had attended a University and then this societal change came, the change of political thought. The Gymnasium became open for, so to say, a wider part of the society, strata that before not even had thought of going to a Gymnasium, or to take the Abitur. [It has not affected us much] we are maybe a bit special, because this is a Humanistic Gymnasium with Latin and old Greek, and the parents that send their children to us, well they already belong to an academic elite, you have to say that [...]. And it's quite strikingly that a sociological study, published recently, found out that parents, especially academic parents, those who have attended a Gymnasium unconditionally send their children to a Gymnasium. This is because of a fear that the children otherwise will undermine their social position if they were sent to a Realschule [...].

The transferred notion of education as the prime predeterminator of future success (spelled secondary path) has resulted in increased stress among students and teachers in primary schools because the result in 4th grade determines the student's options (Reinders, Ehmann, Post and Niemack, 2014, p. 4). The outcome is a more heterogenic mass of students in Gymnasium and a polarization. The amount of middle and low-performers has multiplied resulting in a large number of students that have a hard time coping with the study requirements as also their parents' demands.

(D13) Well, it's a complex issue [stress and burnout]. Of course it has to do, I think, with the level of demand, the requirements in Gymnasium have decreased compared to twenty years ago. Then you have the parents' expectancies that are incredible high [...] today [the students] have to function already in the 5th grade, they must have top scores [...]. But I think it is not because of the school but the surrounding, often from the parents, what they want for their children. [...] It occurs the whole time that we as teachers have to tell the parents to relax, that their children are doing well and that it, usually, works for them.

5.8.3 Quality and evaluation

(D12) [...] we have an extern evaluation [in Bavaria], we had it and it is something important. Of course, many complain and say that it cause a lot of extra work and takes time, but this glance from the outside on yourself, I mean we perform internal evaluations but this glance from the outside which has nothing to do with the students [...] to secure the school quality, not the quality of teaching but the school's quality, that is important and positive.

⁷⁵ Fachhochschule is a German tertiary education institution, translated as University of applied sciences.

⁷⁶ The Abitur in Germany is the name for the final exam at secondary education.

The instigating of quality institutes by the Länder was part of a parcel meant to monitor the quality in education. A plausible scenario would therefore have been that these quality agencies as time went by, had tracked down and defined the variables that had the greatest impacts on the quality of education. Instead, quality themes and quality aspects have changed under the influence from various actors, especially through changed political directives and ,at the school level, by the – for the moment – stated definition of quality (cf. Bayerisches Staatsministerium für Unterricht und Kultus, 2005, p 13; *ibid.*, 2010, p 13). What has become a state of fact is that quality in education cannot be primarily measured and targeted collectively but has to be individually modified: not quantitatively based on standardized tests but qualitatively based on the educational sites specific circumstances. In many ways this demarcation follows the line between research (theory) and practice (education) (cf. Harris-Hummert, 2011, pp. 191-194).

(D12) Quality [in school] is, well, basically every school has to make their own definition. It has very much to do with the school's profile. For example, we are a Humanistic Gymnasium and for us the development of each students personality is very important, that they really get Bildung, not just education but Bildung . To be able to see and appreciate beauty, that is aesthetic, artistic education but also education in natural science, to become part of an enkyklios paideia, a community of people with Bildung [...]. But schools have different and various focus, to speak of any general school quality is very difficult. [...] also to speak of any general school quality at Gymnasiums is difficult because every Gymnasium is a microcosm [...]. It is also politically decided and stated in official document that every school is responsible and must make its own decisions [about quality].

If quality in education is based primarily on the qualitative narratives of individuals or groups, personal interpretations of lived realities, there is always the risk of subjectivity. Hence, evaluation systems that is based on this epistemology, has *automatically* inserted trust as a major conceptional device. This in turn means that sanctions are replaced by – more or less – self managed, and guided, improvements. Though, the motive of which does not imply altruism, the main reason can also be to protect a system.

(D12) Principally it is up to the school management to deal with it [the outcome of extern evaluation], if they don't care then there are no final consequences [...]. I mean the extern evaluation does not work as any Ofsted⁷⁷ [...] every school must decide for themselves that we have these deficits and we want to make improvements, or they can say that nothing will be done [...].

5.9 Natural Science and Technical Gymnasium

The Mathematic and Natural Science or Technical Gymnasium is in Germany a term for a Gymnasium which focuses on physics, chemistry, and partly to biology. Informatics (informatics science or informatics technology) is normally an additional subject. Latin, French and English are the specific languages taught. Many educational reforms and changes have occurred within the school form Gymnasium since the turn of the century: introduction (and conversion) of the G8⁷⁸, expanded opportunities for students' individual choices, dramatic increase of the number of students, more transition opportunities and development of focus areas. Some interpret these changes as improvements, creating more and better education possibilities and as signs of a growing equality. Other focus on the increased burden on school administration and teachers by these changes, depending on the more heterogeneous group of students with the consequence of a deteriorated ability (and interest), as well as an increased competition among schools and students (cf. Köller,

⁷⁷ The Office for Standards in Education, Children's Services and Skills (Ofsted) is a non-ministerial department in UK that monitor education. The task is to `inspect and regulate services that care for children and young people, and services providing education and skills for learners of all ages` (Ofsted, 2015). The consequences can be serious for the institution that fails the inspection (cf. The Guardian, 2014).

⁷⁸ The shortening of time from nine to eight years.

2013, pp. 53-58; Hoffmann, 2013, pp. 49-52). Even prominent members of EB have expressed ambivalent doubts of the positive outcome of the reforms (caused by the results of ILSA):

In spite of its [Gymnasium's qualitative] consistency, it has modernized. [...] Here we must ask us, if the Bildung's expansion has had repercussions on the performance and efficiency. The findings and results of the school achievement studies cannot be interpreted in any explicit direction, however, they do promote the interpretation that the opening up of the Gymnasiums has decreased the performance far less than have been suggested. (Köller, 2013, p. 58.) (Authors translation.)

5.9.1 Teacher (D14) in *Geisteswissenschaften* and change in societal (value) structures

The term *Geisteswissenschaften* refers to humanistic science, *Geiste* is a German word/concept that incorporate mind, spirit and intellect. It is a 'collective name for around 40 different individual disciplines where philosophy, sociology, history and language belong to the more reputable. Since the turn of the century, the field of *Geisteswissenschaften* has been in focus of criticism, both external and internal. It has been aimed at theoretical and institutional dilemmas: an obscure scientific concept and an excessive diversification of institutions with a low degree of cooperation (cf. Gethmann, Langewiesche, Mittelstrass, Simon and Stock, 2005, pp. 2-7; *Süddeutsche Zeitung*, 2005). The outcome has been repeated cut downs and demands on efficiency judged by the same standards as natural science. In many ways, the situation of *Geisteswissenschaften* reflects the problems education in general is facing.

(D14) I also have contacts with the University where changes already are observable. Firstly, students have become much younger and because of that doesn't have the same level. I mean when they start here [...] depending on that we have the G8, which means 8 years study instead of 9 years, the graduates are a lot younger and it does have an impact [...] but I think that to learn subjects in Geisteswissenschaft demands a certain degree of maturity, to be able to work through such issues as humanity, personality and life [...] the students are often not mature enough [...] often they don't have the interest [...]. Well the school doesn't have the same status and position as before. That is our impression here at the Gymnasium and the reason is the [educational] hullabaloo, surely orchestrated, with the outcome that many more students come to the Gymnasium in these days. [...] the system have changed and because of that the parents say that their children need to go at a Gymnasium because they must get a solid education with an Abitur otherwise they can't get any proper and good work [...] a high percentage is not suited for this kind of relatively theoretical work.

The traditional symbolic value of Gymnasium is still prevalent, holding a promise of being selected, but change in society has repercussions on the conditions both for teachers and the 'chosen'.

(D14) [...] this is a second point that the promoting of the students is handed over to the school. Though it is a bit different here, the students come from the countryside, they are a bit less complicated, and their lives are more traditional [...] typical Bavarian, that the mothers still make the dinner for example [...] and the parents do take interest in how well the children does in school [...]. But in spite of it all, we also observe that the parents tend to transfer the responsibility of the fostering of education to the school. That they don't ensure that the children are well prepared, have done their home works [...]. The reason is that nowadays it's common that both parents work [...] each one has enough with oneself and it's also a lot about financial problems, something we notice more, [for example] when we do our excursions and trips with the children.

The illusory trinity – state (Bavaria), countryside and traditional life – is the referential framework where values are made visual and changes can be traced and discovered. Hence, the trinity specifies the said heart values of the state's ruling party and its state dogma; set of principals or articles of faith. 'Bavaria is the incomparable landscape, living tradition, centuries-old culture', 'The family must be in the Centre. Families are the most precious treasure of our society', 'Live in one of the most

modern and most advanced countries of the world, living in a culture of traditions and customs [...], Bavaria is just not a country like any other. Bavaria is something special. And we are proud of` (CSU, no date). Such proclamations when presented in – and masked by – official language, become the legitimate delimitator of what it is to be a citizen and thereby implicit producing a distinct discourse of estrangement, those not qualified, the other, in short the reproduction of dichotomy. In education, as the prime introducer and producer of citizenship and thereby *valued* values, societal changes influence and affect. That is, societal change affects the positions of actors and in the struggle the outcome depends on the actors' amount of capitals: economic, symbolic etc. However, the political level, as the dominant in the education field, has a preparedness to internalize changes since internalization of societal changes is the condition under which all politics are produced and carried out. In that way, the lower levels of actors in education, the practitioners, has but their symbolic and connected social capital to invest in the struggle, capital conferred to them by the values maintained and upheld in the state⁷⁹, such as educational credentials, group and family belongings (cf. Bourdieu, 1991, pp. 220-225; Bourdieu, 1986, p. 248).

5.9.2 An increasingly diverse and complex student community

(D14) No, we say that not much [of ILSA] is of any use, the question is always how these studies are applied [...] in PISA Finland won but why? We say that Finland has a different teacher-student ratio. They have smaller groups and the teachers can spend time and effort on each student [...] it's not possible for us with 30 children to promote each one. This is a Regelschule⁸⁰ and no private school so we have to rely on the parents, that they have adjusted and made certain that the children are apt for the school [...] and when the children are not [as children with difficulties or diagnosis] it cost time [...].

In 2009 the Federal Government of Germany ratified the UN Convention on the Rights of Persons with Disabilities. The so called *Inclusion* means that all people will have the possibility to participate on an equal footing in society. In Bavaria the inclusion has meant that educational transitional paths have been created for disabled children together with extended rights for parents to choose school form (Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, 2014b, pp. 2-8). The percentage of children with special needs in Bavaria has risen from 5.5 % in 2008/09 to 6 % in 2011/12, a trend visual in Germany as a whole⁸¹ (Klemm, 2013, p. 9). The percentage of children with special needs in Bavaria in special schools where 2009 4.7 % compared to 4.6 % in 2012, that is, the inclusion have not affected the proportion of children in special schools, partly due to an increase of children with special needs (Bertelsmann Stiftung, 2014). However, the percentage of children with special needs in inclusion has risen from 16.1 % 2008/09, to 22.4 % 2012/12, an increase of almost 40 % (Klemm, 2013, p. 14). In the debate regarding the inclusion in Bavaria, teacher, school managers and representatives have focused on some main problematic issues: a lack of specialized teachers, the high number of students in each class that affects the stated individual promotion of pupils, the need for longer preparation time for lessons and the necessity of cooperation between teachers and staff (cf. BLLV, no date).

(D14) Inclusion is of course a specific problem [...] inclusion does not refer to a child with a specific problem as ADHD⁸², it refers to a wide variety of disabilities [...]. That is, you have to decide for every case individually, but we are forced by the parents to accept their child because they have decided that their child are not suited for a special school, [the child] is far too intelligent and must be at a Gymnasium. I mean, sometimes this is not to the benefit of

⁷⁹ Do not have to be restricted to the state or have the state as origin, but must be made the value of the state.

⁸⁰ State funded schools with general education, primary or secondary.

⁸¹ For Germany the percentage in 2011/12 was 6.4 %.

⁸² Attention Deficit Hyperactivity Disorder.

the child, sometimes yes but sometimes no [...] and the problem is of course that the parents [now] alone can decide for their children, that is the obstacle.

In a society where human capital is made the canon, competitiveness becomes the prime capital and education the mean to achieve the full `potential` of each individual (cf. OECD, no date/e, p. 4), and if education is the privilege of all, the demand and requirement of the individual will grow and multiply. Hence, equality will become equivalent to the right to choose educational path. In education, this means that groups that before were denounced or made *incompatible*, now become the official sign of the equality of the education system. Exclusion made inclusion. However, since the Bavarian – as well as the German – education system is based on structural selection, this u-turn will be difficult to achieve and could rather be interpreted as a way to effectuate the implementation of a demand and supply market for a group of new costumers. That is, official reform measures in line with the doxa of the day that mask the status quo outcome (e.g Oelkers, 2011, pp. 2-5).

5.9.3 *Elternwille*⁸³ and the notion of Market

(D14) [...] but of course it's said today that we as a school, should act independent and be self responsible in many areas, and that is precisely these self responsible management ideas that exists and comes from the economy [economic theory]`. In a state the education system have mainly two tasks: one collective, to foster the individual into a collective with shared societal values, and one individual, to promote and develop the individual. This is a delicate balance act between collective and individual, in which the dominant political power will shift the center of gravity in the direction consistent with its (pronounced) values. Hence, when education become seen as a market where knowledge is bought and sold, where education of children is spoken in terms of `upgrading human capital` (OECD, 1996, p. 19), the individual will be the hegemonic doctrine. In specific the accumulated value of the individual as visualized in the assessments of individuals and education systems since the market value will be the yardstick (cf. OECD, no date/a, pp. 3-10; Krautz, 2007, pp. 81-87).

(D14) Well, related with the PISA study there were a range of tests introduced [...] we have the VERA test, we have the test for each grade, everything is tested and everything is evaluated. I was part in a school development team, we have already had two external evaluation, and this intern evaluation we had to produce ourselves based on self responsibility. We had to fabricate our own evaluation [...].

Educational reforms will follow in line with an economizing of education by imposing self responsibility and self management which in turn will reinforce the individual's right to choose and decide – as also to fail (cf. Staatliche Schulberatung in Bayern, 2009). The economizing will also require a faster conversion of theoretical knowledge into practical useful *competencies*.

(D14) Yes, the Elternwille counts, and what the teacher think based on pedagogy and achievement, the teacher's experience, doesn't count. Well, it's just those in Bavarian politic, the politicians want to be reelected. [About G8] I would say it started in late 2008 and had to do with changes that occurred in other Länder for example in Baden-Württemberg, and Bavaria did not want to be left behind. In all Europe they have these 12 years of [primary and secondary] schooling and here it used to be 13 years. In that way they wanted to equalize, that the people here [in Bavaria] would come out in the job market after the same amount of time.

The Länder comparison tests were developed by IQB, assigned by KMK, to review the accordance between achieved competence levels and the standards with four specific pinpointed aims: (1) to increase the transparency and objectivity, (2) to secure equal performances in all the *Länder*, (3) to improve the accountability in the educational sector and (4) to provide starting points or indicators

⁸³ The *will of the parents*, the possibility to influence the decision of the educational path of the child.

for the common development of education and teaching (IQB, 2011, p. 9). However, the Länder comparison tests did not become a device for improvement but for prestigious competition. This notion of comparativeness resulted in the development of German and Länder specific tests that together covered all levels of primary education, all part of an intricate and cohesive design (cf. IQB, no date/d).

5.9.4 The task of making practice work

The economizing means to be able to visualize a (thought) value, illustrated by assessments and evaluations, but all that is visual looks best at a distance. The fractions and problems in reality must be invisible for all but those in practice.

(D14) Yes, we [in Bavaria] still have the central exams [...] it has changed a bit and every Land have their specific central exam but it has developed in the way to make a minor part of these exams similar for all the Länder. [...] we had for about one and a half years ago a test on such a similar exam [in mathematics] and because of it we had to spent extra time to correct it, extra work that is, and it was really very poor conducted and performed. The Bavarian students turned out especially bad and the reason were that they weren't allowed to use a calculator. The other Länder had prepared their students because they had previously already often worked without calculator. Bavarian students always use a calculator and suddenly they had to put it away, [it was] the reason for their poor performance. Anyway, we teachers were told that we had to do this extra work, to correct, but the good news was that we were allowed to use the results as part of the grading. However the results turned out to be very bad and then the Ministry [of education] changed their mind just before the grading [...] we were not allowed to use the results. Such things, of course, upset a lot of the teachers because there is always, constantly something new and something novel that is tested but never prepared. They don't think on the problems they create and it is not the only time, it's just an example of such a flop. But we must bear the consequences [...] everything and always is homespun and simply not thought through and that's how it was with the G8 reform according to us [teachers].

Educational systems are competitive systems where diplomas and grades will direct the outcome for the individual. In the same time it is a structure of functions and an organization that produces durable, transposable dispositions in sense of inferring the sense of what it means to belong to the society and a normalization of (Bourdieu and Passeron, 2000, p. 179). Bavaria did not only partake in ILSA, it was a society in which education already had a prime position and the influence of QCE methodology already, to some extent, was present. Therefore partaking meant a reinforced idea of the already strong educational concept and initiation of quality, standards and assessments as synonymous with increased performances, in line with the political proclaimed vision of developing the human resources of economy – the human capital. The cost in form of the extra work load and the (necessary) creative patching and mending (to make practice work) became the task of practitioners.

6. Discussion

The study has thus shown how CE became QCE when education moved into the forefront of powerful IOs such as OECD. The upgrading of education coincided with – and was interdependent on – a positional shift of economic theory. Furthermore, as long as education was referred to a secondary role, the economy of education was related to material assets or *Material Capital*. The concept of humans as capital, *Human Capital*, took hold only when education was elevated into a central part of IOs work. In the early 90s empiric economic theory had progressed from the position of a disputed theory of science into the pragmatic realization and utilization of a dominant epistemological research stance, QCE. It was the creation and incorporation of a subjective methodology that could gain and acquire cultural capital under the premise to be objective and universal (e.g. Bourdieu, 1995, pp. 111-113); to be of common interest for the common good. *‘The possession of a great cultural capital is seen as “something special”, and therefore it is the basis for further material and symbolic profits’* (Bourdieu, 1997, p. 219). It was in this context that Germany, the Länder and the Bund, decided to embark on assessments of the education system in large scale, in accordance with the guidelines of IOs and transnational actors such as OECD and ERT (eg. Lindgard and Greek, 2008, p. 9-14; ERT, 1995⁸⁴, p. 15). The consequences of the decision were far-reaching.

6.1 Change in the balance of power

Bavaria had a key role in the introduction of ILSA and monitoring of education based on ILSA core concepts: Quality based on Standard based on Assessment. The reasons were several but one was that Bavaria already had an education system closely linked to the economy depending on the close ties between the economic and political fields. The education system was formal, hierarchic, stratified and had a reputation of high (top) performance. Since partaking in ILSA could be expected to confirm the education policy of the state party, Bavaria became the promoter. Together, these aspects paved the way for the enactment of a test discourse on the surface familiar but in reality alien (cf. Aljets, 2014, pp. 22-25; Boller, 2009, pp. 31-34), partly because measurable standards were modeled on a national level, not Länder specific, in spite of the cultural, economic and social differences. The introduction of national education standards and the following application of measurement practice also meant that education became output oriented, and that the traditional German *Bildung* concept was transformed⁸⁵; *Bildung* became equivalent with training of basic education competence in preselected subjects. Furthermore, quality became the all-encompassing motto that education would be defined by. In the field of educational research, the former dominant *Erziehungswissenschaft* was challenged by EB, that is, the introduction of ILSA triggered a positional shift in which EB, backed by the political field became the dominant. However, when positions change it affects the exposed (group or groups) relation towards collective in other fields that hold corresponding position. The reason is loss of capital for the exposed and increase of capitals as well as ‘allies’ for the counterpart. For example, in education *Erziehungswissenschaft* with emphasize on qualitative research was the dominant, it was replaced by EB with emphasize on (economic) quantitative research. This shift meant that QCE methodology was promoted, an economic based methodology in line with the policy of Bavaria’s political dominant (cf. Hartung, 2011. pp. 90-91). Bavaria was also very fast to act on the questioning of *Erziehungswissenschaft* by the instigation of the commission of Evaluation of *Erziehungswissenschaft* in 2000 (c.f. Hörmann, 2005, pp. 44-49). Thus, the *rational research* of the day, QCE, became the rationale (cf. Ratdke, 2015, p. 5).

Monitoring of education, specified as the measuring of a set of indicative subjects distinguished from other subjects/areas through their claimed specific characters of prime economic competitive value, marks a change of the German educational context in which the valued standing of education – as part of the individual’s development – has been altered into the measurable individual, though

⁸⁴ See in specific the member list p. 31.

⁸⁵ Even though to what degree and in which institutions it still remained, can be questioned.

collectively performed. In other words, the theory of education as second to the theory of economic (cf. Maritzen and Tränkmann, 2014, pp. 30 -32; Björklund and Lindahl, 2005, pp. 7-9). Tillmann et. al. (2008) have studied the empirical effects of education research, in form of the PISA study, on the education policy in four (German) Länder. The findings showed that even if widely used rhetorically, the outcome in practice were either the implementation of already planned reforms or that the results were used to legitimize and confirm already ongoing changes. However, the study also (unintentionally) displays a set of concepts that moved into the forefront of education: Quality, Standard and recurring evaluation, Assessment (e.g. *ibid.*, pp. 382-383).

6.2 Testing as part of *upgrading* education

KMK implemented the work of binding standards that was to guide the quality work in education. The standards were to be based on norms (*Regel*), not minimal as suggested – at first – by the expertise. The reason was said that the use and level of standards had to be empirically evaluated (by *Regelstandards*) before they could be translated into minimal standards, that is, not to risk a too low or high level of standards (KMK, 2004, pp. 14-15). In the initiation it was said that education and curricula already fulfilled many of the new demands put on the system by the standards. ISB, responsible for the introduction in Bavaria, also underlined that the new standards would improve the equality and permeability of the education system (Hechenleitner and Schwartzkopf, 2005, p. 8). However, the most salient feature have been the increased element of new assessment tests⁸⁶ that has increased the workload of teachers and administration but whose weight differs depending on the status on the education site. Hence, the study's example of a renowned Humanistic Gymnasium where the `new` has been relegated to a secondary role (see 5.8.1). Holders of prestigious symbolic capital – in terms of clientele and traditions – have seen an increase of their capital's economic exchange rate (Bourdieu, 1995, pp. 46-48).

The orchestrating of the presentation of the PISA 2000 result, the *legalizing* of ILSA, was dependent on a negative outcome in the present with future repercussion; if positive it would just have confirmed a widely held belief and not mobilize demands for action. In that way the overtly bad results – in each case described as – of the German Länder acted as a *door opener*, the legitimizer of legitimate claim for change of the educational context. However, Bavaria as one of the few successful did not suffer any media hunt, even if there were indignant outbursts in media about the state of German education⁸⁷. By that, Bavaria or the dominant political actor, CSU, could claim the ratification of the tripartite educational system (see 5.2.1), in spite of the proven lack of equality that had been singled out as a deficit area. Instead of correcting any deficit, the proclaimed reason for Bavaria to partake in ILSA was to maintain their top position. Simultaneously, and correspondently, the increased awareness of the system's inequality and the notion of the importance of education caused a *rush* on and demands of access to Gymnasium as seen in the statistic (see table 3, p. 25).

6.3 The quality concept

Although the Bavarian education system always were a competitive system, the participation in ILSA did not just effect the number of tests, but also brought in a different test culture (format, interval, cohorts etc.), all implied as part of the *Quality* concept (Podium 2000: Bayerisches Staatsministerium für Unterricht und Kultus, 2010, p. 7). Quality became the indication of good education and the term was now state-promoted (despite the lack of definition). Most evident in 2003 when ISBs abbreviation was changed into *Staatsinstitut für Schulqualität und Bildungsforschung* and the establishment of the Quality Agency (see 5.5). However, the elusiveness of the term would soon

⁸⁶ The term assessment is used depending on the repetitive character, in order to guide action.

⁸⁷ These editorials and published letters to the editor were mainly found in newspaper outside of the Bavarian capital Munich, according to the author's nonsignificant survey of media coverage in Bavaria after the publication of the first PISA study (performed 2015 in the Munich City Library).

become apparent. In the last 13 years, QA and QA4 (internal and external evaluations) have searched for those indicators that best describe a school's quality but until this date, there is only one defined scale, one that is used nationally. It was worked out (in Germany) in the form of a decree decided centrally by KMK and monitored by the institute that characteristically has quality inscribed in its name, IQB⁸⁸. The scale is standard (cf. Schmidlein, 2010, p. 121).

How come that quality is possible to work out in the form of performance standards for age specific cohorts but not possible to determine when it comes to indicators of best school practice? ISB/QAs continuous work with quality indicators demonstrates the limitation of the quality concept when transferred into education; comparative estimated values have no counterpart in practice; no specific set of indicators that can guide action. The reason is that the specificity of each context is neglected in ILSA, or more correct, it is a precondition for QCE methodology influenced by economic theory. Thus, partaking in ILSA signify an embedded acceptance and adherence to concepts derived from the theorem of human capital: that there exist anything such as human capital and that it can be refined, and used to secure policy relevant information; the dominance of politic by an economic theoretical discourse (cf. Hartung, 2011. pp. 90-91). Bavaria had already after TIMSS 1995 introduced and extended the testing performances and comparisons, that is, increased its testing procedure, thus data on the educational status were already at hand (see 5.1.2), data that could confirm the competitiveness of Bavarian education. The partaking in ILSA and authorization of EB in Bavaria was thereby no venturesome political act, it was just the extension of something already existing to the form but not content. The partaking also brought in an alien language discourse – psychometric and econometric – mirrored in the Bavarian administrations first task when the PISA 2000 measurement were to be introduced: to learn and (learn how to) apply it (see 5.1.4). This specific language is marked by symbolic important terms implicitly declared as universal concepts, adoptable on all education systems, and thus containing a value of an enriched symbolic capital (cf. Bourdieu and Passeron, 1990, p. 73). This technocratic language (in QCE) based on mathematical calculations acted as an efficient barrier (or firewall) against `outsiders` not familiar with the methodology. Hence, criticism could by that be referred to as lack of knowledge and insights, directed at the `other` or nonconsecrated (e.g. Prenzel et al., 2008), or misunderstandings and misinterpretations directed at those (few) critical consecrated (e.g. Prenzel, 2006).

6.4 The affect on practice

The participation in ILSA greatly strengthened the public perception of the importance of education, the effect was a rush for best education practice to the detriment of the traditional lower level of secondary education which lost position and students. Conversely, the *provider of knowledge*, the practitioners - as in teachers and school administration – should thus have had the best opportunities to improve their position and value. However, at the same time these professions came in focus of the critique, accused of the poor results in the measurements (e.g. Aktionsrat Bildung⁸⁹, 2015, pp. 14-17). This implicit announcement was an interpretation made by the dominant, in politic and economy, which deflected the public discontent towards the lower levels of the field, especially since the education ministers acted fast to show how educational reforms already were implemented or (at least) planned (Tilmann et. al. pp. 381-400). However, the discontent in Bavaria had more to do with the initiated *rush* for best practice which caused, for example, the development of transitional ways between educational paths, even if in reality these were not to form any entrance pass but exit. The response of the dominant had the illusion of upholding objective (universal) values, to ensure the legitimization of state power (Bourdieu and Wacquant, 1992, pp. 224-235). On the other end, the practitioners had to deal with the inflow of a

⁸⁸ Illustrative enough together with the term Bildung.

⁸⁹ The Action Council, established 2005 by the Bavarian Business Association. It is an expert committee of renowned educational researcher that evaluates the state of German education, targeting areas of specific interest.

(more) heterogenic body of students, less destined for the practice of studies, caused by the notion of the beatific of education and the accompanying press, by parents and surroundings, on compulsory teachers to give the students an entrance pass to Gymnasium instead of Realschule or Mittelschule (see 5.7.1-2). In the center of this gravitation stood the economic (self-) interest born in the economization of education: concurrent evaluations, change of curricula, evening classes in the interest of competitiveness, basic competencies instead of specific knowledge, changed content and number of tests; the superficiality of knowledge and the detriment of Bildung. Though Bildung does remain but in those `havens` that are less inclined to have felt the change in education because of their assets, their symbolic and economic capital, the education sites closest related to the higher strata of the dominant fields.

To grant `education` to all, but in a purely formal way, is to exclude from it, under an appearance of humanism, all those that are deprived of the means of realizing it⁹⁰.
(Paraphrasing Bourdieu, 2000, p. 65).

6.5 Altered power relations

The Bavarian state party came to embrace EB as the modern tool to monitor and safeguard the education system. The outcome was more assessments, tests and evaluations, all part in a methodology based on mathematical calculations and numbers. It was an arrangement of mutual benefit in the beginning: EB needed the political backing and support on state level, and CSU wanted the economic theory oriented research stance which fit their description of future challenges. Bavaria was the first Land (in 1998) that introduced reforms in education because of the TIMSS 1997 survey. Within a few years both the concepts and language of EB had become a reality. In recent years this reciprocity between EB and Länder has changed. EB is now a well established research institution with a widespread research infrastructure (institutions, professorships, funding, etc.). The need for state level backing has decreased and in later assessments – which is a food for thought – the test results of Bavaria has not been as successful as before, something CSU has chosen to explain away or even ignore. However, it points at a shift in the relationship and positions between EB and Länder/federal states.

⁹⁰ In original: ` To grant `humanity` to all, but in a purely formal way, is to exclude from it, under an appearance of humanism, all those that are deprived of the means of realizing it. `

Epilogue

Partaking in ILSA means that specific terms that are markers of economic concepts will be introduced as axioms even if they really are postulates – (economic) theory bound but not proven. Quality, being the most illustrative example, became the key specifier of the mass of established quality institutions in the Länder, institutions at the side of the established educational structure, responsible for the quality of education by evaluations. It was therefore a significant event that occurred in 2015 in Nordrhein-Westfalen when the termination of *‘the Agency for Quality assurance, Evaluation and Autonomy of Schools’*, AQS, was announced (Landtag Rheinland-Pfalz, 2015)⁹¹. AQS was established 2004/2005 as a direct response to the German partaking in ILSA and the implemented standards by KMK (Tillmann et. al., pp. 134-135). However, the establishment and the work of AQS became continuously criticized by the political opposition and the teachers’ unions. The targeted issues were foremost the cost and benefit on grounds of the amount of working hours involved, the time frame of and between observation and feed-back, as also the selection of quality criteria (Landtag Rheinland-Pfalz, 2015; Allgemeine Zeitung, 2015; phv, 2012).

⁹¹ In 2007 in Nordrhein-Westfalen the first closure of a quality institute in education took place but it was replaced with a successor in 2013 under the name *‘the Agency for Quality and Support-the State Institute for School’* (QUA-LiS NRW, 2016).

Limitations and directions for future research

Research is a precarious employment and to stay in line with a methodological approach is a conscientious mission, especially for an inexperienced researcher. There is always the risk of crossing the (thin) line between objectivity and bias, and to interpret the results and findings in an increasingly narrow way. Another deficit is that the amateur often get stuck in an idea that, to the reader's weariness, becomes repeated endlessly, and thereby loses the overall perspective which is the base for understanding and interpretation. These are all, in various degrees, characteristics of this study.

One technical aspect that should be mentioned is the change of web addresses that unfortunately occur quite frequently. In the course of the prolonged time for the study, several addresses were found to have been altered or removed, however, the original addresses are those presented.

The outcome for educational research in general by QCE, as described by the participants in the study, is that education research has become increasingly difficult to implement, not by open resistance but by tacit avoidance by schools, burdened by the added and continuously altered, demands of the monitoring of education. What are the consequences when education research is seen as unwanted and unwished by those in focus of it?

The formation of a *'knowledge'* (dependent or maybe addicted) society where education is governed by a measurement's practice can only be justified on grounds of an disinterested self interest, serving the society (state/Land) for *the common good* (cf. Bourdieu, 1996, pp. 379-382). Is there any other possible consequence of the *testing regime* than a hierarchisation/increased *'levelling'* where those in the inferior positions, lowest levels, will lose capital because they are furthest away from the theoretical concepts adopted by the dominant. In other words, they are the *researched* empiri (practice) and by that, holders of a subordinate position. Or can further research find circumstances under which holders of subordinate position in education could gain strength and influence by levelling?

References

- Abgeordnetenhaus Berlin. (2012). *3.Grundlegende Mängel bei der Kooperation der Humboldt-Universität zu Berlin mit einem An-Institut*. Downloaded 2015-03-02 at <http://www.parlament-berlin.de/ados/17/IIIPlen/vorgang/d17-1061.pdf>
- Ackeren, I. van (2002). Von FIMS und FISS bis TIMSS und PISA. Schulleistungen in Deutschland im historischen und internationalen Vergleich. *Die Deutsche Schule*, 94, pp. 157-175.
- Aktionsrat Bildung (2015). *Bildung. Mehr als Fachlichkeit*. Downloaded 2016-04-01 at http://www.aktionsrat-bildung.de/fileadmin/Dokumente/Gutachten_2015_Internet.pdf
- Aljets, E. (2014). *Der Aufstieg der Empirischen Bildungsforschung*. Bremen: Springer Verlag
- Bayern DGB. (2013). *Mitgliederzahlen DGB Bayern*. Downloaded 2015-03-04 at http://bayern.dgb.de/ueber-uns/mitgliederzahlen/?tab=tab_0_0#tabnav
- Avenarius, H., Ditton, H., Döbert, H., Klemm, K., Klieme, E., Rürup, H., Tenorth, H-E., Weishaupt H. & Weiss, M. (2003). *Bildungsbericht für Deutschland*. Opladen: Leske+Budrich
- Baumert, J., Ardelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K-J. & Weiss, M. (2002). *PISA 2000-Die Länder der Bundesrepublik Deutschland in Vergleich*. Wiesbaden: Springer Fachmedien
- The Bavarian Ministry of Economic Affairs and Media, Energy and Technology. (2014). *Bavaria. A Guide for Investors*. Downloaded 2015-03-04 at http://www.invest-in-bavaria.com/fileadmin/media/documents/publications/Investorsguide/Investorsguide_Bavaria_2014.pdf
- Bavarian Ministry of the Interior, for Building and Transport. (no date). *Staat und Kommunen. Fachkräftezuwanderung*. Downloaded 2015-03-06 at <https://www.stmi.bayern.de/suk/asylmigration/fachkraeftezuwanderung/index.php>
- Bayerisches Landesamt für Statistik. (2012). *Pressemitteilungen*. Downloaded 2015-03-06 at https://www.statistik.bayern.de/presse/archiv/2012/259_2012.php
- Bayerisches Landesamt für Statistik. (2014). *Zensus 2011. Pressemitteilungen*. Downloaded 2015-08-24 and 2015-09-18 at https://www.statistik.bayern.de/presse/archiv/2014/177_2014.php
- Bayerische Landtag. (2013). *Kernaussagen zur Bayerischen Bildungspolitik*. Downloaded 2015-04-28 at http://www.csu.de/common/migrated/csucontent/Kernaussagen_zur_Bayerischen_Bildungspolitik_2013-07.pdf
- Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst. (2013a). *Humanistischen Gymnasium: Reich werden mit klassischen Sprachen*. <http://www.km.bayern.de/allgemein/meldung/2341/humanistisches-gymnasium-reich-werden-mit-klassischen-sprachen.html>
- Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst. (2013b). *Pressemitteilung Nr. 348 vom 25.12.2013 Bayerns Bildungsminister Ludwig Spaenle baut*

Ganztagsschulen konsequent aus - flächendeckend und bedarfsorientiert. Downloaded 2015-05-04 at <https://www.km.bayern.de/pressemitteilung/8733/nr-348-vom-25-12-2013.html>

Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst. (2014a). *Reihe A, Bildungsstatistik Heft 59. Schule und Bildung in Bayern 2014*. München: Schriften des Bayerischen Staatsministeriums für Bildung und Kultus, Wissenschaft und Kunst

Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst. (2014b). *Bayerns Schritte auf dem Weg zur Inklusion*. Downloaded 2015-10-07 at <https://www.schulberatung.bayern.de/schulberatung/muenchen/regionalinfo/inklusion/>

Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst. (2015a). *Pressemitteilung Nr. 003 vom 11.01.2015 Bayerns Bildungsministerium: Hohe Bildungschancen für die Schülerinnen und Schüler - Durchlässigkeit in Bayerns Schulwesen massiv ausgebaut – Fakten als Antwort auf SPD-Umfrage*. <http://www.km.bayern.de/pressemitteilung/9296/nr-003-vom-11-01-2015.html%20%20/>

Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst. (2015b). *Prognose zum Lehrerberuf in Bayern 2015*. München.

Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst. (no date/a). *Qualitätssicherung*. Downloaded 2015-05-31 at <https://www.km.bayern.de/lehrer/qualitaetssicherung-und-schulentwicklung/qualitaetssicherung.html>

Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst. (no date/b). *Gestaltung der Gegenwart zur Verbesserung der Zukunft*. Downloaded 2015-05-31 at <http://www.km.bayern.de/ministerium/schule-und-ausbildung/qualitaetssicherung-und-schulentwicklung.html>

Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst. (no date/c). *Übertritt und Schulwechsel in Bayern: So geht die Schulkarriere weiter*. Downloaded 2015-08-19 at <http://www.km.bayern.de/umzug>

Bayerisches Staatsministerium für Unterricht und Kultus. (2010). *Externe Evaluation an Bayerns Schulen. Das Konzept, die Instrumente, die Umsetzung*. München

Bayerische Staatsregierung. (2010). *Wir beraten die Eltern künftig früher und intensiver“ Kultusminister Spaenle zum neuen Übertrittsverfahren für Grundschüler – Erfahrungen der Eltern und Lehrkräfte erheben – Arbeitshilfe vorgestellt*. Downloaded 2015-08-20 at <http://www.bayern.de/wir-beraten-die-eltern-kuenftig-frueher-und-intensiver-kultusminister-spaenle-zum-neuen-uebertrittsverfahren-fuer-grundschueler-erfahrungen-der-eltern-und-lehrkraefte-erheben-arbeitshilfe-vorges/>

BDA (1998). *Mehr ökonomische Bildung in der Schule – Memorandum der unternehmerischen Wirtschaft zur ökonomischen Bildung im allgemeinbildenden Schulwesen*. Downloaded 2015-11-17 at <https://www-user.tu-chemnitz.de/~nean/Alte%20Unterlagen%202006/Lehrerfortbildung%20WTH/Memorandum%20zur%20oekonomischen%20Bildung.pdf>

Bellmann, J. (2006). Bildungsforschung und Bildungspolitik im Zeitalter „Neuer Steuerung“. *Zeitschrift für Pädagogik*, Jahrg. 52 (4), pp. 487-504.

Berner, E., Oelkers, J. & Reusser, K. (2008). Implementation von Bildungsstandards: Bedingungen des Gelingens (und Scheiterns) aus internationaler Sicht. Klieme, E. & Tippet, D. (Eds.). *Qualitätssicherungen in Bildungswesen* (pp. 210-226). Weinheim u.a.: Beltz

Bernt, T., Remme, H. & Langer, T. (2014). Schulentwicklung und Inspektionsprozesse am Beispiel des Hansa-Gymnasiums 2009-2013. Fickermann, D. & Maritzen, N. (Eds.). *Grundlagen für eine daten- und theoriegestützte Schulentwicklung* (pp 253-274). Münster: Waxmann (B)

Bertelsmann Stiftung, Institut für Schulentwicklungsforschung der Technischen Universität Dortmund, Institut für Erziehungswissenschaft der Friedrich-Schiller-Universität Jena (2014). *Chancenspiegel 2014*. Downloaded 2015-08-20 at http://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Chancenspiegel_2014_Kurzfassung.pdf

Bildungsweb. (2015). *Schulen-Vergleich.de*. Downloaded 2015-08-27 at <http://www.schulen-vergleich.de/>

Bildungsstandards der Kultusministerkonferenz. (2004). *Bildungsstandards der Kultusministerkonferenz*. Downloaded 2015-11-02 at http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Konzeption-Entwicklung.pdf

Bildung und Sport (2011). *Information der Bildungsberatung International Übersicht über die Bildungssysteme*. Downloaded 2015-02-09 at <http://www.google.se/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=0CDsQFjAD&url=http%3A%2F%2Fwww.muenchen.de%2FfratHaus%2FHome%2FStadtverwaltung%2FReferat-fuer-Bildung>

Björklund, A. & Lindahl, M. (2005). *Utbildning och ekonomisk utveckling-vad visar den empiriska forskningen om orsakssambanden*. Regeringskansliet: Finansdepartementet

BLK (1997): Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung. *Gutachten zur Vorbereitung des Programms "Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts"*. Bonn.

BLK (2001). *SINUS*. Downloaded 2015-03-15 at <http://blk.mat.uni-bayreuth.de/indexblk.html>

BLLV (2015a). *BLLV. Der kompetente Partner aller Pädagogen*. Downloaded 2015-03-02 at <http://www.bllv.de/Portraet.6626.0.html>

BLLV (2015b). *Testmarathon an Bayerns Schulen*. Downloaded 2015-03-04 at <http://bildungsklick.de/pm/92941/testmarathon-an-bayerns-schulen/>

BLLV. (no date). *Inklusion an Bayerns Schulen-Lehrerbefragung*. Downloaded 2015-10-12 at https://www.bllv.de/fileadmin/Dateien/Land-PDF/Wissenschaft/Befrag_Inklusion_Bericht.pdf

BMBF (2007). *Bekanntmachung des Verwaltungsabkommens über das Zusammenwirken von Bund und Ländern gemäß Artikel 91b Abs 2 des Grundgesetzes (Feststellung der Leistungsfähigkeit des Bildungswesens im internationalen Vergleich)*. Downloaded 2015-11-15 at <http://www.kmk.org/fileadmin/Dateien/pdf/Bildung/AllgBildung/Verwaltungsabkommen-2007-06-13-Bundesanzeiger.pdf>

BMBF (2008). *Bildungsforschung Band 22. Rahmenprogramm zur Förderung der empirischen Bildungsforschung. Framework Programme for the Promotion of Educational Research*. Downloaded 2015-11-08 at https://www.bmbf.de/pub/bildungsforschung_band_zweiundzwanzig.pdf

BMBF. (2014). *Bildung in Deutschland 2014*. Downloaded 2015-09-17 at http://www.bildungsbericht.de/daten2014/bb_2014.pdf

BMBF (no date). *Ganztagsschule braucht Qualität*. Downloaded 2015-05-03 at <http://www.bmbf.de/de/1125.php>

Boller, S (2009). *Kooperation in der Schulentwicklung. Interdisziplinäre Zusammenarbeit in Evaluationsprojekten*. Wiesbaden: VS Verlag für Sozialwissenschaft

Bottani, N. & Tuijnman, A. C. (1994). The design of indicator system. Tuijnman, A. C. & Postlewaithe, T. N. (Eds.). *Monitoring the standards of education* (pp. 47-78). Oxford: Pergamon Press

Bourdieu, P. (1984a). *Homo Academicus*. Stanford: Stanford University press

Bourdieu, P. (1984b). *Distinction. Social Critique of the Judgment of Taste*. Harvard University Press

Bourdieu, P. (1986). The forms of capital. Richardson, J. C. (Ed.). *Handbook of Theory and Research for the Sociology of Education*, pp. 241-258. New York: Greenwood

Bourdieu, P. (1990a). *The Logic of Practice*. Trans. Nice. R. Stanford, CA: Stanford University Press.

Bourdieu, P. (1990b). *In Other Words. Essays Towards a Reflexive Sociology*. Cambridge: Polity Press

Bourdieu, P. (1991). *Language and symbolic power*. Thompson, J. B. (Ed.). Cambridge: Polity Press

Bourdieu, P. (1995). *Praktiskt förnuft. Bidrag till en handlingsteori*. Uddevalla: Daidalos AB

Bourdieu, P. (1996). *The State Nobility*. Cambridge: Polity Press

Bourdieu, P. (1997). Ökonomisches, kulturelles und soziales Kapital. Baumgart, F. (Ed.). *Theorien der Sozialisation* (pp. 217-231). Bad Heilbrunn: Verlag Julius Klinkhardt

Bourdieu, P. (2000). Foreword to the English edition of *Homo Academicus*. Robbins, R. (Ed.) *Pierre Bourdieu*. London: Sage Publications Ltd

Bourdieu, P. (2005). *The Social Structures of the Economy*. Cambridge. Polity Press

Bourdieu, P. & Passeron, J-C. (1977). *Reproduction in Education, society and Culture*. London: Sage Publication

Bourdieu, P. & Wacquant, L. (1992). *An invitation to reflexive sociology*. Chicago: The University of Chicago Press

Brennan, R. L. (Ed.). (2006). *Perspectives on Evolution and the Future of Educational Measurement*. Westport, CT: Praeger Publishers

Brown, M. (1996). FIMS and SIMS: The first two IEA International Mathematics Surveys. *Assessment in Education* 3 (2), pp. 181-200.

Brüsemeister, T., Preuss, B. & Wissinger, J. (2014). Schulentwicklung als Governance. Fickermann, D. & Maritzen, N. (Eds.). (2014). *Grundlagen für eine daten- und theoriegestützte Schulentwicklung* (pp. 215-234). Münster: Waxmann

Bryman, R. J. & Lie, J. (2007). *Sociology: Your compass for a new world* (3rd ed.). Belmont: Thomson Wadsworth

Bryman, A. (2007). Barriers to Integrating Quantitative and Qualitative Research. *Journal of Mixed Methods Research*, 1 (8), pp. 8-22.

Buchhaas-Birkholz, D. (2009). Die "empirische Wende" in der Bildungspolitik und in der Bildungsforschung. Zum Paradigmenwechsel des BMBF im Bereich der Forschungsförderung. *Erziehungswissenschaft* 20 (39), pp. 27-33.

Böhm-Kasper, C. & Weishaupt, H. (2008). Personal für die Bildungsforschung und wissenschaftlicher Nachwuchs. Weishaupt, F., Kraul, M., Böhm-Kasper, C., Schulzeck, U. & Zugenricker, C. (Eds). *Bildungsforschung Band 28. Zur Situation der Bildungsforschung in Deutschland* (pp. 9-34). Bonn, Berlin: BMBF

CDU/CSU. (2013). *Bayern und Sachsen sichern gutes Gesamtergebnis Deutschlands bei PISA-Studie*. Downloaded 2015-03-24 at <https://www.cducsu.de/presse/pressemitteilungen/bayern-und-sachsen-sichern-gutes-gesamtergebnis-deutschlands-bei-pisa-studie>

CERI (1978). *Centre for Educational Research and Innovation (CERI): the first ten years, 1968-1978*. Paris : Organisation for Economic Co-operation and Development

Champagne, P. (2005). "Making the people speak": On the social Uses of and Reaction to Public Opinion Polls. Wacquant, L. (Ed.). *Bourdieu and Democratic Politics* (pp. 111-132). Cambridge: Polity Press

Cheal, D. (2005). *Dimensions of sociological theory*. New York: Palgrave MacMillan

Couldry, N. (2003). Media Meta-capital: Extending the range of Bourdieu's Field Theory. *Theory and Society* 32 (5/6), pp. 653-677.

Cresswell, J. W. & Plano Clark, V. L. (2011). *Designing and Conducting Mixed Methods Research* (2nd edition). London: Sage Publication

CSU. (2013). *Der Bayernplan*. Downloaded 2015-04-29 at <http://www.csu.de/politik/beschluesse/der-bayernplan-2013-2018/>

CSU. (No date). *Was uns am Herz Liegt*. Downloaded 2015-03-04 at <http://www.csu.de/politik/themen-werte/bayern/>

DEGÖB (2015). *Publikationen in der ökonomischen Bildung*. Downloaded 2015-11-17 at <http://degoeb.de/index.php?page=publikationen>

Der Deutsche Schulpreis (no date/a). Downloaded 2015-03-12 at <http://schulpreis.bosch-stiftung.de/content/language1/html/index.asp>

der Deutsche Schulpreis (no date/b). *Ziele*. Downloaded 2015-03-12 at <http://schulpreis.bosch-stiftung.de/content/language1/html/53126.asp>

Deutscher Lehrerverband. (2000). *Memorandum. Ökonomische Grundbildung ist Teil der Allgemeinbildung*. Downloaded 2015-11-17 at <http://www.lehrerverband.de/memoekon.htm>

Dietrich, F., Heinrich, M. & Thierme, N. (Eds.). (2011). Neue Steuerung-alte Ungleichheiten? Überlegungen zu einer relevanten und dennoch wenig diskutierten Fragestellung. *Neue Steuerung-alte Ungleichheiten? Steuerung und Entwicklung im Bildungssystem* (pp. 9-23). Münster: Waxmann Verlag

DIPF. (2002). *Neuer Bildungsschock für Deutschland?* Downloaded 2015-10-30 at <http://www.bildungsserver.de/innovationsportal/bildungplus.html?artid=138>

Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. Helmke, A., Hornstein, W. & Terhart, E. (Eds.). *Qualität und Qualitätssicherung im Bildungsbereich. Zeitschrift für Pädagogik, Beiheft 41*, pp. 73-92

Eide, K. (1990). *30 Years of Educational Collaboration in the OECD*. International Congress: Planning and Management of Educational Development. UNESCO

Erdiek-Rave, U. (2006). Grußwort der Präsidentin der Kultusministerkonferenz. Blum, W., Drücke-Noe, C., Hartung, R. & Köller, O. (Eds.). *Bildungsstandards Mathematik: konkret*. (2010). 4th edition. IQB.

Erhard, J. (2010). Vorwort. Schaal, B & Huber, F. (Eds.). *Qualitätssicherung in Bildungswesen: Auftrag und Anspruch der bayerischen Qualitätsagentur* (pp 7-12). Münster: Waxmann Verlag

ERT (1995). *Education for Europeans. Towards the Learning Society*. Downloaded 2015-11-16 at http://www.ert.eu/sites/ert/files/generated/files/document/1995_education_for_europeans_-_towards_the_learning_society.pdf

EU. 2007. (no date). *From agriculture to high technology*. Downloaded 2015-10-24 at <http://www.eu2007.de/en/Germany/Bundeslaender/Bayern/technology.html>

European Parliament (2000). *Lisbon European Council 23 and 24 March 2000. Presidency conclusions*. Downloaded 15-04-23 at http://www.europarl.europa.eu/summits/lis1_en.htm

Fatke, R. & Oelkers, J. (Eds.). (2014). Das Selbstverständnis der Erziehungswissenschaft: Geschichte und Gegenwart. *Zeitschrift für Pädagogik*, 60, Beiheft, pp 7-13.

Fickermann, D. & Maritzen, N. (Eds.). (2014). Einführung. *Grundlagen für eine daten- und theoriegestützte Schulentwicklung* (pp. 11-26). Münster: Waxmann

Flyvberg, B. (2006). Five misunderstandings about case-study research. *Qualitative Enquiry*, vol. 12 (2), pp. 219-245.

Foshay, A.W., Thorndike, R. L., Hotyat, F., Pidgeon D. A. & Walker, D. A. (1962). *Educational Achievements of Thirteen-Year-Olds in Twelve Countries. Results of an international research project, 1959–1961*. Hamburg: UNESCO Institute for Education

Forskningsetiska Rådet. (1990). *Forskningsetiska principer inom humanistisk-samhällsvetenskaplig forskning*. Downloaded 2015-10-22 at <http://www.codex.vr.se/texts/HSFR.pdf>

Friedrich, A. L. (2008). *Kompendium: Intervjumetodik, kvalitativa analyser och rapportering av kvalitativa undersökningar*. Uppsala universitet: Institutionen för psykologi

Friedrich Ebert Stiftung (2009). *BILDUNG und SCHULE in BAYERN: Reformschritte für ein sozial gerechtes, innovatives und zukunftsfähiges Bildungssystem in Bayern*. München

GEBF (no date). *Vorstand der GEBF*. Downloaded 2015-04-24 at <http://www.gebf-ev.de/vorstand/>

GEW (2015). *GEW. Gewerkschaft, Erziehung und Wissenschaft*. Downloaded 2015-03-02 at http://www.gew.de/The_German_Education_Union_GEW.html

GEW-Bayern. (2015). *GEW. Gewerkschaft, Erziehung und Wissenschaft. Bayern*. Downloaded 2015-11-04 at <http://www.gew-bayern.de/index.php?id=7>

Gethmann, C-F., Langewiesche, D., Mittelstrass, J., Simon, D. & Stock, G. (2005). *Manifest Geisteswissenschaft*. Downloaded 2015-10-01 at <https://edoc.bbaw.de/frontdoor/index/index/docId/446>

GHR (2005). *Scientific Literacy*. Downloaded 2015-04-11 at <http://hypersoil.uni-muenster.de/2/01/05.htm>

Gläser, J., Enno, A., Gorga, A., Hedmo, T., Håkansson, I. & Laudel, G. (2014). Path Dependence and Policy Steering in the Social Sciences: The Varied Impact of International Large Scale Student Assessment on the Educational Sciences in Four European Countries. Whitley, R. & Gläser, J. (Eds.). *Organizational Transformation and Scientific Change: The Impact of Institutional Restructuring on Universities and Intellectual Innovation* (pp. 267-296). Bingley: Emerald Group

Gorder, K. L. (2000). Understanding school knowledge: A Critical Appraisal of Basil Bernstein and Pierre Bourdieu. Robbins, D. (Ed.). *Pierre Bourdieu* (pp. 335-346). London: Sage Publications

Grek, S. & Rinne, R. (2011). Fabricating Europe. Ozga, J., Dahler-Larsen, P., Segerholm, C. & Simola, H. (Eds.). *Fabricating Quality in Education* (pp. 19-31). London: Routledge

Gräsel, C. (2011). Die Verbreitung von Innovationen als Aufgabe der Unterrichtsforschung. Zlatkin-Troitschanskaia, O. (Eds.). *Empirische Bildungsforschung-ein Überblick aus interdisziplinärer Perspektive* (pp. 320-328). Wiesbaden: VS Verlag

Harris-Hummert, S. (2011). *Evaluating Evaluators*. (Dissertation). Wiesbaden: VS Verlag für Sozialwissenschaft

Hartong, S. (2011). *Basiskompetenzen statt Bildung*. Frankfurt: Campus Verlag

Harvey, L. & Greene, D. (1993). Defining quality. *Assessment and Evaluations*. 18 (1), pp 9-34. Downloaded 14-12-02 at <http://www.tandfonline.com/doi/abs/10.1080/0260293930180102#.VHh2iGc22Ak>

Hechenleitner, A. & Schwarzkopf, K. (2005). *ISB. KMK- Bildungsstandards. Konsequenzen für die Arbeit an die bayerischen Schulen*. Downloaded 2016-09-15 at http://www.anglistik.uni-wuerzburg.de/fileadmin/99050601/8_ek_standards_READER.pdf

- Hegarty, S. (2014). From Opinion to Evidence in Education: Torsten Husén's contribution. Nordin, A. & Sundberg, D. (Eds.) *Transnational Policy Flows in European Education* (pp. 47-56). Oxford: Symposium Books
- Herrlitz, H.-G. (2013). *Die Streit um die Gesamtschule in Niedersachsen. Göttinger Beiträge zur erziehungswissenschaftlichen Forschung*. Downloaded 14-12-03 at <http://d-nb.info/1048410153/34>
- Hessisches Kultusministerium . (2006). *Die Bildungsstandards, die Outputsteuerung und Ihre Kritiker*. Downloaded 2015-08-07 at http://www.rhs-giessen.de/data/intern/bildungsstandards_und_kritiker.pdf?cid=612167c60e0abb6af32beba76e7caff7
- Heyneman, S. P. (1993). Quantity, Quality and Source. *Comparative Education Review*, vol. 37 (4), pp. 372-388.
- Hilgers, M. & Mangez, E. (2015). Introduction of Pierre Bourdieu's theory of social fields. Hilgers, M. & Mangez, E. (Eds.). *Bourdieu's Theory of Social Fields: Concepts and applications* (pp. 1-36). New York: Routledge
- Hoffmann, E. (2013). Erfolgsmodell unter Druck. *Die Politische Meinung*, Jahrgang 58 (Nr. 522), pp. 53-59.
- Hopmann, T. S., Brienek, G. & Retzl, M. (2007). *PISA According to PISA: Does PISA Keep What It Promises*. Wien: LIT Verlag
- Hornbostel, S & Keiner, T. (2002). Evaluation der Erziehungswissenschaft. *Zeitschrift für Erziehungswissenschaft*, 5 (4), pp. 634-653.
- HRK. (2003). Downloaded 2015-08-20 at <http://www.hrk.de/positionen/gesamtliste-beschluesse/position/convention/empfehlung-des-bmbf-der-kmk-und-der-hrk-an-die-hochschulen-zur-vergabe-von-leistungspunkten-in-der/>
- Humboldt Universität zu Berlin. (2004). *Protokoll Nr. 04/05*. <https://gremien.hu-berlin.de/de/lisk/protokolle/2004/prot2004-03-08.pdf>
- Husén, T. (1973). Foreword. Comber, L. C. & Keeves, J. P. *Science Education in Nineteen Countries. An empirical study*. Stockholm: Almqvist & Wiksell
- Husén, T. (1987). *Higher education and social stratification: an international comparative study*. (34). UNESCO: International Institute for Educational Planning
- Husén, T. & Postlethwaite, T. N. (1967). Intentions and Background of the project. Husén, T. (Ed.). *International Study of Achievement in Mathematics-A Comparison of Twelve Countries*, vol. I (pp. 25-34). Stockholm: Almqvist och Wiksell.
- Husén, T & and Tuijnman, A. (1994). Monitoring standards in education : why and how it came about. A. C. Tuijnman, A. C. & Postlethwaite, T. N. (Eds.). *Monitoring the Standards of Education* (pp. 1-22). Oxford: Pergamon Press
- Husén, T., Tuijnman, A. C. & Halls, W. D. (1992). *Schooling in Modern European Society. A Report of the Academia Europaea*. Oxford: Pergamon Press

Härnqvist, K. (1975). The International Study of Educational Achievement. *Review of Research in Education*, 3 (85), pp. 85-109.

Hörmann, G. (2005). Zur Messbarkeit von Erziehung und Erziehungswissenschaft. Klement, K. (Ed). *Das Messbare und das Eigentliche* (pp. 41-60). Innsbruck: Innverlag

IEA. (2014). *Pilot Twelve-Country Study*. Downloaded 2014-06-12 at http://www.iea.nl/pilot_twelve-country_study.html

IEA. (no date). *SIMSS*. Downloaded 2014-06-12 at <http://www.iea.nl/sims.html>

IQB (2011). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufen in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann

IQB (no date/a). *Ländervergleich 2008/2009*. Downloaded 2015-03-17 at https://www.iqb.hu-berlin.de/bt/LV08_09

IQB (no date/b). *Ländervergleich*. Downloaded 2015-11-21 at <https://www.iqb.hu-berlin.de/bt/lv2012>

IQB (no date/c). *The Institute*. Downloaded 2015-11-22 at <https://www.iqb.hu-berlin.de/institut>

IQB. (no date/d). *VERA*. Downloaded 2015-07-17 at <https://www.iqb.hu-berlin.de/vera>

IQB (no date/e). *FDZ*. Downloaded 2015-07-17 at <https://www.iqb.hu-berlin.de/fdz>

ISB. (2012). *Bindungsbericht Bayern 2012*. München: Staatsinstitut für Schulqualität und Bildungsforschung

ISBa (no date/a). *Über das ISB. Partner in Sachen Bildung*. Downloaded 2015-05-08 at <https://www.isb.bayern.de/ueber-das-isb/>

ISB (no date/b). *Qualitätsagentur*. Downloaded 2015-11-06 at <https://www.isb.bayern.de/ueber-das-isb/organisation/qualitaetsagentur/>

Isserstedt, W., Middendorff, E., Fabian, G. & Wolter, A. (2007). *Die wirtschaftliche und soziale Lage der Studierenden in der Bundesrepublik Deutschland 2006*. 18. Sozialerhebung des Deutschen Studentenwerks durchgeführt durch HIS Hochschul-Informationssystem. BMBF: Berlin

Jahnke, T. (2015). Zur Kritik empirischer Unvernunft. Rademacher, S. & Wernet, A. (Eds.). *Bildungsqualen* (pp. 121-137). Wiesbaden: Springer VS

Jarl, M., Fredriksson, A. & Persson, S. (2011). New Public Management in Public Education: A catalyst for the Professionalization of Swedish School Principals. *Public Administration*, vol. 90 (2), pp. 429-444.

Jenkins, R. (2007). *Pierre Bourdieu*. London: Routledge

Josephson, P., Karlsruhn, T. & Östling, J. (2014). *The Humboldtian tradition: origins and legacies*. Leiden: Brill

Kane, M. T. (2006). Validation. R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17-64). Westport, CT: Praeger

Kangourou sans Frontières. (no date). *History*. Downloaded 2015-08-28 at <http://www.aksf.org/history.xhtml>

Kellaghan, T. (1996) IEA Studies and Educational Policy. *Assessment in Education. Principles, Policy & Practice*, 3 (2), pp. 143-160.

Klemm, K. (2013). *Inklusion in Deutschland-eine Bildungsstatistische Analyse*. Downloaded 2015-10-12 at https://www.unesco.de/fileadmin/medien/Dokumente/Bildung/Studie_Inklusion_Klemm_2013.pdf

Klemm, K. (2014). *Lehrerinnen und Lehrer der MINT-Fächer: Zur Bedarfs- und Angebotsentwicklung in den allgemein bildenden Schulen der Sekundarstufen I und II am Beispiel Nordrhein-Westfalens*. Downloaded 2015-08-14 at http://www.telekom-stiftung.de/dts-cms/sites/default/files/dts-library/body-files/rechte-spalte/05_Impulse/sonstiges/MINT-Lehrerbedarf_Studie_gesamt.pdf

Klieme, E. (2006). Empirische Unterrichtsforschung: aktuelle Entwicklungen, theoretische Grundlagen und fachspezifische Befunde. Einführung in den Thementeil. *Zeitschrift für Pädagogik* 52 (6), pp. 765-773.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Grubel, H., Prenzel, M., Reiss, K., Riquartz, K., Ross, J., Tenorth, H-E. & Vollmer, H. I. (2007). *Zur Entwicklung nationaler Bildungsstandards*. Downloaded 2015-10-30 at https://www.bmbf.de/pub/zur_entwicklung_nationaler_bildungsstandards.pdf

Klieme, E., Köller, O. & Prenzel, M. (2013). *Wo die PISA-Kritik irrt-eine Stellungnahme zu den Thesen Wolfram Meyerhöfers*. Downloaded 2015-04-07 at http://zib.education/fileadmin/user_upload/PDFs/PISA/Replik_zu_Meyerhoefer.pdf

Klieme, E. & Tippelt, R. (Eds.). (2008). Qualitätssicherung im Bildungswesen. Eine aktuelle Zwischenbilanz. *Zeitschrift für Pädagogik* (Beiheft 53), pp. 7-15.

KMK (1997a). *Konstanzer Beschluss*. Downloaded 15-01-08 at http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1997/1997_10_24-Konstanzer-Beschluss.pdf

KMK (1997b). *Eine Kultur der Anstrengung entwickeln*. Downloaded 2015-10-08 at <https://www.kmk.org/presse/pressearchiv/mitteilung/eine-kultur-der-anstrengung-entwickeln.html>

KMK (2002). *PISA 2000-Zentrale Handlungsfelder*. Downloaded 2015-05-04 at <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2002/massnahmen.pdf>

KMK (2003). *Bildungsbericht für Deutschland. Erste Befunde*. Downloaded 15-06-08 at http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_01_01-Bildungsbericht-erste-Befunde.pdf

KMK (2004). *Bildungsstandards der Kultusministerkonferenz*. Downloaded 15-01-09 at http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Konzeption-Entwicklung.pdf

- KMK (2006). *Bildungsstandards*. Downloaded 15-01-09 at http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2006/2006_06_02-Bildungsmonitoring.pdf
- KMK (2007). *Gemeinsame Bildungsforschung soll unmittelbar wirken*. Downloaded 2015-11-15 at <https://www.kmk.org/presse/pressearchiv/mitteilung/neue-steuerungselemente-fuer-weitere-verbesserung-des-bildungswesens.html>
- KMK (2010a). *„Zentrum für Internationale Vergleichsstudien“ von Bund und Ländern gegründet*. Downloaded 2015-03-12 at <http://www.kmk.org/presse-und-aktuelles/meldung/zentrum-fuer-internationale-bildungsvergleichsstudien-von-bund-und-laendern-gegruendet.html>
- KMK (2010b). *Ergebnisse der 330. Plenarsitzung der Kultusministerkonferenz in München. Mitgliederversammlung des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB)* <http://www.kmk.org/presse-und-aktuelles/meldung/ergebnisse-der-330-plenarsitzung-der-kultusministerkonferenz-in-muenchen.html>
- KMK (2012/2014). *Verfahren zur Entscheidung der Kultusministerkonferenz über Datennutzungsanträge für neuartige Ländervergleiche*. Downloaded 2015-06-21 at http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_09_20-FDZ-Datennutzungsantraege.pdf
- KMK. (2015a). *Qualitätssicherung in Schulen*. Downloaded 2015-11-02 at <https://www.kmk.org/themen/qualitaetssicherung-in-schulen.html>
- KMK (2015b). *IQB*. Downloaded 2015-06-15 at <https://www.kmk.org/themen/qualitaetssicherung-in-schulen/iqb.html>
- KMK (2015c). *VERA 3 und VERA 8. 11. Wer erhält die Ergebnisse?* Downloaded 2015-08-06 at <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsmonitoring/vergleichsarbeiten/faq.html#c10755>
- KMK (2015d). *Ländervergleich 2011*. Downloaded 2015-11-24 at <http://www.kmk.org/index.php?id=2410>
- KMK (2015e). *Verfahren zur Entscheidung der Kultusministerkonferenz über Datennutzungsanträge für Neuartige Ländervergleiche*. Berlin/Bonn
- KMK (2015f). *Lehrereinstellungsbedarf und -angebot in der Bundesrepublik Deutschland 2014–2025–Zusammengefasste Modellrechnungen der Länder*. Downloaded 2015-08-14 at http://www.kmk.org/fileadmin/pdf/Statistik/Dokumentationen/Dok_208_LEB_LEA_2015.pdf
- KMK (2015g). *Qualitätssicherungen in Schulen*. Downloaded 2015-08-30 at <https://www.kmk.org/themen/qualitaetssicherung-in-schulen.html>
- KMK (no date). *Aufgaben der Kulturministerkonferenz*. Downloaded 2015-08-02 at <https://www.kmk.org/kmk/aufgaben.html>
- Krautz, J. (2007). *Pädagogik unter dem Druck der Ökonomisierung. Zum Hintergrund von Standards, Kompetenzen und Modulen. Pädagogische Rundschau*, 1, pp. 81-93.

- Kuhlmann, C. (2012). *Bildungspolitik und Leistungsvergleichsstudien*. Wiesbaden: Springer VS
- Köller, O. (2013). Die Konstante Des Deutsche Bildungssystems. *Die Politische Meinung*, Konrad Adenauer Stiftung , Jahrgang 58 (Nr. 522), pp. 53-59.
- Köller, O. & Pant, H. (2010). Die Rolle von Bildungsstandards in einem System der Qualitätssicherung und Qualitätsentwicklung. Schaal, B. & Huber, F. (Eds.). *Qualitätssicherung im Bildungswesen* (pp. 55-67). Münster: Waxmann
- Lankes, E-M. & Huber, F. (2014). *Leitfaden für die Erstellung von Zielvereinbarungen*. München: ISB
- Landtag Rheinland-Pfalz, (2015). *Drucksache 16/5675*. Downloaded 2016-08-02 at <http://www.landtag.rlp.de/landtag/drucksachen/5675-16.pdf>
- Lee, K. (2013). *Bourdieuian Analysis on African Americans underrepresentation at parks and outdoor recreation*. (Doctoral thesis, Texas A & M University). Ann Arbor: ProQuest LLC
- Lindblad, S. (2011). Turn taking in large scale reforms: "Scandinavian neo-institutionalism" as an example. Gogolin, I., Baumert, J. & Scheunpflug, A. (Eds.). *Transforming Education. Umbau des Bildungswesens* (pp. 67-82). Zeitschrift für Erziehungswissenschaft, 14 (Sonderheft 13).
- Lindgard, B. & Grek, S. (2008). *The OECD, Indicators and PISA: An exploration of events and theoretical perspectives*. Downloaded 2015-10-26 at http://www.researchgate.net/profile/Bob_Lingard/publication/241619070
- Lücken, M., Thonke, F., Pohlmann, B., Hofmann, H., Golecki, R., Rosendahl, J., Benzing, M. & Poerschke, J. (2014). KERMIT-Kompetenzen ermitteln. Fickermann, D. & Maritzen, N. (Eds.). *Grundlagen für eine daten- und theoriegestützte Schulentwicklung* (pp. 127-154). Münster: Waxmann
- Mauritzen, N. & Tränkmann, J. (2014). Zwischen Empirie und Normativität: Elemente eine Theorie des Bildungsmonitorings. Fickermann, D. & Mauritzen, N. (Eds.). *Grundlagen für eine daten- und theoriegestutzte Schulentwicklung* (pp 27-50). Münster: Waxmann
- Mazzeo, J., Lazer, S. & Zieky, M. J. (2006) Monitoring Educational Progress with Group-Score Assessments. Brennan, R-L. (Ed.). *Educational Measurement*, pp. 681-700. Westport, CT : Praeger Publishers
- Miles, M. B. & Huberman, A. M. (1994). *Qualitative Data Analysis*. (2nd edition). London: Sage Publications
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. (2008). *Kompetenzorientierung – Eine veränderte Sichtweise auf das Lehren und Lernen in der Grundschule*. Downloaded 2015-08-18 at http://www.schulentwicklung.nrw.de/lehrplaene/upload/klp_gs/LP_GS_Handreichung.pdf
- Muders, S. & Weishaupt, H. (2012). Personalsituation der Bildungsforschung. *Bildungsforschung* 37, pp. 155-166
- Münch, R. (2009). *Globale Eliten, locale Autoritäten, Bildung und Wissenschaft unter dem Regime von PISA, McKinsey & Co*. Frankfurt am Main: Suhrkamp Verlag

- Meyer, H-D. & Zahedi, K. (2014). *Open letter to Andreas Schleicher, OECD, Paris*. Downloaded 2016-02-16 at <https://bildung-wissen.eu/fachbeitraege/basistexte/open-letter-to-andreas-schleicher-oecd-paris.html>
- Nickel-Bacon, I. (2006). Positionen der Literaturdidaktik. Groeben, N. & Hurrelmann, B. (Eds.). *Empirische Unterrichtsforschung in der Literatur- und der Lesedidaktik* (pp 95-114). München: Weinheim
- Niessen, M. (2011). Building structures in research funding. Gogolin, I. & Baumert, J. (Eds.). *Transforming Education. Umbau des Bildungswesen* (pp 161-169). Zeitschrift für Erziehungswissenschaft. Sonderheft 13.
- Novoa, A. & Yariv-Mashal, T. (2003). Comparative Research in Education: a mode of governance or a historical journey? *Comparative Education*, 39 (4), pp. 423-438.
- OECD. (1996). *The Knowledge Based Economy*. Downloaded 2015-10-12 at <http://www.oecd.org/sti/sci-tech/1913021.pdf>
- OECD. (2005). *Teachers Matter. Attracting, Developing and Retaining Effective Teachers*. Downloaded 2015-08-17 at <http://www.oecd.org/edu/school/34990905.pdf>
- OECD. (2014). *Education at a Glance 2014. OECD Indicators*. Downloaded 2015-08-27 at <http://www.oecd.org/edu/Education-at-a-Glance-2014.pdf>
- OECD. (2015). *Evaluations and assessment frameworks for improving school*. Downloaded 2015-01-11 at <http://www.oecd.org/education/school/46927511.pdf>
- OECD. (no date/a). *PISA-The OECD Programme for International Student Assessment*. Downloaded 2015-10-13 at <http://www.oecd.org/pisa/pisaproducts/37474503.pdf>
- OECD. (no date/b). *PISA Sample Questions*. Downloaded 2015-08-02 at <http://pisa-sq.acer.edu.au/>
- OECD (no date/c). *PISA 2012. Creative Results. Creative problem Solving: Student's skills in tackling real life problems*. Downloaded 2015-08-06 at <http://www.oecd.org/pisa/keyfindings/pisa-2012-results-volume-v.htm>
- OECD (no date/d). *OECD work on Education & Skills*. Downloaded 2015-12-12 at <http://www.oecd.org/edu/Directorate-for-education-and-skills-brochure.pdf>
- OECD-CERI, no date. *International Indicators of Education Systems (INES)*, pp. 2-3. Downloaded 2015-12-24 at <http://www.oecd.org/edu/ceri/50234306.pdf>
- Oelkers, J. (2011). *Inklusion als Aufgabe der öffentlichen Schule?* Downloaded 2015-10-08 at <http://www.ife.uzh.ch/research/emeriti/oelkersjuergen/vortraegeprofoelkers/vortraege2011/BremenInklusionSchule.pdf>
- Overesch, A. (2007). *Wie die Schulpolitik ihre Probleme (nicht) löst: Deutschland und Finnland im Vergleich*. Münster: Waxman Verlag
- Ozga, J., Dahler-Larsen, P., Segerholm, C. & Simola, H. (Eds.). (2011). *Fabricating Quality in Education*. London: Routledge

Pant, H. A. (2014). Aufbereitung von Evidenz für bildungspolitische und pädagogische Entscheidungen: Metaanalysen in der Bildungsforschung. *Zeitschrift für Erziehungswissenschaft*, vol. 17 (4) , pp 79-99.

Papadoupoulos, G. S. (1994). *Education 1960-1990 The OECD Perspective*. OECD

phv, (2012). *Positionspapier-Wozu brauchen wir eine Agentur für Qualitätssicherung, Evaluation und Selbstständigkeit von Schulen (AQS), die keine inhaltliche Qualität eruiert?* Downloaded 2016-05-04 at [http://www.philologenverband.de/diverses/texte/?tx_ttnews\[tt_news\]=318&cHash=c911a7e3b74310f7a2291b907bf01256](http://www.philologenverband.de/diverses/texte/?tx_ttnews[tt_news]=318&cHash=c911a7e3b74310f7a2291b907bf01256)

Perkins, J. A. (1968). The International Conference on the World Crisis in Education: Summary Report of the Conference Chairman. *Journal of Educational Administration*, vol. 6 (1), pp. 3-12.

PISA Ergebnisse. (2013). *PISA und IGLU*. Downloaded 2015-01-12 at <http://www.tresselt.de/pisa.htm>

Piopiunik, M. & Wößmann, L. (2011). *Folgekosten unzureichender Bildung durch entgangenes Wirtschaftswachstum*. *Wirtschaftsdienst*, vol 91 (1), pp. 34-41.

Podium (2000). *Bayerischen Schulen auf dem Weg*. Downloaded 2015-05-10 at <http://www.fundacionbertelsmann.org/cps/rde/xbcr/SID-74FE54BE-739B366B/bst/podiumSchule2-2000sonder.pdf>

Prenzel, M. (2006). Wie solide ist PISA? Oder Ist die Kritik von Wuttke begründet? Downloaded 2016-03-24 at http://archiv.ipn.uni-kiel.de/PISA/Wie_solide_ist_PISA.pdf

Prenzel, M., Baumert, J. & Klieme, E. (2008). *Steuerungswissen, Erkenntnisse und Wahlkampfmunition: Was liefert die empirische Bildungsforschung? – Eine Antwort auf Klaus Klemm*. Downloaded 2015-11-16 at http://www.pedocs.de/volltexte/2011/1489/pdf/replikklemm_vollversionD_A.pdf

QUA-LiS. (2016). *Qualitäts- und UnterstützungsAgentur-Landesinstitut für Schule*. Downloaded 2016-09-03 at <http://www.qua-lis.nrw.de/>

Radtke, F-O. (2015). Methodologischer Ökonomismus – Organische Experten im Erziehungssystem. *Erziehungswissenschaft*, Heft 50, 2015/1, pp. 7-16.

Raither, J., Dollinger, B. & Hörmann, G. (2009). *Einführung Pädagogik*. Berlin: Springer Verlag

Reinders, H., Ehmann, T., Post, I. & Niemack, J. (2014). *Stress: Stressfaktoren bei Eltern und Schülern am Übergang zur Sekundarstufe*. Downloaded 2015-09-21 at <http://www.jugendforschung.de/pdf/SEB-33.pdf>

Robson, C. (2005). *Real World Research: A Resource for Social Scientists and Practitioner-Researchers* (2nd ed.). Oxford: Blackwell Publishing

Rotte, R & Rotte, U. (2005). *Discussion Paper nr 13*. RWTH: Institut für Politische Wissenschaften

Rotte, R & Rotte, U. (2007). Recent Education Policy and School Reform in Bavaria: A Critical Overview. *German Politics*, vol 16 (2), pp. 292-313.

Saha, L & Tuijnman, A. (2003). Comparative Indicators in Education. Keeves, J. P. & Watanabe, R. (Eds.), *International Handbook of Educational Research in the Asia-Pacific Region* (pp. 1123-1136). Netherlands: Kluwer Academic Publishers

Schmidtlein, A. F. (2010). Assumptions commonly underlying government quality assessments practices. *Tertiary Education and management*, vol. 10 (4), pp. 263-285.

Schneider, S. (2006). Qualitätssicherung an Bayerns Schulen. ISB. *Dokumentation zum Fachkongress am 25./26. November in Eichstätt* (pp. 11-25). München: Bayerisches Staatsministerium für Unterricht und Kunst

Schneider, W., Preckel, F. & Stumpf, E. (2014). Schulische Förderung von Hochbegabten: Ergebnisse nationaler und internationaler Studien. Schneider, W., Preckel, F. & Stumpf, E. (Eds.). *Hochbegabtenförderung in der Sekundarstufe. Ergebnisse der PULSS-Studie zur Untersuchung der gymnasialen Begabtenklassen in Bayern und Baden-Württemberg*. Frankfurt am Main, pp. 10-20

Schulentwicklung in Bayern (no date). *Links. Stiftungen/Weitere*. Downloaded 2015-05-20 at <http://www.schulentwicklung.bayern.de/index.php?Seite=134&PHPSESSID=dedbb3913b48645fbbb9999fdcd0635a>

Schulz, T. (1971). *Investment in human capital. The role of education and research*. New York: The Free Press

Secolsky, C. & Denison, D. B. (2012). *Handbook on Measurement, Assessment and Evaluation in Higher Education*. New York & London: Routledge

Sekreteriat, KMK. (2008). *Anrechnung von außerhalb des Hochschulwesens erworbenen Kenntnissen und Fähigkeiten auf ein Hochschulstudium (II)*. Downloaded 2015-08-20 at http://www.akkreditierungsrat.de/fileadmin/Seiteninhalte/KMK/Vorgaben/KMK_Anrechnung_ausserhochschulisch_II.pdf

Sirvanci, M. B. (2004). Critical issues for TQM implementation in higher education, *The TQM Magazine*, vol. 16 (6), pp. 382-386.

Skolverket, 2014. *This is the Swedish National Agency for Education*. Downloaded 2014-12-29 at <http://www.skolverket.se/om-skolverket/andra-sprak-och-lattlast/in-english>

Spiel, C., Lüftenegger, M., Wagner, P., Schober, B. & Finsterwald, M. (2011). Förderung von Lebenslangem Lernen-eine Aufgabe der Schule. Zlatkin-Troitschanskaia, O. (Ed.). *Empirische Bildungsforschung-ein Überblick aus interdisziplinärer Perspektive* (pp. 305-319). Wiesbaden: VS Verlag

Staatliches Schulberatung in Bayern. (2009). *Das „neue“ Übertrittsverfahren in Bayern*. Downloaded 2015-10-13 at http://www.schulberatung.bayern.de/imperia/md/content/schulberatung/pdf/uebertrittsverfahren_2009.pdf

Stake, R.-E., Contreras, G. & Arbesú, I. (2012). Assessing the Quality of an University-Particularly its Teaching. Secolsky, C. & Denison, B. (Eds.). *Handbook on Measurement, Assessment and Evaluation in Higher Education* (chp. 1). New York & London: Routledge

Stanat, P., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Schümer, G., Tillmann, K.-J. & Weiss, M. (2002) *PISA 2000. Die Studie in Überblick*. Berlin: Max-Planck-Institut für Bildungsforschung

Statistischen Bundesamt (2012). *Schulen Auf Einen Blick. Ausgabe 2012*. Downloaded 2015-03-13 at https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/Broschuer_eSchulenBlick0110018129004.pdf?__blob=publicationFile

Statistischen Bundesamt (2014). *Schulen Auf Einen Blick. Ausgabe 2014*. Downloaded 2015-03-13 at https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/Broschuer_eSchulenBlick0110018149004.pdf?__blob=publicationFile

Stifterverband für die Deutsche Wissenschaft. (no date). *Ländercheck. Allgemeine Entwicklung der Drittmittel*. Downloaded 2015-06-17 at http://www.laendercheck-wissenschaft.de/drittmittel/drittmittel_allgemein/index.html

Stiftung Bildungspakt Bayern (2015). *Willkommen bei der Stiftung Bildungspakt Bayern*. Downloaded 2015-11-16 at <http://bildungspakt-bayern.de/>

Suskie, L. (2004). *Assessing student learning: A common sense guide*. Bolton, MA: Anker Publishing Company

Swartz, D. (1997). *Culture and Power. The sociology of Pierre Bourdieu*. Chicago: The University of Chicago Press

Swartz, D. (2000). Pierre Bourdieu: The Cultural Transmission of Social Inequality. Robbins, D. (Ed). *Pierre Bourdieu*. London: Sage Publication

Taylor, F.W. (1911). *The Principles of Scientific Management*. Downloaded 2015-01-08 at <http://wps.pearsoncustom.com/wps/media/objects/2429/2487430/pdfs/taylor.pdf>

The Bavarian Ministry of Economic Affairs and Media, Energy and Technology (2014). *Bavaria A Guide for Investors*. Downloaded 2015-03-04 at http://www.invest-in-bavaria.com/fileadmin/media/documents/publications/Investorsguide/Investorsguide_Bavaria_2014.pdf

Tillman, K.-J., Dederich, K., Kneuper, D., Kuhlmann, C. & Nessel, I. (2008). *PISA als bildungspolitisches Ereignis: Fallstudien in vier Bundesländer*. Wiesbaden: VS Verlag für Sozialwissenschaft.

TIMSS (1997). *Mathematisch naturwissenschaftlicher Unterricht in internationalen Vergleich*. Downloaded 15-06-16 at <http://www.staedel.de/ressourcen/TIMSSII-Broschuere.pdf>

Torrance, A. (2012). Triangulation, Respondent Validation, and Democratic Participation in Mixed Methods Research. *Journal of Mixed Methods Research*, 6 (2), pp. 111-123.

Tröhler, D. (2013). The OECD and Cold War Culture: thinking historically about PISA. Meyer, H.-D. & Benavot, A. (Eds.). *PISA, Power, and Policy y: the emergence of global educational governance* (pp. 141-162). Oxford: Symposium Books

TUM (2015). *Susanne Klatten Endowed Chair for Empirical Educational Research*. <http://www.ebf.edu.tum.de/en/team/>

Varjo, J., Simola, H. & Rinne, R. (2013). Finland's PISA Results: an analysis of dynamics in education Politics. Meyer, H-D. & Benavot, A. (Eds.). *PISA, Power, and Policy y: the emergence of global educational governance* (pp. 51-76). Oxford: Symposium Books

VBE (2015). *Deutsches Verbände Forum/Mainauer Erklärung von VBE (Deutschland), GÖD-APS (Österreich), LCH (Schweiz) / Mehr finanzielles Engagement für Schule tut not.*

Downloaded 2015-08-17 at <http://www.verbaende.com/news.php/Mainauer-Erklaerung-von-VBE-Deutschland-GOeD-APS-Oesterreich-LCH-Schweiz-Mehr-finanzielles-Engagement-fuer-Schule-tut-not?m=103040>

Veld, R., Fussel, H-P. & Neave, G.(1996). *Relations between state and higher education*. The Hague/London/Boston: Kluwer Law International

VERA (no date). *Projektgruppe VERA am zepf*. Downloaded 2015-11-23 at <http://www.uni-landau.de/vera/>

Webb, J., Schirato, T. & Danaher, G. (2002). *Understanding Bourdieu*. London: Sage Publications

Weiniger, E. B. (2005). Foundation of Pierr Bordieu's class analysis. Wright, E. O. (Ed.). *Approaches to class analysis* (pp. 82-118). Cambridge: University Press

Weishaupt, H. (2008). Vorbemerkung. Weishaupt, F., Kraul, M., Böhm-Kasper, C., Schulzeck, U. & Zugenricker, C. (Eds). *Zur Situation der Bildungsforschung in Deutschlan* (pp.5-8) Bonn, Berlin: BMBF

Wittmann, E. C. (2010). *Lernstandserhebung VerA 3 Mathematik /2010*. Downloaded 2015-07-17 at <http://www.mathematik.uni-dortmund.de/ieem/mathe2000/pdf/vera3/KM%2005-10.pdf>

Wuttke, J. (2009). PISA: Nachträge zu einer nicht geführten Debatte. *Mitteilungen der Gesellschaft für Didaktik der Mathematik*, 87, 22-34.

Jahnke, T. & Meyerhöfer, W. (2007). (Eds.). *PISA & Co. Kritik eines Programms*. Fraunzbecker: Hildesheim

ZIB (2012a). *About the ZIB*. Downloaded 2015-03-12 at <http://zib.education/en/ueber-das-zib.html>

ZIB (2012b). *Kooperation mit dem IQB*. Downloaded 2015-03-11 at <http://zib.education/ueber-das-zib/kooperation-mit-iqb.html>

ZIB (2012c). *Research*. Downloaded 2015-03-11 at <http://zib.education/en/forschung.html>

Zlatkin-Troitschanskaia, O. & Gräsel, C. (2011). Empirische Bildungsforschung-ein Überblick aus interdisziplinärer Perspektive. Zlatkin-Troitschanskaia, O. (Ed.). *Stationen Empirischer Bildungsforschung* (pp. 9-20). Wiesbaden: VS Verlag

Zugenrucker, I. & Weishaupt, H. (2008). Finanzierung der Bildungsforschung. Weishaupt, F., Kraul, M., Böhm-Kasper, C., Schulzeck, U. & Zugenrucker, I. (Eds). *Zur Situation der Bildungsforschung in Deutschland, Band 28* (pp.23-34). Bonn, Berlin: BMBF

Secondary literature/sources

Känguru der Mathematik (no date). *Känguru der Mathematik-das ist*. Downloaded 2015-10-11 at <http://www.mathe-kaenguru.de/>

Tenorth, H-E. (2011). *Zehn Jahre PISA. Bilanz und Ausblick mit Professor Tenorth. Die Probleme Drittens*. Downloaded 2015-08-05 at https://www.youtube.com/watch?v=wcjgkBS_OvA

Save Our Schools NZ. (2014). Downloaded 2015-08-30 at <http://saveourschoolsnz.com/category/lead-expert-teachers/>

Sundell, P. (2014). *Who's on first. Teacher professionalism in the wake of the förstelärare-reformen*. Master Thesis. Downloaded 2016-05-10 at <https://gupea.ub.gu.se/handle/2077/37103>

Zweites Positionspapier des EMSE-Netzwerkes. (2008). Downloaded 2015-07-17 at http://www.emse-netzwerk.de/uploads/Main/EMSE_Positionsp2_Rueckmeldungen.pdf

Articles

Allgemeine Zeitung, 2015. Downloaded 2016-05-04 at http://www.allgemeine-zeitung.de/politik/rheinland-pfalz/schul-tuev-in-rheinland-pfalz-zum-schuljahr-201617-aufgeloest-cdu-und-lehrerverbaende-triumphieren_16152748.htm

Augsburger Allgemein. (2015). *Vier von fünf Abiturienten in Bayern beginnen ein Studium*. <http://www.augsburger-allgemeine.de/bayern/Vier-von-fuenf-Abiturienten-in-Bayern-beginnen-ein-Studium-id33103057.html>

Bildungsklick (2015). *Pressemitteilung. Anmeldungen am Gymnasium leicht rückläufig - Übertrittsquote stagniert*. Downloaded 2015-09-18 at <http://bildungsklick.de/pm/73894/anmeldungen-am-gymnasium-leicht-ruecklaeufig-uebertrittsquote-stagniert/>

Spiegel. (2001). *Mangelhaft. Setzen*. Downloaded 2015-02-24 at <http://www.spiegel.de/spiegel/print/d-20960899.html>

Der Tagesspiegel (2012). *Humboldts Risiko*. <http://www.tagesspiegel.de/wissen/rechnungshof-ruegt-die-hu-humboldts-risiko/6705578.html>

Süddeutsche Zeitung. (2005). *Geisteswissenschaften in Gefahr*. Downloaded 2015-10-01 at <http://www.sueddeutsche.de/karriere/sparkurs-in-bayern-geisteswissenschaften-in-gefahr-1.550944>

Süddeutsche Zeitung. (2013). *Warum Bayern nicht mehr Bildungsprimus ist*. Downloaded 2015-03-26 at <http://www.sueddeutsche.de/bildung/schulleistungsvergleich-der-bundeslaender-warum-bayern-nicht-mehr-bildungsprimus-ist-1.1792491>

The Guardian. (2014). *A poor Ofsted report could lead to headteachers being 'disappeared'*. Downloaded 2015-09-23 at <http://www.theguardian.com/education/2014/mar/11/heads-poor-ofsted-report-dismissal-shortages>

WAZ. (2015). Downloaded 2015-07-17 at <http://www.derwesten.de/politik/lehrer-wollen-vera-tests-fuer-drittklaessler-abschaffen-id9312801.html>

Appendix 1

PISA-a reborn programme

The PISA studies have had a considerable and confirmed impact on educational systems globally. It is provided by the Educational Committee (OECD/CERI) and it is a longitudinal study conducted since 2000. The Programme for International Student Assessment (PISA) is a triennial international survey which aims to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students. To date, students representing more than [70](#) economies have participated in the assessment. (OECD, 2014)

PISA is coordinated by the governments of participating countries under the auspices of the OECD. Samples of 15-year old students are assessed in three main subjects- reading, mathematics and science- based on a matrix sampling approach and students and teacher also have an additional questionnaire to fill in. The test is said to measure not just academic but also the ability to use knowledge in practice (Mazzeo, Lazer, and Zieky, 2006, p. 683). The PISA studies are a part of an evaluative idea that different educational systems can be compared if just corresponding and reliable statistic, (indicators) of the performance of each educational system can be obtained. The abbreviation PISA was first used in the Learning Sciences Programme implemented in 1973 by CERI. It was part of OECDs early work to consider the feasibility of obtaining educational performance indicators., It was *‘designed to foster the development of research on factors affecting formal and informal learning, taking into account the full range of relevant research in the social and biological sciences’* (Papadopoulos, 1994, p 88; CERI, 1978, p 18). But the term *relevant* proved to be highly problematic because it meant to sort out, and grade, sciences relevant for education and involved the questions of who had the authority base over the knowledge and who regulates the value structures in education. CERI never went through with the scheme, the program became highly controversial among the member states depending on the questions of scientific relevance and the values behind that it raised, and was discontinued already in 1977 (Papadopoulos, 1994, p. 89; CERI, 1978, p. 28). However, in 2000 it reappeared in modified guise.

Appendix 2

Example of questionnaire

Fragestellungen-

Allgemeines über das Bayrische Studiensystem

Geschichtliche Übersicht
Beschreibung der aktuellen Studiensysteme
Dessen Besonderheiten / besonders bezeichnend für Bayern

Aus welchem Grund fasste Bayern den Entschluss zur Teilnahme an Pisa 2000 ?

Eventuelle Erfahrungen von anderen Studien
Einigkeit oder unterschiedliche Ansichten für ein Teilnehmen (politisch, Verwaltung)
Welche Personen waren involviert (bzw. hatten Einfluss)

Erwartungen und Gedanken zur Teilnahme im PISA-Projekt.

Das große Ansehen des bayrischen Studiensystems
Eventuelle Erfahrungen von anderen Studien
Welche Vor-oder Nachteile diskutierte man?

PISA 2000,praktische Erfahrungen

Administrative Erfahrungen
Welche Teile der Auswertung /„Messung“ gingen gut/weniger gut
Erwartete/unerwartete Dilemma

Resultat des Projektes PISA 2000

Ansichten betreffend Feed-back der PISA-I,PISA-E Resultate
Wichtige Teile die man besonders beachtete/analyisierte
Positive resp. negative Ansichten
Wie war der Effekt auf die Debatte ueber die Messungen

PISA `s Teilaspekte

3-Jahresintervall : angemessen oder nicht
Änderungen in der Ausführung der Messungen im Laufe der Jahre
Was misst PISA generell

Allgemeiner Teil

Erfahrungen von der Teilnahme an solchen internationalen Auswertungen wie PISA.
Kontrolle und Manipulation der Ausbildungssysteme, eventuelle Veränderungen während der Jahre (2000-)
Erweiterung und Verminderung von Bedeutung für den Test und dessen Auswertung im Ausbildungssystem
Haben sich die Ansichten über die internationalen Messungen auf Bund-resp. Länderniveau während der Jahre geändert?
Gibt /gab es Akteure die mit fortschreitenden PISA-Studien an Einfluss gewann/ verlor

Appendix 3

Three forms of all-day-schools in Germany

In Germany there are three kinds of all-day schools: (1) Bound all-day school where the students for three days a a week minimum, must partake in the activities for at least 7 hours, (2) partial bound all-day school where a part of the students for three days a a week minimum, must partake in the activities for at least 7 hours, and (3) open all-day school where the students voluntarily can attain the activities on the afternoon (Tillmann et al. 2008, p. 185). In Bavaria the percentage in 2012/2013 of the Bavarian students in all-day schools were 12,4 %, the lowest number of all German Länder and a majority were in open all-day schools (Klemm, 2014, pp. 15-16). In spite of the declared promotion of this specific educational reform agenda and a massive funding as a part of a large national reform initiative program, IZBB (*Investitionsprogramm Zukunft, Bildung und Betreuung*⁹²) that included all-day schooling in the Länder in accordance with the PISA 2000 results (Tillmann et al., 2008, pp. 190-191). However, CSU instead used the funding to another educational reform program, the G8 which stand for the shortening of the Gymnasium from 9 year to 8 year, another reform in which also was a consequence of the PISA results (e.g. Friedrich Ebert Stiftung , 2009, p. 3). In 2013 the Bavarian education minister Ludwig Spanlae launched a massive expansion plan of all-day schools in Bavaria but the majority were to be in the form of open (Bayerische Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, 2013b). Hence, the education reform of all-day schooling in Bavaria has been influenced by the larger German context where the KMK has stated the promotion of all-day schooling and a number of Länder has significantly expanded these school forms. Nevertheless the political dominant in Bavaria has been able to encompass and adjust the effect in the Bavarian educational discourse, individual development as opposed to equality of opportunity, with altered content, open all-day schools in line with traditional CSU family values, and with substantial delay (ibid.)

⁹² The Investmentprogram Future, `Bildung` and Care.