



UiT The Arctic University of Norway

Faculty of Humanities, Social Sciences and Education

'PISA effects' on Norwegian education

A critical realist perspective

Terje André Bringeland

A dissertation for the degree of philosophiae doctor – May 2023



Thesis submitted for the degree of philosophiae doctor (ph.d.)

‘PISA effects’ on Norwegian education
A critical realist perspective

By

Terje André Bringeland

UiT - The Arctic University of Norway

Faculty of Humanities, Social Sciences and Education

Acknowledgements

I am grateful to the University in Tromsø (UiT), the Faculty of Humanities, Social Sciences and Education and the Department of Education for giving me the opportunity to earn a doctoral degree in comparative education related to educational measurements and politics. I am especially grateful to my supervisor, Professor Tone Skinningsrud, for dedicated supervision. Moreover, thanks to Tone for introducing me to Professor Margaret Archer and her work, and letting me do my work in the spirit of Bhaskarian realism. I also want to thank my second supervisor, Associate Professor Jens Breivik, for feedback on this doctoral thesis and for organising evaluations and meetings. I also want to say thank you to my mid-term evaluator Associate Professor Lene Nyhus and end-term evaluator Professor Eyvind Elstad, for taking the time to give me feedback on this thesis.

I further want to thank the research group Space and Time in Education (STED) for the collaboration and discussions related to our anthology, led by Professor Unn-Doris Karlsen Bæck. I also want to thank the Norwegian National Research School in Teacher Education (NAFOL), at the Norwegian University of Science and Technology (NTNU), for letting me be a part of their programme. These communities have given me the opportunity to meet other academics and explore other academic institutions. The national and local research cohorts have generated further reflections and development in my research process.

I want to thank Professor Gry Paulgaard for a professional cooperation in teaching courses at University of Tromsø and my fellow doctoral students Inger Lise, Siv, Daniel, Anna-Maria, Lise-Mari, Oda, Anna, Johannes and Sibylle, for all of the academic and non-academic discussions and meetings throughout this period. A special thanks to Anna-Maria and Daniel for translating the Internal Conversation Indicator (ICONI) to Norwegian.

Preface

The relationship between structure and agency has taken a central place in sociology for decades. Is structure determining agency, or the other way around, or do they meet in the middle? If so, how is social change explained? The answer is that social change cannot be explained based on these conflated logics. This is the reason why Margaret S. Archer has stated that '[t]he problem of structure and agency [...] denote central dilemmas in social theory – especially the rival claims of Voluntarism versus Determinism, Subjectivism versus Objectivism, and the micro- versus the macroscopic in sociology.' (Archer [1988] 1996, xi). The way to explain change is to acknowledge that '[...] structure and agency constitute different levels of the stratified social reality, each possesses distinctive emergent properties which are real and causally efficacious but irreducible to one another.' (Archer 1995, first page). Therefore, Archer's morphogenetic cycle and conceptualisations of the educational system and agency are crucial constructs of real phenomena, that enables studying and explaining change or reproduction.

Table of contents

Acknowledgements	v
Preface	vii
Table of contents	ix
List of tables	xiii
List of figures	xv
Summary in English	xvii
Summary in Norwegian	xix
1. Introduction	1
1.1 Background for studying PISA effects.....	1
1.2 PISA – a decisive cause for effects?	2
1.3 What is PISA?	3
1.3.1 A culturally neutral non-curricular achievement test.....	7
1.3.2 A policy instrument (with flaws)	9
1.3.3 Entangled with global edu-business (the global education industry – GEI).....	16
1.4 Norwegian PISA researchers assessing ‘PISA effects’ and future prospects for international large-scale tests	17
1.5 A contribution to the (enlarged) field of PISA effects	20
1.6 A gateway within the PISA effects field.....	25
1.7 Focus on two social forms and a reconceptualisation of causality	28
1.8 Reconnecting the reconceptualisation on the two social forms in the Norwegian case	30
1.9 Aim, overarching research questions, and the scope of the research	31
1.10 Thesis outline	32
1.11 Summary	33
2. State of the art: macro-level and micro-level research related to PISA effects	36
2.1 The literature search and sectioning.....	36
2.2 State of the art in macro-level research on PISA effects.....	37
2.2.1 Pons’ review of international PISA research	37
2.2.2 Hopfenbeck’s review of international PISA research (including Norway)	38
2.2.3 Previous research on ‘PISA effects’.....	40
2.2.3.1 International comparisons of PISA-effects	40
2.2.3.2 PISA effects in Norway.....	41

2.2.3.2.1	Elaborations on findings of PISA effects in Norwegian studies	42
2.2.3.2.2	Summing up: the need for a conceptualisation of the educational system	47
2.3	State of the art in micro-level research related to PISA effects	48
2.3.1	Other findings on micro-level PISA effects in Norway	50
2.3.2	The quest for a concept of (teacher) agency (with properties).....	55
2.4	Summary	59
3.	Meta-theoretical underpinnings: critical realism.....	60
3.1	Critical realism	60
3.2	Intransitive and transitive dimensions of reality	62
3.3	Three domains of reality	63
3.4	Generative mechanisms.....	64
3.5	Tendencies, causality, and effects in open systems	64
3.6	Power1 and power2 relationships	66
3.7	Reasons as causes.....	66
3.8	Abduction and retroduction.....	67
3.9	Summary	68
4.	The key conceptual framework.....	72
4.1	A system theory and the causal agent	72
4.1.1	Morphogenetic cycles	72
4.1.2	The concept of the educational system	73
4.2	Educational systems: centralized and decentralized, structures and processes.....	75
4.2.1	Unification.....	75
4.2.2	Systematization	76
4.2.3	Differentiation	76
4.2.4	Specialization	78
4.3	The concept of agency: personal and social identity.....	80
4.3.1	The human ‘self’ and ‘agency’ in different theoretical traditions.....	80
4.3.2	Agency, personal identity, and ultimate concerns.....	82
4.3.3	Reflexivity – the internal conversation	82
4.3.4	Modes of reflexivity	84
4.3.4.1	Communicative reflexives.....	85
4.3.4.2	Autonomous reflexives	85

4.3.4.3	Meta-reflexives.....	85
4.3.4.4	Fractured reflexives.....	86
4.4	The causal agent with ideas and materials	86
4.5	Research questions: The Norwegian case	87
4.6	Summary	89
5.	A case study framework	91
5.1	Case study research	91
5.2	Design and context.....	91
5.3	Evidence and analysis	93
5.4	Data collection and piloting	95
5.5	Research ethics.....	98
5.6	Validation and generalisation.....	100
5.7	Critical remarks on the theory, design, and methods	102
5.8	Summary	104
6.	Results and analytical discussions	106
6.1	A specific reply	106
6.2	The overarching research question I	106
6.2.1	Discussing findings about structures and processes in the Norwegian educational system before and after the ‘PISA shock’	107
6.2.2	Subsidiary research questions - article I: The Impact of PISA on Education in Norway: A Morphogenetic Perspective on Structural Elaboration in an Education System 115	
6.2.3	A short comment	118
6.3	The overarching research questions II	118
6.3.1	Subsidiary research questions - article II: PISA and Teachers’ Reflexivities. A Mixed Methods Case Study	120
6.3.2	Subsidiary research questions - article III: School Leaders’ Reflexive Mode in their Internal Conversations on PISA.....	121
6.3.3	Discussing previous Norwegian micro-level findings with PISA before and after the use of reflexivity and reflexive modes	122
6.4	The key factors and variables.....	124
6.5	Summary	125
7.	The thesis’ contributions to the PISA research	127
7.1	Reconceptualising ‘PISA effects’	127

7.1.1	Macro-level ‘PISA effects’	127
7.1.2	Micro-level ‘PISA effects’	128
7.1.3	Macro-level and micro-level ‘PISA effects’	128
7.2	Empirical contributions	129
7.3	Methodological contributions	130
7.4	Theoretical contributions.....	131
7.5	Strengths and limitations of the research	132
7.6	Ideas for further research.....	134
7.7	Summary	135
	Works cited	137
	Appendixes.....	147
	Appendix A: The Norwegian Centre for Research Data (NSD) approval	147
	Appendix B: Informed letter of consent—mathematics teacher	149
	Appendix C: The interview guide with the mathematics teacher	154
	Appended papers.....	158
	Article I: The Impact of PISA on Education in Norway: A Morphogenetic Perspective on Structural Elaboration in an Education System.....	158
	Article II: PISA and Teachers’ Reflexivities. A Mixed Methods Case Study	195
	Article III: School Leaders’ Reflexive Mode in their Internal Conversations on PISA	233

List of tables

Table 1: Schematised overview over the macro-level theoretisations related to PISA effects.	47
Table 2: Schematised overview over the micro-level theoretisations related to PISA effects.	55
Table 3: The Three Domains of Reality. Source: Bhaskar ([1975] 2008, 13).	63

List of figures

Figure 1: The Basic Morphogenetic/Static Cycle with its Three Phases. Source: Archer (1995, 157).....	73
Figure 2: Illustration of the Internal Systemic Structure of Centralized Educational Systems. Source: Author's illustration..	75
Figure 3: The Structural Conditioning of Educational Interaction in the Centralized System. Source: Archer ([1979] 2013, 266, adjusted illustration by the author).....	79
Figure 4: The Structural Conditioning of Educational Interaction in the Decentralized System. Source: (Archer [1979] 2013, 267, adjusted illustration by the author).....	80
Figure 5: The Hierarchy of the Overarching and Subsidiary Research Questions. Source: Author's illustration.	89

Summary in English

The objective of the thesis is to examine, from a critical realist perspective and by using Margaret Archer's conceptual framework, possible effects that the Programme for International Student Assessment (PISA) test has on the Norwegian education system and on school personnel. The intention has been to reconceptualise the notion of 'PISA effects' both at the systemic (macro) and individual (micro) level.

The thesis is based on a total of five cases (art. I = 1, art. II = 1, art. III = 3) that represent entities at different levels: the national education system (the macro-level); an individual mathematics teacher (the micro-level); and school leaders (micro-level). Two of the articles employ qualitative methodology (art. I and art. III), while article II employs both a qualitative and quantitative methodology, i.e., method triangulation. The data consist of relevant policy documents and secondary literature (art. I), and semi-structured interviews (art. II, art. III) with a mathematics teacher (n=1) and three school leaders (n=3), the former in combination with a Likert scale measuring modes of reflexivity.

My investigation of educational reforms and other state interventions for two decades, since the beginning of the 1990s, before and after the PISA-shock (art. I) indicated that basic system characteristics, internal structures, and processes in the system, were maintained. Contrary to what some previous research has claimed, the PISA shock did not fundamentally change the Norwegian educational system. This analysis employs Archer's definition of educational systems, centralized and decentralized, her specification of internal structures and processes in state educational systems, and her model of morphogenetic cycles.

The four micro-level case studies presented in article II and III, employ Archer's concepts reflexivity and reflexive modes, which can explain differences among school personnel (teachers and school leaders) in how they react to the PISA test. The use of reflexivity and reflexive modes as analytical concepts enables the alignment of teaching staffs' personal concerns and values with their reactions to the PISA test and shows how agents mediate the effects of external systemic structures.

Picking up on Xavier Pons quest for a new theoretical framework for the study of PISA effects, which could establish it as a 'normal science', I argue that critical realism and Archer's conceptual framework satisfies his basic requirements, i.e., ontological realism, epistemic relativism, and judgemental rationality.

Summary in Norwegian

Målet med oppgaven er å undersøke, fra et kritisk realistisk perspektiv og ved å bruke Margaret Archers konseptuelle rammeverk, mulige effekter som PISA-testen har på det norske utdanningssystemet og på skolepersonell. Hensikten har vært å rekonseptualisere forestillinger om 'PISA-effekter' både på systemisk (makro) og individuelt (mikro) nivå.

Arbeidet er basert på totalt fem case-studier (art. I = 1, art. II = 1, art. III = 3) som representerer enheter på ulike nivåer: det nasjonale utdanningssystemet (makronivået), en individuell matematikk lærer (mikronivå) og tre skoleledere (mikronivå). To av artiklene benytter kvalitativ metodikk (art. I og art. III), og artikkel II benytter både en kvalitativ og kvantitativ metode, dvs. metodetriangulering. Dataene består av relevante policydokumenter og sekundærlitteratur (art. I) og semistrukturerte intervjuer (art. II, art. III) med en matematikklærer (n=1) og tre skoleledere (n=3). Førstnevnte i kombinasjon med en Likert-skala som måler refleksivitetsmoduser.

Min undersøkelse av utdanningsreformer og andre nasjonale tiltak i skolen i løpet av to ti-år, fra begynnelsen av 1990-årene, det vil si før og etter PISA-sjokket (art. I), viser at de grunnleggende egenskapene ved utdanningssystemet, forble de samme, til tross for PISA-begrunnede tiltak. Analysen anvender Archers definisjon av utdanningssystemer, sentraliserte og desentraliserte, begreper knyttet til interne strukturer og prosesser og dessuten hennes modell for morfogenetiske sykluser.

De fire case-studiene fra mikro-nivået (matematikklærer og skoleledere) i artikkel II og III anvender Archers begreper om refleksivitet og refleksive moduser, som er egnet til å forklare hvorfor enkeltpersoner har ulike reaksjoner, selv om de befinner seg i den samme strukturelle konteksten og hvordan disse reaksjonene henger sammen med personenes personlige verdioppfatninger og anliggender.

Jeg hevder at ønsket om et nytt teoretisk rammeverk for studiet av PISA effekter, fremsatt av Xavier Pons, som kan etablere feltet som en 'normalvitenskap', innfris av kritisk realisme og Archers konseptuelle rammeverk som baserer seg på prinsippene ontologisk realisme, epistemologisk relativisme og vurderingsrasjonalitet.

1. Introduction¹

1.1 Background for studying PISA effects

My original interest in PISA (Programme for International Student Assessment) was kindled by discussions about the test with my fellow students in the master programme in teacher education in Bergen. One thing we agreed upon was that we did not know much about how the test was perceived and experienced by school personnel. Moreover, after reading an article by Svein Sjøberg from the University of Oslo about the PISA test, I became even more gripped by curiosity about this topic. Sjøberg was undoubtedly writing from a critical stance but had not himself carried out any empirical studies at the educational meso or micro-level. When looking more closely into PISA research, it became obvious to me that this field of research was dominated by macro-level perspectives examining policy agents and national reforms.² I have been mostly interested in PISA research at the micro-level, since there was a lack of research conducted at that level and more had been called for (Sjøberg 2014a, c)³. As a former substitute (mathematics) teacher in compulsory school, I wanted to illuminate professional reactions to the PISA test. But gradually, as I became more known with Margaret Archer's theories, I understood that I could do research on the educational system too. I wanted thus to explore the possible effects of PISA on parts⁴ of and people in the Norwegian education system (the structure and its incumbents).

¹ On the website of the University of Tromsø (UiT), it is written that a synopsis should normally be 40–80 pages. This meta-text is longer, which appears permissible since the term 'normally' is used rather than a maximum number of pages (UiT n.d). The reason for this extension is that explanation, discussions, and transparency increase the number of pages. Another reason is that Archer's work is comprehensive to elaborate on as it is connected to different levels of theory.

² The notions of the macro-, meso- and micro-level of social analysis are related to different emergent strata in social reality. These terms are relational terms in the sense that a given stratum can be micro to another and macro to a third. In this context I consider structures and processes in individual schools the micro-level; the municipality, the school owner and decision maker for municipal schools the meso level; and national decision-making bodies the macro-level. For further discussion on the distinction between micro- and macro-level analyses, see (Archer 1995, 6–12).

³ For consistency with the published articles in this thesis, this synopsis has followed the Chicago Author-Date reference style 16th Edition. In this summary where there have been several of (the same) references in a paragraph frequently, I have used footnotes with the reference as a strategy to avoid too much reading interruptions. Further, I have used single apostrophes to signal citations or (key) phrases. Double apostrophes signal already written citations in books or articles, or work titles.

⁴ 'Parts' here refers to social structure, i.e., systemic structures.

An important reason for studying PISA effects in Norway is that international tests (including PISA) are justified by the fact that they must provide information to political authorities. PISA and other tests are legislated by law⁵ and it is compulsory for schools to participate in these tests when they are randomly chosen (NOU [Green Paper] 2023:1, 103). The results from these tests makes it possible to follow developments in the Norwegian schools; giving information to politicians, school owners, school leaders and teachers about the average competence or skill level of Norwegian students (NOU [Green Paper] 2023:1, 104). Thus, PISA is one test that provides information to political authorities and school personnel for school development. This increases the chances of there being (claimed) PISA effects in the Norwegian setting. For example, the effect of PISA on national testing since the turn of the millennium is described as severe: ‘the magnitude and impact of testing took a dramatic turn when the PISA tests entered Norwegian grounds in 2001’ (Nordkvelle and Nyhus 2017, 234). The two scholars argue that the national-cultural perspective was challenged by a global perspective on education (Nordkvelle and Nyhus 2017). A way to decode this, is that the PISA’s framework, focusing on literacies in some test domains, holds a relative strong global definition power over education based on economic interests; what students’ ought to learn for economic productivity and thereof development. This could trigger fundamental questions: what is a human? (cf. Solberg 2021), and what is schooling about? (cf. Biesta 2009).

1.2 PISA – a decisive cause for effects?

In the social reality causes are causal powers creating (an) effect(s) (Bhaskar [1975] 2008). The effect can occur in vivo or in vitro. As agency is the mediator of causal powers, besides being a causal power itself, will entail that effects of PISA can be exercised and not actualised, and exercised and actualised. This includes effects that are also exercised or actualised unperceived i.e., by other agents. Hence, effects of PISA can be within the agent, mediated through the agent, and actualised in material ways i.e., on various structures. The title of this thesis is «‘PISA effects’ on Norwegian education». This is an inclusive title for investigating PISA effects outside a system or within a system corresponding to former and forthcoming research. The title embraces different understanding of the concept causality i.e., different theorising and argumentation of PISA effects. Moreover, the title hints that ‘PISA effects’ are not caused by PISA alone, but with agency as the mediator. Hence, PISA effects

⁵ Regulation to the Education Act § 2-7

involve multi-causality. Therefore, the denotation ‘PISA effects’ involves co-causation, which is compatible with critical realist conceptualisation of causal power to cultural and structural objects (Bhaskar [1975] 2008). However, some causal powers will be more prominent than others for causing the effect. Thus, co-causing must not be construed as mutually constitutive of an effect or that the same causes produce the same effects on a regular law-like basis. However, due to PISA’s dispositional power it can produce certain tendencies e.g., how it is used among school personnel. In this thesis I will sometimes use apostrophes to ‘PISA effects’ beside referring to the field, to remind that PISA effects are not possible without agency, or to stress the role of PISA in causal questions.

In this thesis units, factors, structures are understood as synonyms, and so are variables, (generative) mechanisms and (emergent) properties. Both divisions can be understood as being causes, however, the latter would be more specific. Hence, when effects are established, they can revisit agency again as for example as memorisation (of ideas or actions), either by inner conversation per se, or by others influence, when engaging with PISA. For example, if you are critical about PISA, this will characterise your thinking and actions to some extent until you become more or less critical, either increasing or decreasing your engagement with PISA, due to realisation of other sides of the test. In other words, established effects on yourself or on a system can become new co-causes for change or reproduction, and as Archer (2000) outlines, this can manifest as various emotional commentaries within different reflexive agents: ‘I feel bad about this way of operating with PISA, it is not compatible with me anymore, so I need to change my direction with PISA’, or ‘PISA makes me happy, PISA is still compatible with my perspective and values, so I am continuing as before’, or ‘I am not happy with the PISA test, it troubles me, I take distance from it, but I do not have any choice than to conduct it’, or ‘what should I do with PISA, it so hard, last time I just conducted it, it makes me worry, it has positive and negative sides, I need my colleagues for deciding what to do with it’ and lastly ‘hm, PISA, I have not looked into it. I don’t know what it is. Maybe I should look into it’.

1.3 What is PISA?

PISA is an International Large-scale Assessment (ILSA) study of knowledge and skills in the domains of reading, mathematics, and natural science. It has been developed by the OECD (Organisation for Co-operation and Development), which is an intergovernmental organisation with 38 member countries (OECD n.d-a). Besides the US and European

countries, two Asian countries, South Korea and Japan, are members. In the Middle East, Israel is the only member, and in the South-Eastern hemisphere, New Zealand, and Australia. The combined GDP of the member countries constitutes two thirds of the world's total (MBN n.d).

Lundgren (2010, 42–43) put forward that the intentions with PISA can be traced back to the Cold War which evoked international competition. He points to the 'Space Race' where the Soviet Union had an advantage prior to the US. The Soviet Union had the first satellite in 1957 and the first dog Laika in space at the end of the year. After these two events, and a year later, the IEA (International Association for the Evaluation of Educational Achievement) emerged as a European initiative for comparing educational systems. The American President Kennedy response was that he promised that in the end of the decade in 1960 there would be a man on the moon. To achieve this goal the education system played a major role and needed strengthening. Especially, knowledge in mathematics and science were imperative. Human capital was seen as an economic prosperity. There was a need for an efficacious education system that also emphasised more knowledge about the students learning and development. The Organisation of Economic Co-operation and Development was established in 1961, mainly for stronger economic cooperation and productivity after the World War II (Elstad and Sivesind 2010; Bürgi 2019; Elfert 2019; Centeno 2019; Krejsler 2019).

Through the sixties school reforms developed as a response to larger competition with other countries. At this time, there were also a realisation of a teaching-technology with a material-method-system (input-output) with a certain material to be learned and taught, consequently learning material that followed the 'Space Race' between the Soviet Union and the US. Later, in the beginning of the 1980s neo-liberal thoughts from the UK and US dominated with a free market. This challenged the welfare society with a pressure on privatisation and competition. New political parties were also established, such as the Green Party which advocated for the environment and lifestyle-change in opposition to a more globalised world with increased production (Lundgren 2010, 47).

Between the 1970s and 1980s Ulf Lundgren was involved with CERIs Education Indicators Programme (INES), where the US played a large part (Lundgren 2010, 48). The US wanted comparative international statistics between the OECD-countries for surveillance and control of human capital. At that time, the numbers of countries participating in IEA had increased. In the 1990s result and goal management were implemented more heavily with a global market

economy. By this time, the IEA had no test that could provide trend analyses. But there were negotiations between the IEA and CERI (INES) for a new comparative test or analysis, without any success. This changed with a steering group in the 1990s, which included Ulf Lundgren, who decided that the new test should be independent and called Programme for International Student Assessment (PISA).

The PISA test was initiated in 199[6]/7 in Norway and is an OECD initiative to establish ‘international standard-setting’ in education (Elstad 2010, 63). It was administered for the first time in 2000 and has since been run every third year, except in 2021, due to the Covid pandemic. The last round was in 2022. The size of the Norwegian sample of students taking the test was approximately 8,500 from 271 schools for PISA 2022, consisting of randomly selected students from randomly selected schools (ILS n.d-b). Whole school classes are not included. PISA is a triennial test, but will become conducted each fourth year from 2025 (NOU [Green Paper] 2023:1, 105).

The PISA test itself takes about two hours to complete. PISA provides a snapshot of students’ knowledge and does not follow individual students over time. Alongside the test, an additional questionnaire is administered to the students, which takes about 50 minutes. It contains questions about students’ background, attitudes, learning strategies and the learning environment at school. A questionnaire is also distributed to the school management concerning educational leadership, quality assessment, learning environment and student diversity (ILS n.d-a). The PISA test was mainly computer-based from 2015 (OECD n.d-e, f).

In PISA, students are mainly tested on their ability to apply knowledge from certain ‘domains’ in real life situations, rather than their knowledge of school subjects. In each PISA test round, one of the three core domains is emphasised more than the other two. In addition to the three core domains, students may also be tested in the optional domains of financial literacy, collaborative problem solving, global competence, school organisation and creative thinking (OECD n.d-d). Upcoming optional domain in 2025 is learning in the digital world (OECD n.d-c). Over the years, the number of countries participating in PISA has increased. At the start, in 2000, there were a total of 43 (non-member and member) countries taking the PISA 2000 test (OECD n.d-g). In 2022, 83 different countries and economies are expected to take part, among these are 45 non-OECD countries and economies (OECD n.d-b).

Lingard (2020) names spin-off products of PISA from the OECD. The first was the PIAAC (Programme for the International Assessment of Adult Competencies) study targeting 16–65-year-olds, launched in 2003. Further, TALIS (Teaching and Learning International Survey) launched in 2008 for examination of school leaders and teachers working conditions and learning environments for lower secondary schools. PISA for Schools launched as a trial in 2011 and is a local school test on demand. Next, PISA for Development (PISA-D) was launched in 2013 for low- and middle-income countries targeting 14–16-year-olds in schools and out of schools. Moreover, PISA4U was launched in 2016 and consists of online learning modules and collaborative activities for teachers, and also giving credentials. The IELS (International Early Learning and Child Well-being Study) study is aimed for 5-year-olds and was launched in 2016. This test is often called ‘Baby-PISA’. The Teaching and Learning International Survey (TALIS Starting strong) for school leaders and teachers in kindergarten was launched in 2018.⁶ The SSES (Study of Social and Emotional Skills) study targets 10– and 15-year-olds and was launched as a trial in 2018. The OECD has also a Future of Knowledge and Skills 2030 project for curriculum standards. Consequently, the OECD is one of the leading institutions on testing and education governance, offering tests from pre-school children to seniors (Zhao 2020). PISA can be understood as a moving target; one that is being moved and moves (Komatsu and Rappleye 2021).

Arguably, PISA has changed in scope (what is being tested), scale (number of participating countries) and explanatory power (linking with other OECD data and establishing stronger explanations) (Lingard 2020, vi). Despite there being other spin-off products, the main PISA test has obtained a reputation as ‘the gold standard’⁷ or ‘the flagship’⁸ from the OECD among International Large-scale Assessments (ILSAs), of which there are many other tests.⁹ PISA is also called a ‘horse race’ where countries are competing to be on the top of the ranking list (Rutkowski and Rutkowski 2016). In the 2000s the reference society was Finland in the Nordic context, a country topping the first PISA cycles. Finland later fell in placement and the reference society changed to Shanghai (Sellar and Lingard 2013). Whether policy borrowing

⁶ (Utdanningsdirektoratet n.d)

⁷ (Sjøberg 2017, 17).

⁸ (Rutkowski and Rutkowski 2016, 252).

⁹ Trends in Mathematics and Science Studies (TIMSS), Progress in International Reading Literacy (PIRLS), and others. Though all of these are ILSAs, they have their own history and properties, and have been incorporated into the Norwegian educational system at different times, which suggests that they deserve their own examination.

from the ‘reference society’ actually occurs is disputed (Sivesind 2019), but top performing countries naturally attracts attention. PISA 2018, the latest PISA test to date, reveals that Singapore with Chinese provinces (Beijing, Shanghai, Jiangsu and Zhejiang) and one special administrative region (Macau), are among the top three on the ranking list (OECD 2019). PISA has become a key instrument for education governance (Ydesen 2019b; Zhao 2020). However, the historical roots of OECD challenge the organisation and the PISA test as political impartial (Ørskov 2019; Verger, Fontdevila, and Parcerisa 2019; Krejsler 2019).

1.3.1 A culturally neutral non-curricular achievement test

The PISA test claims to measure 15-years-olds’ ability to use their reading, mathematics, science knowledge and skills to tackle ‘real-life situations’. Thus, PISA is a so-called ‘non-curricular’ test, which avoids using items that may be too close to any country’s national curriculum and does not measure ‘school knowledge’ but instead ‘literacy’ in various knowledge domains. Thus, it avoids favouring students from particular countries and aims to be culturally neutral.

The question of test ‘neutrality’ and whether students from some countries are either disadvantaged or favoured in their test performance on PISA has been addressed by several researchers (Sjøberg 2019; Gjone 2010; Kjærnsli et al. 2004; Nortvedt et al. 2016; Jensen, Mork, and Kjærnsli 2018). The Norwegian researcher Svein Sjøberg has claimed that producing a culturally impartial test is an impossibility, and moreover, that this aim is impossible to reconcile with current ideals in, for example, science teaching, which emphasises localised curricula that situate science in a context. The ambition of test neutrality across cultures results in a test that is decontextualised: ‘the context is contrived or historical, based on statistical necessity or concern for “fairness”’ (Sjøberg 2019, 37, 47).

Sjøberg substantiates his allegations of decontextualised items by referring to procedures for item selection, which have been outlined in OECD publications. The procedures entail that items too close to national curricula measuring ‘school knowledge’ are excluded in the process.¹⁰ Despite these descriptions of the item selection process, several Norwegian studies have investigated the correspondence between the ‘PISA curriculum’ and the Norwegian

¹⁰ The processual conception of mathematics and the notion of competences were developed by mathematics teachers in the US, Denmark, and the Netherlands during the 1970s and 1980s (Gjone 2010).

National Curriculum, the ‘PISA curriculum’ being the ‘analytical framework’¹¹ for each domain worked out by appointed international experts. Norwegian studies have disagreed on the extent to which the national curriculum differs from the ‘PISA curriculum’, but some claim that the Norwegian mathematics¹² and science¹³ curricula are approximately the same.

A study made by Gunnar Gjone (2010) addressed whether Norwegian students were prepared for the type of mathematics encountered in the PISA test. He compared the Norwegian mathematics curriculum plan from 1997 (L97) with the analytical framework for the PISA test, which is based on a processual conception of mathematics and defines specific competences. Gjone concluded that in the Norwegian curriculum, notions like ‘competences’ were poorly developed and lacked precision.

He also compared concrete items from the final national exam in the Norwegian compulsory school with apparently similar PISA items and found that the Norwegian test items were not constructed to measure processual competence. Although at a superficial level items from PISA and from the Norwegian National Exam seemed quite similar, upon closer examination the measurement of a processual understanding of mathematics was lacking in the Norwegian examples.¹⁴ This difference in measured competences was not mentioned, and perhaps not noticed, by the Norwegian researchers who wrote the Norwegian report on PISA 2003 (Kjærnsli et al. 2004). They claimed that ‘[a]lthough the four central ideas in PISA and the four central ideas in the Norwegian L97 are not identical, their common content warrants the conclusion that they cover approximately the same content’ (Kjærnsli et al. 2004, 47). Thus, during the first decade after the first PISA test, Norwegian researchers differed on whether students in Norway were taught the types of mathematical competences emphasised by PISA, which stress problem solving and generalisations more than reproduction, definitions, and calculations.

¹¹ PISA’s analytical framework contains definitions of ‘literacies’ in the various domains measured by the test, i.e., reading literacy, mathematical literacy, and scientific literacy. The analytical framework also specifies competences within each domain.

¹² (Nortvedt et al. 2016).

¹³ (Jensen, Mork, and Kjærnsli 2018).

¹⁴ Gjone (2010, 174) also pointed out that the differences between the Norwegian National curriculum and the PISA ‘curriculum’ were in fact reflected in the Norwegian results on the PISA test, because Norwegian students scored lower on ‘problem solving and generalisations’ than on ‘reproduction, definitions, and calculations’.

Later studies, among them one done by Nortvedt et al. (2016), raised a similar question about the PISA mathematics test. These authors discussed the relevance of PISA 2012 to mathematics education in Norway and Sweden, comparing the PISA 2012 mathematics framework with the Norwegian and Swedish National mathematics curricula. Their focus was on learning goals for students aged 13-15. These scholars conclude that the mathematical knowledge, abilities, and skills described in the national curricula were relevant for solving PISA items, and consequently did not disadvantage Norwegian students in obtaining high scores on the test.

The PISA test in the domain of science has also been investigated. Jensen, Mork, and Kjærnsli (2018) compared the analytical framework for this domain in PISA 2015 with the Norwegian science curriculum, focusing in particular on the competence goals but also on the other three aspects of scientific literacy (knowledge, attitudes, and contexts), i.e., they compared competences, knowledge, and attitudes in the PISA framework with the competence goals and descriptions of basic skills in the Norwegian curriculum. Their comparison of science competences in the framework of PISA 2015 and the Norwegian science curriculum indicated agreement between the two.

Thus, among investigations that have compared the PISA test with the Norwegian National Curriculum, the majority of studies have concluded that there are no significant differences between the ‘standards’ measured by the test and the standards defined by the National Curriculum in Norway. In this sense, the PISA test is compatible with the Norwegian curriculum and is considered as a fair test for Norwegian students.

1.3.2 A policy instrument (with flaws)

In addition to assessing 15-year-olds’ ability to apply knowledge to ‘real life situations’ and assessing the school environment from the perspective of students and school leaders, OECD also recommends using PISA results to assess the quality of the educational system in participating countries. Andreas Schleicher, the director of the PISA programme in the OECD, claims that the test results may serve as valuable input for countries’ educational policy:

PISA is not only the world’s most comprehensive and reliable indicator of students’ capabilities, it is also a powerful tool that countries and economies can use to fine-tune their education policies. (Schleicher 2018).

The PISA test has been acclaimed and criticised. Some of the critique imbricates. Hence, in the international sphere, the Chinese-American Yong Zhao has summarised a(n) (established) critique of PISA. He claims that PISA is not measuring skills for modern economics and the future, that it holds a monolithic view of education, and is distorting the purpose of education (Zhao 2020). That PISA is measuring skills for modern economics and the future is not scientifically proven and is a made-up claim. Additionally, there is not ‘research showing that PISA covers enough to be representative of the school subjects involved or the general school knowledge-base’ (Hopmann 2008, 438, referred in Zhao 2020). Today, this claim can be challenged through the previous sub-chapter discussed. The monolithic view of education, meaning the PISA applicability of PISA is universal and not taking consideration that there are different societies and cultures with education that is imposed by it. PISA is constructed from a Western world-view that influence other countries way of living and leaving little room for local variations and needs. This challenge educational democracy. That PISA is distorting the purpose of education is tied to the fact that education is already established from the OECD through PISA. Hence, PISA’s measuring might come at cost with different responsibilities of education (Zhao 2020).

The staunchest critic of PISA in Norway is Svein Sjøberg, he emphasises that ‘OECD’s PISA project is not an educational project. It is a political project [...], it tells what young people should learn, regardless of the nation’s culture, nature, traditions and values.’ (Sjøberg 2017, 17). Moreover, Sjøberg is critical of the biased image portrayed of Norwegian education in the media debate after the announcement of the first PISA-results in 2001. As Sjøberg points out, Norway scored above France, the US, Denmark, Switzerland, and Germany. Still, the average score for Norway was described by conservative politicians and the media as scandalously low.

Sjøberg claims that ‘PISA has, in fundamental ways, changed Norwegian education and how we talk about it’ (Sjøberg 2014c, 196). In an edited volume on *Reform ideas in Norwegian education (Reformideer i norsk skole)*, he calls ‘the PISA-fication’ of Norwegian education ‘a success story from OECD’. Sjøberg’s assertion is that PISA has not only changed education in Norway, but has contributed to globalising the education sector, based on the ‘belief that markets are always good and that competition promotes quality among students, schools, teachers, and not least, among nations.’ (Sjøberg 2014c, 196). Underlining the broad impact of PISA, Sjøberg claims that this testing program has influenced our educational ideas and

educational practices as well as resulted in structural educational reforms based on neo-liberal and New Public Management ideas (Sjøberg 2014c, 197).

Sjøberg points out that since the beginning, leading Norwegian politicians, both the leader of the Conservative Party, Erna Solberg, and the leader of the Labour Party, Jens Stoltenberg, have legitimised PISA results as valid evidence of the quality of Norwegian education. During the campaign for the 2009 Parliamentary election, Solberg publicly pledged that with the conservatives in government she could promise a rise in Norwegian scores on international tests. In a similar vein, Prime Minister Stoltenberg in his 2008 New Year television address to the Norwegian people mentioned ‘the grave concern’ that the recently received 2006 PISA results had caused him. Norway had scored below average on the test. He assured the public that his government had ‘got the message’ (Sjøberg 2014c, 198). In 2013, Stoltenberg, after some improvement in Norwegian PISA scores in PISA 2009, proudly declared that the Labour Party, not the Conservatives, had initiated the measurement of quality in Norwegian education.¹⁵

Also, government policy documents from the early 2000s show that PISA, TIMSS and ‘OECD’s experts’ were frequently referenced in support of policy suggestions, although the emphasis on competitiveness and ‘being best in test’ are contrary to the values emphasised in the purpose clause of the Norwegian Education Act, which highlights values such as equality, solidarity, democracy, equal rights, freedom of expression, and neighbourly love (Sjøberg 2014c, 197). Sjøberg considers the emphasis among leading Norwegian politicians and successive Norwegian governments – of various political persuasions – on Norway’s PISA score as a confirmation of OECD’s ‘take-over’ of Norwegian educational policy making (Sjøberg 2014c, 208). In a recent publication he points out that 21st century skills, which entail ‘creativity, communication, collaboration, critical and scientific thinking, problem solving, social and language skills and interdisciplinarity’ have been embraced by PISA, and are central competences in the new Norwegian National Curriculum LK20 (Sjøberg 2022, 150).

One example of how PISA ideas have influenced our way of thinking about education is the introduction of the notion ‘learning pressure’ (*læringstrykk*), which is used in policy documents, newspaper reporting, and in educational research, where it has even been

¹⁵ This turned out not to be true, according to newspaper reporting. See Sjøberg (2014c, 198).

operationalised to enable the quantitative measurement of this ‘pressure’. This notion is inextricably connected with competition, and an emphasis on being the best, which in the PISA context is tied to human capital considerations, seeing educational spending as investments for economic growth (Sjøberg 2014c, 198–199).

Among the structural educational reforms resulting from PISA, Sjøberg mentions the Knowledge Promotion Reform (K06), and the introduction of a national quality assessment system in education (NQAS), which included National Tests. Before leaving office in 2005, the coalition government with the Conservative education minister Kristin Clemet, who had presented the first PISA results as ‘medium good’,¹⁶ but also expressed that Norway was a ‘school loser’ compared to the first PISA results in 2001, laid the foundations for these reforms. The red-green government which held office from 2006 to 2014, followed up the plans prepared by Clemet. Notably, the policy papers from the red-green government continued Clemet’s practice of using international assessment studies as their most important source of information about Norwegian schools. Sjøberg concludes his account of central politicians ‘across the aisle’ using PISA data and policy advice in justifying educational reforms by asserting that ‘[i]t is obvious that PISA and the OECD delivers the central premises for Norwegian educational policy’ (Sjøberg 2014c, 207–208).

However, the credibility of PISA as a valid instrument for assessing the quality of educational systems has been challenged by international researchers. In May 2014, an international group of more than 80 academics from more than 5 countries published an open letter to the Director of PISA in OECD, Andreas Schleicher, voicing their concern about the damaging effects of PISA rankings on education worldwide (The Guardian 2014). The letter points out that ‘[a]s a result of PISA, countries are overhauling their education systems in the hopes of improving their rankings, [by introducing] far-reaching reforms in accordance with PISA precepts.’ The public letter also claims that ‘PISA has contributed to an escalation [of] testing and a dramatically increased reliance on quantitative measures’, relying on test results that are ‘widely known to be imperfect.’

Regarding PISA’s influence on educational policy, the open letter points to a ‘shift of attention to short-term fixes designed to help a country quickly climb on the rankings’. The

¹⁶ The term ‘medium good’ was carefully selected by the Minister of Education and her associates (Bergesen 2006, 47).

letter also draws attention to PISA ignoring ‘the less measurable or immeasurable educational objectives like physical, moral, civic and artistic development, thereby dangerously narrowing our collective imagination regarding what education is and ought to be about’. The letter sees PISA’s emphasis on a narrow range of school subjects as an expression of OECD’s mission, which is to promote economic development. However, as the letter underlines, the main goal of public education is not only to prepare for gainful employment, but to ‘prepare students for participation in democratic self-government, moral action and a life of personal development, growth and wellbeing’ (The Guardian 2014).

This letter from over 80 academics threw doubt on OECD’s legitimacy as the educational ‘policy maker’ on a global scale, asserting that OECD, as opposed to United Nations organisations such as UNESCO and UNICEF, has no formal mandate to ‘improve education and the lives of children around the world’. Moreover, OECD also lacks ‘mechanisms of effective democratic participation in its education decision-making process’. OECD ‘has become the global arbiter of the means and ends of education around the world’; it has ‘led many governments into an international competition for higher test scores,’ and has ‘assumed the power to shape education policy around the world, with no debate about the necessity or limitations of OECD’s goals.’ (The Guardian 2014).

Schleicher retorted, in an exchange with Heintz-Dieter Meyer and Katie Zahedi, two of the signatories of the open letter, that ‘OECDs’ mandate is provided by the member countries of the OECD. Objecting to the allegation that PISA lacks democratic mechanisms, he clarifies that decisions on PISA and all other OECD activities are made by member countries, and decisions concerning PISA are made by the PISA Governing Board, with representation from all member countries (Meyer and Zahedi 2014).

The question regarding how OECD and PISA can exert their broad global influence without having formal legislative authority in any country has been addressed by various studies of how ‘soft governance’ works. Sjøberg, in his 2014-article on the ‘PISA-fication’ of Norwegian education, suggests that, along with other themes, ‘[h]ow does PISA influence educational policy?’ should be a topic for research investigation (Sjøberg 2014c, 213). He insists that among all the issues and critiques that have been raised about PISA, the far most important is ‘how PISA functions as an instrument for exerting political power’ (Sjøberg 2014c, 223). PISA seems to have achieved the power to outplay national educational goals, priorities, and curricula.

PISA has also been seen as an instrument for the Global Educational Reform Movement (GERM), which advocates increased competition and free choice, the standardisation of curricula, more use of tests and formal grading, test-based accountability, the ranking of schools and teachers based on test results, and achievement-based pay for teachers (Sahlberg 2011). Sjøberg suggests that the Norwegian PISA debate should change its focus from test results and ranking lists and instead examine more fundamental issues. His final conclusion is that the GERM ideas represented by PISA are a threat to the basic values of Norwegian education (Sjøberg 2014c, 223).

Sjøberg, who claims that PISA has heavily influenced educational policy in Norway, has also provided some explanations for this. One suggested explanation is that OECD is generally held in high regard among Norwegian politicians, and the organisation's policy advice is highly valued. As evidence, Sjøberg refers to the address given by Prime Minister Stoltenberg at the 50th Anniversary of OECD, delivered in the presence of the Norwegian Finance Minister and Foreign Minister as well as the Norwegian press corps. On this occasion, Stoltenberg emphasised that OECD has been a particularly important organisation for small countries like Norway. 'Through 50 years it has been one of the world's most important provider of knowledge and foundational principles. Very many countries listen to the main messages from OECD, and in that way the organisation has contributed to changes in the world.' (Sjøberg 2014c, 222).¹⁷ The credibility and relevance attributed to the PISA test by leading Norwegian politicians might be due to the general esteem enjoyed by the OECD.

Sjøberg also explains the acceptance of PISA as well as other International Large-scale Assessments (ILSA) at the national level by their contribution to creating job markets for academics with statistical and educational expertise. In Norway, it is the ILS (Institute for Teacher Education and School Research) at the University of Oslo that manages international and national tests through tenders, not research applications. Historically, forty percent of this department's funding came from such test-enterprises (Sjøberg 2022, 135).

Hopmann (2007, 15) has written that PISA has a large 'market share' to uphold. Public money is being put into PISA and similar tests and many educationalists have these tests and

¹⁷ There are also examples of the opposite. For instance, the OECD-report from 1988/89 recommended the use of national tests. However, the Education Minister from the Labour Party at that time, Gudmund Hernes, did not listen to this advice. It took many years before national tests were implemented (Elstad 2010, 60).

topics as work, an industry that is expanding with collaborations in the research field. According to Hopmann, this is a ‘too big and too seductive’ business to turn down just because of critics who do not support the whole ‘enterprise’ or parts, such as weaknesses in the methodology. Extensive critique of the methodology of PISA has been forwarded on test constructs, design, sampling, IRT (Item-Response-Theory), data processing and questionnaires, and even curriculum and culture curriculum fairness (Hopfenbeck et al. 2018; Zhao 2020). For instance, sampling problems are connected to school start age which varies between countries, meaning that some have more school exposure than others (Zhao 2020). Furthermore, the PISA tasks are not the same for all students, and content varies (Sjøberg 2014b, 36; 2022, 144). Another element with the tasks is that verbs are translated differently between countries that have similar languages (Sjøberg 2022, 144). The PISA tasks have also been critiqued, partly because some of the tasks are secret due to reuse and therefore not published.

Clockwise, the PISA booklet consists of ‘multiple-matrix-sampling’, which entails that only 20 percent of the PISA tasks are included. Based on this design, one can decide the score for the whole population. Sjøberg (2022, 145) mentions that the way from students answers to the PISA score is complicated, and even statisticians find it hard to assess details around the PISA score for a population. The data analysis consists of Item Response Theory and Rasch-modelling. Professor Svend Kreiner, a Danish veteran Rasch-statistician, says that his country can be ranked as 2–42 on the ranking list with this method. Kreiner’s method criticism has not been challenged. Another element is that the technical reports are published a year after the PISA rankings (Rutkowski and Rutkowski, 2016, referred in Sjøberg (2022, 146)).

Rutkowski and Rutkowski (2016) also address weaknesses in the participant sampling (sampling error), i.e., exclusion rates beyond the maximum 5 % for some countries. This exclusion entails students with disabilities and migrant backgrounds (Zhao 2020), and rural backgrounds (Thomas 2021). Moreover, Rutkowski and Rutkowski (2016) writes that the percentage of students included in the samples between countries differ, there are issues with the achievement estimation model (measurement error), i.e., item parameters are not equal across measured populations (measurement equivalence – different booklets¹⁸ – different clusters and items and degrees of difficulty). There are also missing and error-prone

¹⁸ (Rutkowski, Rutkowski, and Zhou 2016, 4).

background data as well as issues with the measuring trends (linking error), i.e., poor content overlap (few common items measured between cycles where the domain is minor). Even the UK education magazine TES asked ‘Is PISA Fundamentally Flawed’ in 2013 (Zhao 2020).

1.3.3 Entangled with global edu-business (the global education industry – GEI)

The open letter on PISA addressed to Andreas Schleicher, and signed by over 80 academics, claimed that OECD has ‘entered into alliances with multi-national for-profit companies, which stand to gain financially from any deficits – real or perceived – unearthed by PISA’ (Meyer and Zahedi 2014, 872). American schools and school districts are among the customers buying these services on a ‘massive’ scale, but the engagement of for-profit enterprises in a global education market where educational products and services are bought and sold is expanding to an extent which warrants the name the global education industry (GEI).

Stephen J. Ball is one among several researchers who have investigated the proliferation and growth of multinational private enterprises selling products and services to the education sector (Ball 2012). The entire World Yearbook of Education for 2016 was dedicated to the rapidly growing global education industry (GEI).¹⁹ Ball documents that Pearson Education is the world’s largest education company selling services and products on a global scale, even recipes for ‘one-off’ educational reforms (Ball 2012, 134).

In his analysis of big business in education, Ball draws attention to Pearson in particular – who won the bid for important parts of PISA 2015, and for developing the Framework for PISA 2018 – and who operates on a global scale in selling products and services in the area of pedagogy, curriculum, assessment and ‘joining these up’. The company’s curriculum and assessment work ‘contributes to define what cultural knowledge is most worthwhile’ (Ball 2012, 127). Being situated in a new ‘policy space somewhere between multilateral agencies, national governments, NGOs, think tanks and advocacy groups, consultants, social entrepreneurs and international business, in and beyond the traditional sites [...] of policy-making’ (Ball 2012, 10), big international commercial enterprises are increasingly positioned to decide what knowledge is of value and how it is going to be tested and presented. Commercial enterprises become educational policy players with their own agendas, which is

¹⁹ Lubienski, Christopher, Gita Steiner-Khamsi, and Antoni Verger. 2016. World yearbook of education 2016 : the global education industry. Abingdon, Oxon: Routledge.

primarily to expand the market for their own educational products and services. In an article from 2016 on the Global Education Industry (GEI), Gita Steiner-Khamsi points out that ‘the adoption of the same [de-nationalised] set of competency-based standards across countries, benefits GEI, and she mentions PISA as an example (Steiner-Khamsi 2016).

OECD’s expansion of the PISA test to include spin-off products such as PISA for individual schools (PISA for schools), PISA for adults (PIAAC) (PISA for 16–64 year-olds) and PISA for developing countries (PISA-D — PISA for development), has raised concerns about whose interests are served by this expansion – business interests in growing new markets for selling profitable educational products and services, or the wish to provide education for deprived groups (Sjøberg 2022, 147–150). However, spin-off products from PISA entails more than studies on solely the three PISA domains: mathematics, science and reading.

1.4 Norwegian PISA researchers assessing ‘PISA effects’ and future prospects for international large-scale tests

Marking the 20th anniversary of PISA research in Norway, Björnsson and Olsen (2018a), who are researchers at the Department of Teacher Education and School Research (ILS) and the Centre for Educational Measurement at the University of Oslo (CEMO),²⁰ edited a collection of articles that summed up findings and learning points from working with International Large-scale Assessments (ILSAs) for 20 years.²¹ In this overview article they briefly assess the general question concerning the causal effects of PISA and other International Large-scale Assessments (ILSAs) on educational change in Norway during this 20 year period. They also discuss the strong and weak point of such tests, as well as the dilemmas confronted in the test construction, examining the future prospects of international large-scale assessments, such as TIMSS and PISA.

Regarding the quality of international tests, Björnsson and Olsen point out that in recent years, after serious critique, it has been increasingly recognised that the validity of the test used for international comparisons may be lower than previously assumed (Björnsson and

²⁰ CEMO acts as an advisory body to the Ministry of Education, the Directorate of Education and international units (UiO n.d).

²¹ Norwegian researchers have been working on analysing PISA (and TIMSS) data for more than 20 years (Lie et al. 2001; Kjærnsli et al. 2004; Kjærnsli et al. 2007; Kjærnsli and Roe 2010; Kjærnsli and Olsen 2013; Kjærnsli and Jensen 2016; Björnsson and Olsen 2018b; Frønes and Jensen 2020). In this thesis a summary of one of the latest publications is highlighted, which is a summary of 20 years with PISA and TIMSS.

Olsen 2018a, 17). This does not imply that international comparisons are meaningless but that test items related to national contexts should be included in future tests.

So far, such tests have been largely driven by consensus processes. There are several stakeholders or partners—which have increased over the years—that must agree on the terms. The aim for the tests is to facilitate international comparability, which makes it necessary to ensure that the tests are standardised and conducted in the same way in every country (Björnsson and Olsen 2018a, 29). A small country like Norway will unfortunately have little influence on the structure of international tests. A possible solution for future test versions could be a modularisation of the tests, creating modules with items that are suitable for specific regions of the world, in combination with general modules. This would enable both the diachronic comparison of results, showing how average country scores vary or stay the same over time, as well as synchronic comparisons between different countries. Another solution would be to connect data from future international tests with data from national tests, which requires personal identity information of individual participants. This would make it possible to follow students both backwards and forwards in time and would provide much better information for national policy. Such linking of data sources have already been put in place in some countries (Björnsson and Olsen 2018a, 29–30).

On the future prospects for largescale internal assessments, Björnsson and Olsen (2018a, 26) consider it is very unlikely that international tests, in the coming decades, will lose their importance or even vanish. One reason is that many countries have incorporated such tests in their national quality monitoring system. In Norway both TIMSS and PISA are part of the National Quality Assessment System (NQAS/NKVS), and it is compulsory for Norwegian students to participate in the test when they have been randomly selected to be in the test sample. Indicators of quality on international large-scale tests will not become less significant over the next 20 years, but they may undergo changes. Adaptations are bound to occur due to developments in technology, methods, and design, and also because the school content is constantly changing (e.g., changed curricula), and there are continuous discussions on what kinds of qualifications and competences are needed in future society.

These Norwegian PISA researchers reject the idea of finding causal links between Norway's participation in international large-scale tests, among them PISA, and changes in Norwegian education, claiming that:

it is obviously not possible to establish any clear causal link between Norwegian participation in international assessments and the changes that have taken place [in Norwegian education] during the same period [the 20 years since PISA was first introduced in Norway]. (Björnsson and Olsen 2018a, 20).

In support of this rejection, they mention a recently published study by Nortvedt (2018), which describes how national educational policy is shaped in a complex cultural context, where recommendations from international organisations are not necessarily followed up. Björnsson and Olsen underline this point by referring to other studies that resonate with Nortvedt's finding (Takayama 2008), confirming that results and analyses from international largescale studies are used to affirm or even strengthen existing policy orientations in national contexts.

To these researchers the methodology for data collection in international largescale studies also leads to the rejection of clear causal links based on PISA data. Björnsson and Olsen assert that ILSAs are 'cross sectional studies that hardly can lead to firm conclusions about causes, and the studies only include some selected items for measuring education systems. But the potential for cautious causal interpretations also exists for this type of study.' (Björnsson and Olsen 2018a, 17). Cross sectional studies, where all data are collected at the same point in time, are often contrasted with longitudinal studies that follow individuals or groups over time. The explanation for why clear causal links cannot be established by analysing PISA data is apparently the premise of events. Cause and effect are understood as a sequence of events, where the cause occurs before the event – since data that are collected at the same point in time cannot be used to draw such conclusions.

However, not only does the procedure for data collection, i.e., collecting all the data at the same point in time, prohibit the drawing of causal inferences from large-scale international studies, the statistical models typically used in the analysis of such data also prohibit causal inferences. This is because the statistical models measure the association between variables and not their causal connections. An article written by other PISA researchers in the same volume, Nilsen and Blömke (2018), confirms the view that, since the data are cross-sectional, i.e., all data are collected at the same time, statistical methods measure associations rather than causality, and drawing causal inferences is prohibited:

all large-scale international surveys such as TIMSS and PISA are cross-sectional and collect all their data at one point in time. Therefore, it is not possible to draw causal conclusions about whether a factor A led to B. For instance, we cannot investigate whether teacher quality leads to higher learning outcomes. We can only investigate whether teacher quality is associated with learning outcome. But to facilitate understanding and language fluency and to emphasise the direction in which we have run the regression, we still use words as influence and effect even though this is strictly causal language. (Nilsen and Blömeke 2018, 65).

Whether PISA has had any effects on Norwegian education is therefore rejected in the quantitative realm. It is just not any clear causal links. Causal language is just used for directional guidance of the regression analysis. However, PISA effects remain a contested issue, as the following subchapter address.

1.5 A contribution to the (enlarged) field of PISA effects

The French researcher Xavier Pons recommends establishing systematic and cumulative research with key factors and variables within the field of PISA effects (Pons 2017). Supportively, there are numerous of scholars using the notion ‘effect’ in the PISA research when addressing the influence of PISA. Examples are ‘effect(s) of PISA’, ‘PISA and its effects’ (Rautalin, Alasuutari, and Vento 2019), or ‘PISA effect’ (Grek 2009). PISA has in research reviews also been categorised as ‘PISA effects’, ‘effects of PISA’, ‘influence of PISA’ (Pons 2017), and ‘PISA’s impact’ (Hopfenbeck et al. 2018). This entails that the research field is divided between different terminologies despite addressing approximately the same: PISA effects. Hence, PISA effects is an incorporated term that refers to a large research literature. However, the different denotions of PISA effects can be misleading, it can easily result in a belief that PISA alone was the cause of an event. Nevertheless, the denotions can be retained, but with the caveat that the effects are mediated by humans/social agents and that a number of reasons other than PISA results may have contributed to the design of specific educational reforms or the use or rejection of PISA in school. Based on these observations, argumentations and Pons’s request of systematic and cumulative research within this research area, this research will give a contribute to the PISA effects field.

The question is what specific contribution it will bring to the field (see subchapter 1.6). Pons has already stated that the PISA effects field is ‘fuelled by many individual contributions from various disciplines and academic traditions’ (Pons 2017, 133). A claim that is based on

explorations of a selection of PISA effect studies on education governance and policy processes. Thus, Pons has already pinpointed absences which I will embark on in this thesis, that still persists in the PISA effects field. However, I will give a brief introduction to the PISA effects field with various of peer-reviewed PISA studies, which is wider scoped (e.g., quantitative, micro-level, and Norwegian studies are for example included) than Pons's literature review that was limited to education governance and policy processes. An argument for enlarging the PISA effects field, it is not because Pons's pragmatic systematic reduction of it, it is grounded in how I have defined an effect in the introduction of this thesis which gives a wider inclusion of studies.

My enlarged PISA effects field gives an overview of PISA studies and adds studies for grounding this research. The overview also indicates what is absent, even within Pons's literature review. For instance, despite mentioning that context matters, Pons (2017) literature review is seemingly built on studies from the macro-level, leaving out the entanglement with PISA on the micro-level. Pons mentioning of key factors might therefore be understood as corporate agents, not primary agency. The difference is their bargaining power. However, this is a construe not a rejection of factors at the micro-level. Therefore, continuing with the same screening procedure as Pons, from where he left, could exclude vital contributions of PISA studies concerning the micro-level.

My own observation is that the PISA effect field is heterogeneous with different units, focuses and conceptualisations. For example in the quantitative realm there are studies based on secondary data analysis of national PISA student data measuring various correlations and/or trends (see e.g., Hopfenbeck et al. 2018; Björnsson and Olsen 2018b; Zheng, Cheung, and Sit 2022). Most quantitative studies are occupied with associations that can explain differences in PISA scores. Some have sought the 'generative mechanisms' for high(-skilled) occupational expectations using PISA data using regression analysis and covariate modelling (Jiang, Chen, and Fang 2021). In the quantitative realm, there is also studies on PISA Governing Board members (e.g., Breakspear 2012), school leaders and teachers (e.g., Utdanningsforbundet 2008), and students (e.g., Hopfenbeck and Kjærnsli 2016) with PISA as a topic using survey as a method. There are also mixed methods studies with PISA (data) (see e.g., Hopfenbeck and Kjærnsli 2016).

In the qualitative realm there are studies of effects on national educational policy (includes educational systems) with document analysis (see e.g., Baird et al. 2016; Pons 2017; Morgan

2018; Hopfenbeck et al. 2018; Yang and Fan 2019; Thomas 2021). There are studies of PISA team members (e.g., Aursand 2018), policymakers (e.g., Adamson et al. 2017), policy officials (e.g., Hossain 2023), school leaders and/or teachers (e.g., Eggen 2010; Bringeland 2015; Radišić and Baucal 2018; Aursand 2018; Aursand and Rutkowski 2021; Dilekçi 2022; Andersson and Sandgren Massih 2023), students (e.g., Hopfenbeck 2010) and teachers' parents' representatives (Hossain 2023), with interview as a method. There are also ethnographic observational studies of when Norwegian students are taking the PISA test (Hopfenbeck 2010).

In the qualitative realm one will also find that PISA products has had effects on 'the emergence of topological mechanisms' (i.e., Lewis, Sellar, and Lingard 2016). Moreover, like the quantitative realm, PISA has effects on finding the 'generative mechanisms' limited to different capabilities and capital of the student affecting equity in education and educational performance in PISA (Pham 2019). There are PISA critique studies holding qualitative and quantitative arguments (see e.g., Hopfenbeck et al. 2018; Zhao 2020). There are also studies comparing the PISA framework with the national curriculum (e.g., Gjone 2010; Nortvedt et al. 2016; Jensen, Mork, and Kjærnsli 2018).

PISA studies has also used concepts from relatively known philosophers such as Foucault (e.g., Kanes, Morgan, and Tsatsaroni 2014), Fairclough (e.g., Thomas 2021), Bernstein (e.g., Kanes, Morgan, and Tsatsaroni 2014; Kelly and Kotthoff 2017), Bourdieu (e.g., Stray and Wood 2020; Andrews 2021), and Luhmann (Santos, Carvalho, and Portugal e Melo 2022) and Latour (e.g., Gorur 2011; Serder and Ideland 2016), for explaining phenomena related to PISA. There is also research on PISA spin-off products such as PISA for Schools (e.g., Lewis, Sellar, and Lingard 2016; Lewis and Lingard 2023) and PISA for Development (e.g., Addey 2020).

PISA studies, besides the use of PISA data and the PISA framework, also involves the use of PISA tasks (Giberti and Maffia 2020). PISA tasks are used for problem-solving teaching sequences with primary students (O'Shea and Leavy 2013), or designing PISA-like tasks for 15-years-olds (Nusantara, Zulkardi, and Putri 2021), and 7th graders (Aini et al. 2023). Additionally, for testing school mathematical knowledge (contextual, conceptual and procedural knowledge) amongst pre-service teachers to solve PISA tasks (Sáenz 2009).

There are many factors and variables activated in the PISA effects field. The field consist of relative few studies from ‘the ground’, and their ‘opinions regarding this [ILSA] remain silent or unheard’ (Hossain 2023, 5). This could for example be how PISA results are being used in educational settings. However, Hossain limits his respondents to parents’ and teachers’ representatives from their collective associations since ‘neither parents and teachers nor education officials in the government are accountable to participate in ILSAs, they would not be likely to be concerned about the results of these assessments’ (Hossain 2023, 5). Other scholars seeks ‘the role of policy officials in mediating global-local policy interactions’ (Stray and Wood 2020, 267). The question is if school personnel in an abstract way also can be viewed as ‘policy officials’ and become integrated under this umbrella for responding to global-local policy interactions’? Altogether, agency is needed with PISA since they are either policy mediators (Stray and Wood 2020), or policy implementers (Hossain 2023). Notably, PISA studies are undertheorised from the micro-level (Hossain 2023). In the Norwegian context there was a quest with PISA from almost a decade ago about focusing on schools (Sjøberg 2014a, c) and teachers work (Sjøberg 2014c).

Besides focus on agency, deep research (Pons: overarching theoretisations) is quested related to education governance and policy (Pons 2017; Stray and Wood 2020). Previous publications have not focused on the interplay between deep structure-agency mechanisms and their interplay. Some might argue that Fairclough critical discourse analysis would cover it, but it won’t alone embrace such deep mechanisms due to focus on discourses. Neither would Bourdieu, Bernstein, Luhmann, Foucault and Latour because of their focuses—they don’t have focus on emergent properties and their interplay. Actually, neither would meta-theories such as naive (empirical) realism (sense-observation), social constructivism (social constructs) and interpretivism (construes) alone because they do not focus on real ontological structures and their mechanisms based on emergence and counterfactual thinking of a necessary ‘laminated system’ (Bhaskar and Danermark 2006).

However, some scholars are using terminology that is familiar with critical realism (e.g., Lewis, Sellar, and Lingard 2016; Pham 2019; Jiang, Chen, and Fang 2021). These works focus on causal powers with agency and/or some causal powers with the school which they identify as the ‘generative mechanisms’ for explaining their problem (i.e., Pham 2019; Jiang, Chen, and Fang 2021). Some are focusing on ‘emergence of topological mechanisms’ (i.e., Lewis, Sellar, and Lingard 2016). However, critical realism understands reality as stratified, based on emergence, this should include deeper generative mechanisms that explains the

interplay of change between social objects i.e., between structure and agency (Bhaskar [1975] 2008).

But first, picking up Pons again; PISA effects studies are ‘rarely confronted in a dialogical and cumulative way’ (Pons 2017, 133). In my own screening of the field I did not come across any confrontations, apart from critique studies on PISA. Also, according to Pons; one should better conceptualise PISA effects (Pons 2017, 131). This implies causality. Pons states that it takes an extensive theoretical framework to conceptualise an effect (Pons 2017, 141). In the introduction part of this thesis I defined an effect based on a critical realist foundation. However to define an effect based on meta-theory is not enough to meet Pons’s request. This study needs to embrace a selection of previous studies with another way of understanding causality that gives another perspective of what PISA effects there has been and/or where deep PISA effects are situated. This is what will differentiate the former PISA studies (on generative mechanisms) from this one. Thus, a necessity will be the generative mechanisms that explains the interplay of change between social objects i.e., includes change at macro and micro level (structure-agency interplay). To fulfill Pons’s request for confrontation, there should be an overarching theoretical gap for recontextualisation (of one or more factors) in the PISA effects field.

Pons wrote a critical review of ‘Fifteen Years of Research on PISA Effects on Education Governance’ in 2017. The review was based on 87 references derived from searches in the data base ERIC (Education Resources Information Centre) (Pons 2017). Pons was dissatisfied with many PISA studies making the ‘PISA shock’ a central theme. He also lamented the apparent lack of knowledge, even within the OECD, about the already large body of PISA research which avoided focusing on the ‘PISA shock’. Instead of using the notion of ‘shock’, which Pons regards as a metaphor rather than a concept, he recommends the scientifically established concepts ‘reception, uses and effects of an international survey’ (Pons 2017, 133). A property with this model is that it can embrace socio-cultural interaction in a field e.g., documented interaction.

Apart from Pons critique of researchers’ using the metaphor ‘PISA shock’, Pons claim that ‘PISA introduced major changes in the governance of education worldwide’ and that ‘PISA has a strong influence on a variety of national reforms’ (Pons 2017, 131). He underlines that changes in the governance of education effected by PISA had been driven by ‘soft power strategies’ and ‘new policy transfers’ based on ‘data and measurement tools which redefine

the scales of education policies' (Pons 2017, 131). Moreover, he points out that PISA's strong influence on a variety of national reforms, which had been documented in numerous case studies, strongly depended on domestic policy contexts. The analyses of such policy contexts, however, employed different conceptual frameworks. Pons emphasises that missing from these studies were 'overarching theorisations of the political meaning of PISA effects on education governance and policy processes' (Pons 2017, 131), with key factors and variables (Pons 2017). What is construed requested from Pons, or if you prefer Stray and Wood (2020), is deep research.

Pons wants upcoming research to transcend PISA effects beyond the input-output discourse of the 'PISA shock' which entailed implementation or justification of new policies and devices on educational systems. Pons (2017, 132) states that the scientific consequences of the input-output discourse were studies on effects of the survey creating national standards and curriculum reforms. Further studying different reference societies and their school models, writing about the absence of shocks in other surveys, and comparing shocks between school systems. In addition, discussing different policy reactions to PISA results. Hence, Pons want to overcome such effects with the model 'reception, uses and effects'. There is no standard script on how to do this, so ingenuity and innovation is needed. What Pons's model can do in this thesis is to redirect (stated) PISA effects to deep effects (for discussions).

1.6 A gateway within the PISA effects field

Pons states that the PISA effects field is missing 'overarching theoretisations of the political meaning of PISA effects on education governance and policy processes' (Pons 2017, 131). This do not mean that there are not studies using conceptual frameworks, as exemplified earlier, but they 'rarely conceptualise the PISA effects themselves and do not always distinguish the key variables or factors that can explain why and how such effects occur or not' (Pons 2017, 138). Again, causality has been previously introduced. The need for overarching conceptualisations suggests the application of fundamental theories, but theories that feature education governance and policy processes with key factors and variables. Pons points to three main challenges for 'subsequent studies' of PISA effects: better conceptualisations of these effects, preservation of an epistemology of uncertainty, that is, avoiding 'taken for granted views', and normalisation of research on PISA effects, not to perpetuate 'its so-called novelty' of the input and output discourse (Pons 2017, 131). In this study I will undertake this gateway to the field and give Pons a reply. Altogether, I will follow Pons's propose for a better conceptualisation of PISA effects, i.e., capture key factors and

variables that can explain the ‘political meaning of PISA effects on education governance and policy processes’, provided by a new conceptual framework (includes theoretical frameworks and methodological approaches) to the PISA effects field.

I was reintroduced to Archer’s educational system theory that includes agency by my supervisor, and introduced to critical realism at a conference. Afterwards I found out that the field of PISA effects have not utilised Archer’s conceptual framework which is underpinned by critical realism. This was underscored by pragmatic searches in Oria (Norwegian database) and ERIC (Education Research Information Centre) combining the keywords ‘PISA’ and ‘Archer’. None of the search engines provided any results for this combination. Not even my master’s thesis was identified, which is included later in this thesis since it is relevant for this study and to mention that some of Archer’s concept are used. I also double-checked for PISA and reflexive modes and Archer’s two types of educational systems with PISA in the same search engines. There were no relevant matches. This indicates that there is no use of Archer’s theoretical concepts within the field of PISA effects besides my Master thesis.²² Neither did I find any educational system theories, except from Bourdieu and Luhmann, used in this field which tells us that the PISA effects field is under-theorised i.e., insufficiently theoretically studied, which Pons literature review underscores (Pons 2017). As will be recalled; Pons asks for systematic and cumulative research within this field. This entail research that is built on the same foundations. A way to start is therefore with Archer’s conceptual framework, which will enable theoretical, methodological, and empirical contributions to the PISA effects field.

Dealing with Archer’s theory and legitimising the use of it in the upcoming overarching research questions, except from the theory being absent in the PISA effects field, is further grounded in her argument of avoiding theoretical conflationism. In downwards conflation, causal power is denied to the ‘people’ while in upwards conflation causal power is denied to the structure or system. The model of ‘the rational human being’ represents upwards conflation, i.e., ‘the single property of rationality is held to *make* both human beings and also their society’ (Archer 2000, 5). *Society becomes an epiphenomenon of agency*. The social constructivist model is a downwards conflationary model in seeing human beings as totally

²² It is not a problem if Archer’s theory was identified if new research questions is posed or one starts systematic and cumulative research jointly because more data is needed in another context.

malleable ‘indeterminate material’ (cf. Émile Durkheim, and in works of Michel Foucault, Basil Bernstein, and Niklas Luhmann) who are totally shaped by the *pressures of socialisation* and where *agency becomes an epiphenomenon*. However, although upwards and downwards conflation may in some sense be strategies for theoretical reductionism (not only found in qualitative research but also in quantitative research with associated variables on a stratum). Archer prefers to use the term ‘conflation’, because the characteristic ‘conflationary theory’ also encompasses another type of social theory, which is not reductionist in the sense that ‘the people’ nor ‘the parts’ are left out per se. The problem is they insist on the inseparability of ‘the parts’ and ‘the people’. In these types of theories as well, emergence, autonomy and causal power are denied either to the parts or the people. Thus, Archer claims: ‘conflation is the more generic error and reductionism is merely a form of it’ (Archer 2000, 5–6). Since structure and agency are conceived as intertwined, their separate effects cannot be studied. Archer points out that this type of conflationism is encountered in Pierre Bourdieu’s and Anthony Giddens’ works. I argue that this type of conflation is also found in Bruno Latour’s work on actor-network theory due to a realism perspective. Archer claims that analyses relying on central conflationist theorising tend to vacillate between extreme voluntaristic accounts, i.e., the agent is the only causal force, and determinism, i.e., structure, is the only causal force (Archer 2000, 6). Recalling Pons epistemology of uncertainty i.e., avoiding ‘taken for granted views’ with PISA effects studies, is also to avoid conflationist theorising as such approaches prevent the study of which conditions give agents greater degrees of freedom and which conditions are experienced as constraining. In practice such approaches preclude autonomy and causal powers either to structure or agency, making it hard to explain change, but also to embrace different ways of operating at different levels of the educational system. Decoded, the different outcomes with conflationist approaches are either that the agent decide their over own working operations (upwards), or the structure decides the agential working operations (downwards), or lastly, both structure and agency are mutually constitutive of the working operations (central), that deciding the determinant of an operation could therefore entail vacillation between agency and structure as emergent properties (autonomy) from both levels are withheld. Hence, ‘the two elements cannot be untied and therefore their reciprocal influences cannot be teased out’ (Archer 2000, 6).

Archer recommends analytical dualism to preserve the autonomy and causal power of structure, culture, and agency. Archer developed her own approach to research, the morphogenetic approach, where one can study social emergent properties (SEP), cultural

emergent properties (CEP), and personal (agential) emergent properties (PEP) of change or reproduction (Archer 1995). Archer conceptual framework includes the morphogenetic approach: the morphogenetic cycles, real conceptualisations of the structural system, the cultural system and personal (agential) system, and the Internal Conversation Indicator (ICONI). Deconstructed, for example, Archer's real conceptualisations of the educational system (SEP) and reflexive agency (PEP) opens for doing specific comparisons within and between countries. Compared to other systemic theories on offer, Archer provides levels of theorising based on emergence, autonomy and causal powers (Cruickshank 2003, 143–145). Archer theorising involves a four-level theorising offering transcendental perspectives: meta-theory (critical realism), general theory (structure-agency-culture), domain theory (the educational system and agency) and specific theory (for example her own investigations of France, Russia, Denmark, and England).

An advantage of Archer's conceptual framework is that it is inclusive. Archer's conceptualisations contain entities both at the macro and the micro levels. Her conceptual framework addresses both structural change (macro-level) and how the structural and cultural context condition individual (school leaders and teachers) actions, practices, and reflections (micro-level). Archer's real conceptualisations and definitions considers particularities, similarities, and system changes i.e., changes in and of mechanisms caused by agency as a mediator of ideas and materials. One can possibly reclaim reality if there any alleged system claims, and if intended, make predications of tendencies.

1.7 Focus on two social forms²³ and a reconceptualisation of causality

Recalling Pons finding on missing overarching theoretisations in PISA effects studies, this study will focus on Archer's social forms of the educational system and agency, and their interplay. The missing structure-agency interplay was also pinpointed by my own findings in the PISA effects field, but referred as social objects (and their generative mechanisms). Altogether, these might be the fundamental key factors (with their variables) that Pons seeks for better understanding PISA effects on education governance and policy processes. Theorising on critical realism and Archer's entails understanding causality as a necessary connection instead of sequences events i.e., if certain events follow (in time) after another

²³ Agency as a social form can be discussed as humans are natural and social manufactured. For instance, reflexivity is a natural and trained social property, that is exercised throughout life, and that filters the social reality.

event, the first event is seen as causing the following event. The latter is the Humean way of understanding causality. According to critical realism and Archer, events that follow each other in chronological time are not necessarily causally connected (Bhaskar [1975] 2008; Archer 1995). Therefore, when advocacy of educational reforms follows (in time) after the announcement and debates about PISA results, they are not necessarily a PISA effect in the sense that PISA caused it, they might have come without PISA. Thus, PISA results may simply be used to argue for policies that have already been decided upon, and may be distorted in the policy making process.

According to critical realism and Archer's conceptions of causality, PISA and PISA results by themselves have not effects, they have only causal power that needs to be activated (Bhaskar [1975] 2008; Archer 1995). It is only when PISA ideas and PISA results are adopted by influential agents as valid knowledge about the quality of educational systems that PISA effects may occur on the macro-level. Such influential agents may encompass leading politicians, political parties, professional associations with influence on educational decision making. This do not exclude the fact that effects of PISA can be present amongst other stakeholders as they interact with PISA, either as a cultural or material phenomenon, and this can create effects in the educational system based on properties with PISA i.e., they can equally as stakeholders be convinced that PISA is a valid test for knowledge. Again, the sequence of events, from the dissemination of PISA results to the use in schools, might not be a causal association. In opposition to what caused educational reforms, PISA can cause effects on agency itself, despite one can argue that international testing and focus on results are seen more relevant causes for the use of PISA in school. However, to reduce effects of PISA would exclude the fact that PISA is a cultural phenomenon per se and neglect differences in 'reception, uses and effects' at different levels of the educational system.

Conversely, the mismatch with a necessary connection is that the same novelty of PISA effects can reoccur i.e., that PISA again is being reduced to an input-output discourse. To avoid this simplicity we can gather information on how PISA is being used, in the modus operandi of the agent, and thereby establish which effects there have been on agency with PISA. This would be compatible with Pons model. It could entail a clearer focus on teachers' practice when encountering PISA ideas, depending on their personal concerns and values. Again, to avoid effects on agency itself, is simply to avoid the necessary connection between the phenomenon and the agent and to understand possible change on different levels of the educational system. Thus, two types of fundamental effects are possible with PISA: on the

parts and on the people. That there are effects on the parts and the people can explain why certain change and ways of dealing with PISA happens on the macro and micro level of the educational system. A reconceptualisation of causality also entails a reconceptualisation of alleged PISA effects i.e., either claimed PISA effects and/or the understanding of where PISA effects are situated (f.ex., Sjøberg: on schools and teachers work). This also challenge the claim of no PISA effects, as rationality instead of statistical models, can give us another perspective. However, multi-causality is a keyword, meaning that some causal powers will be more dominating than others for a given effect. Asking counterfactual questions can challenge the already claimed PISA effects or even where they are situated.

1.8 Reconnecting the reconceptualisation on the two social forms in the Norwegian case

Pons model ‘reception, uses and effects of an international survey’ amongst actors can embrace reactions, social positions, ideology, values, and ontological mainstays on how PISA is used. In the PISA effects field there has been statements about PISA effects on the Norwegian educational system. There are claims of the system being more decentralized (Mausethagen 2013; Baird et al. 2016; Imsen, Blossing, and Moos 2017; Nortvedt 2018; Camphuijsen, Møller, and Skedsmo 2021) or (re)centralized (Imsen, Blossing, and Moos 2017; Nortvedt 2018) because of PISA reforms. Such claims can be grounded in agents ontological mainstays and their way of understanding causality. Since Archer’s conceptual framework is based on necessary connections, emergence and change in educational systems, one can discuss whether the Norwegian educational system has fundamentally changed after the introduction of the PISA test. This will fulfill Pons’s request for confrontation of PISA studies where the educational system is a key factor (with its variables) for explaining change.

A reason for studying PISA effects amongst the people is PISA’s resurrection each third year with modifications e.g., optional modules. PISA is a moving target which is not absolute. PISA can be dealt with quite differently. Teachers and school leaders are key factors of interest on the micro level as they are responsible for conducting the test. Eggen (2010) reported explicitly that school leaders need to have a plan and strategy with PISA. In schools there are ongoing interests with PISA for formative purposes. Bringeland (2015) study indicates that school leaders are interested in using PISA beyond summative intentions, and some school leaders are seeking PISA courses for development and translation. Thus, teachers and school leaders reactions to using PISA results in their own educational setting is of

interest. Reactions can be past present and present. Therefore, the involvement of reflexivity is vital for understanding the modus operandi with PISA, and change of it, which can be indicated by different reflexive modes (variables). Hence, teachers and school leaders are key factors holding reflexive properties (variables) indicating orientations with PISA.

1.9 Aim, overarching research questions, and the scope of the research

A recall for this thesis is to view causality as a necessary connection, which comes with critical realism and Archer's theories. In the Norwegian context there are claims with PISA, both on the macro-level and micro-level that deserves examination. Thus, the aim of this thesis is to reconceptualise alleged PISA effects on the Norwegian educational system by discussing whether the Norwegian educational system has fundamentally changed after the introduction of the PISA test by using Archer's theoretical approach. Furthermore, to get an increased understanding of how school personnel react to the PISA test by using Archer's concepts of reflexivity and reflexive modes. The overall research questions are constructed as following:

- i. How does Archer's theoretical approach enable a reconceptualisation of alleged 'PISA-effects' on the Norwegian educational system?
- ii. How can Archer's concept of 'reflexivity' and 'reflexive modes' increase our understanding of how school personnel (teachers and school leaders) react to the idea of using PISA-test results in their own educational setting?

The first research question concerns a reconceptualisation of previous conceptions of 'PISA effects' on the Norwegian educational system. Among the theoretical concepts used to assess possible 'PISA effects' at the systemic level of the Norwegian educational system are Archer's conceptualisations of structures and processes in state educational systems: unification, systematization, differentiation and specialization, whose predominance varies in different systems. These conceptualisations enable a historical educational analysis, and are often used with Archer's model of morphogenetic cycles to guide the analysis. Thus, the first article (article I) deals with the social macro-level between 1990–2010 in the Norwegian educational system for analysing systemic processes before and after the Norwegian 'PISA shock', which followed the announcement of the first PISA results in 2001 and was accompanied by an intense and extensive public debate.

The second research question address micro-level phenomena, i.e., individuals or incumbents of the system, including one mathematics teacher (article II) and three school leaders (article

III), all working in Norwegian lower secondary schools under a specific type of educational system. The two articles that address micro-level phenomena at the level of the school employ Archer's concepts of reflexivity and reflexive modes, which are related to a person's agency. Article II, which presents the case study of one teacher, additionally illustrates the use of a mixed methods approach in analysing reflexive modes.

Throughout my doctoral period, I realised that Archer's work was comprehensive i.e., her theories on the educational system and reflexive modes, and it would take time to understand her conceptual framework and becoming introduced to her methodologies. In addition, a historical analysis would take time for me as a novice researcher for discussing alleged PISA effects on (internal structure of) the educational system. This article was also time-consuming and scheduled in an anthology. To accelerate the article production, I realised that I could do a reanalysis of my empirical material from my Master thesis to provide new perspectives with the empirical material (Bhaskar [1975] 2008; Bhaskar and Danermark 2006). This enabled a second article for the anthology based on Archer's conceptualisations of reflexivity and reflexive modes. Later, I realised that I did not need a severe group of respondents as this research is a starting point for doing systematic and cumulative research within a specific research paradigm with the PISA phenomenon (Pons 2017). This was one reason for settling with having one respondent in the last article with the same Archerian conceptualisations as the second article. All these decisions were consistent with the previous posed overarching research questions. In this thesis, the chronology of the two articles on reflexivity and reflexive modes have changed order of presentation as the last article to date argues for the use of reflexive modes. This is the reason for it being presented first after article I.

This research should be understood as placing the first pieces to the puzzle of doing systematic and cumulative research within a paradigm that is new to the PISA effects field. This research can't be reduced to illustrative examples of a conceptual framework because the empirical data is claimed to identify real phenomena which can be transferred to other cases, which enables hypothesis and further investigations.

1.10 Thesis outline

Each chapter in this thesis includes a summary. In chapter 1, the background, PISA as a cause, a synopsis of 'what is PISA?', Norwegian PISA researchers assessing 'PISA effects' and future prospects for international large-scale tests, the contribution to the PISA effects field, the theoretical gap within the field, the focus on two social forms and a

reconceptualisation of causality, the reconnection of the reconceptualisation on the two social forms in the Norwegian case, the aim and overarching research questions, and the scope of the research are presented. In chapter 2, the pragmatic literature review related to PISA effects on the macro and micro-level is presented, mainly concerning the Norwegian educational system and incumbents (agency). Chapter 3 outlines the meta-theoretical critical realist foundations and the overall philosophical science outlook, while chapter 4 further elaborates on social realist theory and the ‘common yardstick’ (real definitions and conceptualisations) that guided the research. This includes the morphogenetic cycle for diachronic change, the interplay between the two educational systems, reflexivity and reflexive modes and the causal agent with ideas and materials. Towards the end, this chapter presents the subsidiary research questions that are addressed in the individual articles. Chapter 5 presents the research approach, in terms of case study design, context, evidence and analysis, data collection and piloting, and research ethics. Validation of the research is also discussed in terms of internal and external validity, reliability, and generalisation in this chapter. Lastly, a few critical remarks on theory, design and methods are presented. Chapter 6 revisits the overarching research questions and subsidiary research questions for each individual article with the aim of presenting and discussing results from this research. Finally, chapter 7 details the implications of the research, such as its theoretical, methodological, and empirical contributions. The chapter also sheds light on the limitations of this research and makes suggestions for further research.

1.11 Summary

The first round of the PISA test was conducted in 2000 among a group of 32 OECD member countries. Since then, the number of participating countries has increased to 83 as of 2022, including 45 non-OECD members. Participants are located in all parts of the world. PISA has gained a reputation as ‘the gold-standard’ or ‘the flagship’ from the OECD amongst large-scale international tests, and despite methodological flaws, Norway has participated in each test round since the beginning. In recent years, spin-off products of PISA such as TALIS, PISA for Schools and PISA4U has been introduced from the OECD.

The four characteristics of PISA as an international large-scale assessment that were highlighted in the research literature, and which I found most significant for the reader, were the following: i) PISA’s historical roots back to the World War II, ii) its claim to be a culturally neutral non-curricular test, iii) its role as a policy instrument (with flaws) and iiiii)

its entanglement with global edu-business. However, all these characteristics of PISA have been or can be contested.

PISA and its alleged effects on education are contested both in international and national debates. PISA's damaging effects on education worldwide has been raised as a concern in the international public sphere. An open letter addressed to the OECD director for PISA Dr. Andreas Schleicher signed by over 80 academics from various countries in the world expressed this concern. The open letter was published in the British newspaper *The Guardian* in 2014. Norwegian top politicians from 'across the aisle' have, however, expressed their beliefs in PISA results as an adequate measure of Norwegian educational quality. The Norwegian researcher Svein Sjøberg has even claimed that OECD directs Norwegian educational policy.

However, Norwegian PISA researchers, who for 20 years have worked on the PISA datasets, have on the other hand claimed that the causal effect of PISA on Norwegian education cannot be substantiated. They point to research indicating that the domestic social, cultural, and political context influences national policy more than PISA data and policy advice from the OECD, which substantiates the question if there actually is policy borrowing from the 'reference society'. Furthermore, these researchers underline that no kinds of causal effects can be established from the PISA data, since these data are cross sectional, all data gathered at the same time, and besides, the statistical analyses of these data measure associations (correlations, explained variance, etc.) rather than causes. When researchers use words such as 'influence' and 'cause' when presenting their statistical findings in analysing PISA data, it is just to facilitate communication with their audiences. They are using concepts that strictly belong to another sphere of knowledge. Despite there being no clear causal links in the quantitative realm with PISA and educational changes, there is a dedicated research area named 'PISA effects'. So, whether PISA has had (an) effect(s) on Norwegian education, therefore, remains a contested issue.

However, this research is a specific reply to Xavier Pons for a reconceptualisation of PISA effects to avoid the same novelty of input-output discourse of the 'PISA shock' and to better understand the political meaning of PISA effects on education governance and policy process. Pons propose the use of the model 'reception, uses and effects' and the use of overarching theoretical conceptualisations that held key factors and variables that explains how and why PISA effects occur or not. This research will use an extensive conceptual framework provided

by Margaret Archer for enabling a reply to Pons. Already in my first explorations of the research literature on PISA, I found that a significant number of studies on so-called ‘PISA effects’ focus on policy issues and changes at the macro-level of national educational systems with claims of the Norwegian educational system being impacted by the PISA test. Micro-level PISA effects are fewer, and I wished to make teachers and school leaders reactions to the PISA test clearer. Archer’s conceptual framework will enable comparative systematic and cumulative research which Pons requests within the field of PISA effects.

2. State of the art: macro-level and micro-level research related to PISA effects

2.1 The literature search and sectioning

This literature review is based on already known literature, recommended literature, chain-searches, and searches in databases. It is characterised as being a pragmatic literature review for addressing PISA effects and absences on the macro and micro-level, related to the educational system. The part of the literature review that involved searches were conducted in databases such as ERIC and Oria. The key terms that were used were ‘review’, ‘effect’, ‘countries’, ‘policy’, ‘comparative’, ‘lower secondary schools’, ‘school leaders’, and ‘teachers’. The criteria for inclusion were that the works contained information about PISA effects on countries and school personnel.²⁴ A feature for inclusion was that PISA was being discussed and it was published during the two last decades. Another feature for inclusion were that the publications were normally peer-reviewed²⁵ and that the works were written in English or Norwegian. The intention is to exhibit various research on effects of the main PISA test in different countries on the macro-level, especially on educational systems. Moreover, micro-level effects of PISA on agency, especially on school leaders and teachers. The Norwegian case will especially be emphasised for further legitimising this research and for starting discussions on the macro and micro level.

The first part (macro-level) deals with the research status on macro issues, i.e., whether and how PISA has had effects on educational systems. PISA’s potential impact on the educational system is investigated in article I. This part briefly discusses the current theoretical dispute in international research on the need for a definition of the educational system and pinpoints some reasons for this need. The argument is that Archer’s conceptualisation of an educational system can resolve unclarities and differences about the educational system being more or less centralized or decentralized because of PISA. The educational system as a social form is a central concept in article I. The second part (micro-level) of this chapter deals with research on how PISA has impacted the micro-level of schools, focusing on school personnel. The lack of a concept of agency is not reduced to the PISA research. Thus, this part also discusses

²⁴ In cases where there was little information on PISA in Norway, especially the effects on school personnel, I needed to zoom out from PISA to focus on (the use of) other tests, which is done in article III and mirrored in the discussion in chapter 2.3.

²⁵ I have included my master’s thesis and the survey from the Union of Education Norway, as I see them as vital contributions to the topic in providing information about PISA.

theoretical disputes in international research on (teachers') agency. It argues that Archer's conceptualisations of agency and reflexivity resolve some basic issues that have been raised in these international theoretical debates. The PISA research and other research can possibly benefit from such an application. Agency and reflexivity are central concepts in article II and III.

2.2 State of the art in macro-level research on PISA effects

2.2.1 Pons' review of international PISA research

In 2017, Xavier Pons summed up "Fifteen Years of Research on PISA Effects on Education Governance". He called it "A critical review" and argued for a new turn in research on PISA and its effects (Pons 2017). His review contains 87 references on educational policy from 2003–2016 collected through the peer-reviewed database ERIC (Education Resources Information Center). Three aspects of each article were examined: their theoretical approach, methodology and main findings. Only studies written in the English language were included. The intention was to 'review the literature of PISA effects on education governance and policy processes.' Pons (2017, 132). Pons claims that until 2009 there were few publications on PISA effects, but the number increased after 2010 (Pons 2017, 134). PISA's methodological and ontological mainstays and '[...] education systems'' main features' are not centered in his review (Pons 2017, 132).

There are three reasons that justify Pons review: 1) The 'PISA shock' simplified discourse. In this category there has been a focus on fast implementation of policy that legitimates OECDs 'soft policy'. This has simplified the PISA research to an input-output discourse with the 'PISA shock'. The second 2) justification is that scientific literature on PISA policy effects is little known and little used by the OECD themselves and amongst political actors. Finally, 3) the scientific research on PISA effects is heterogeneous with individual contributions from different disciplines and traditions making their conclusions difficult to compare. Moreover, the 'PISA effect' field is 'fuelled by many individual contributions from various disciplines and academic traditions, or by some specific groups of scholars whose works are rarely confronted in a dialogical and cumulative way' (Pons 2017, 133).

Pons lists 11 themes that enlighten these 87 articles by placing them into one or several of the listed themes. One of these dominant themes is that 'context matters'. For instance,

PISA results do not have the same impact, salience, and influence according to the contexts in which they are disseminated and that characterise each education system. Thus, the aim is to go further in the analysis of domestic education policy process in order to better appreciate the influence of PISA on governments. (Pons 2017, 138).

Pons enlisted 9 contextual factors that explain how countries receive PISA, f.ex., instrumental and strategic uses by the government and the degree of the assessment system's institutionalisation. The scholar concludes that the PISA research needs better conceptualisations of PISA effects, with an epistemology of uncertainty to avoid absolutism, and a normalisation of the research on PISA effects to avoid the same novelty. Pons states that PISA has brought about major changes to education governance worldwide. Pons also claims that 'PISA does not inevitably bring about radical changes in the governance of an education system. [And] PISA is not inevitable per se to talk about current education policy processes.' (Pons 2017, 141). Another element Pons states is what PISA effects are:

Is it sufficient to observe that policy actors use PISA in their speeches or in policy texts in a specific historical period to conclude that there is a PISA net or marginal effect? The French case that we studied is an interesting counter-example: an increasing number of speeches on PISA in a country does not necessarily mean that this survey has an effect or influence on domestic education policy. (Pons 2017, 141).

Pons argues against trivialisation of PISA effects and writes that 'conceptualising a PISA effect requires a deep knowledge [...] and [...] conceptual tools and theoretical frameworks that allow one to take into account both changes at different policy levels and the variety of reception régimes at work in education systems.' (Pons 2017, 141).

2.2.2 Hopfenbeck's review of international PISA research (including Norway)

One year after Pons' review was published, in 2018, Therese N. Hopfenbeck et al. published a more comprehensive review of the international research literature entitled "Lessons learned from PISA".

While Pons' review only included articles on 'effects on education governance', Hopfenbeck et al.'s review covered several other themes addressed by PISA research. Hopfenbeck et al.'s review included articles published in English language peer-reviewed journals from the period of 15 years from January 1999 to September 2015 (the first PISA-cycle in 2000 to the

6th cycle in 2015). Five literature bases (ERIC, PsycINFO, Scopus, Web of Science, and Zetoc) were searched, and three content categories across journals from various disciplines were identified: i) secondary analysis of PISA datasets, ii) policy impact, and iii) critiques.

The content category ‘secondary analysis of PISA datasets’, which contained 404 articles, centres around inequalities related to socio-economic status (SES), such as SES gaps, systemic and institutional parameters, urban-rural locale, family cultural capital, and family structure.

In the content category ‘policy impact and governance’, which contained 144 articles, many of them analysed potential factors, mechanisms, networks, and dynamics driving PISA’s influence on policy and governance. Besides addressing policy and governance impact, articles in this category also addressed curriculum issues, country performance, economics, the media, and national assessments. The articles on ‘policy and governance’ focused mostly on the implications of policy borrowing, shifts in accountability structures, and increasing demands for standardisation as results of PISA.

The content category ‘critiques’, which contained 106 articles, focuses upon cognitive test constructs, design, data processing and questionnaires as well as technical issues (sampling, Item Response Theory, measurement invariance (if there is equal perception of a concept)) and bias in PISA, and curriculum and culture curriculum fairness. The latter involves the OECD claim that there will be no curriculum bias towards the participating countries, as PISA is not built upon these frameworks (Hopfenbeck et al. 2018, 345).

‘[T]he literature lacks a similarly extensive exploration of changes induced across national assessment systems as a result of PISA’. (Hopfenbeck et al. 2018, 346). Hopfenbeck et al. (2018) state that the PISA literature appears to be conflicted: ‘[T]he authors of *secondary data analysis* publications are often building upon PISA data, and the *critique* and *impact/policy* authors pointing out structural weaknesses and cracks in the foundations of ongoing PISA constructions.’ (Hopfenbeck et al. 2018, 347).

Hopfenbeck et al. note PISA’s central role in education policy debates in many countries, but warns against uncritically using PISA-research in policy making, stating that ‘studies based on PISA datasets have led to progress in educational research while simultaneously pointing

to need for caution when using this research to inform educational policy.’²⁶ Further research can examine school climate, assessment cultures in schools and students’ approaches to learning (Hopfenbeck et al. 2018, 348).

2.2.3 Previous research on ‘PISA effects’

2.2.3.1 International comparisons of PISA-effects²⁷

Studies comparing countries illustrate that PISA has had different effects since the turn of the millennium (Grek 2009; Ringarp and Rothland 2010; Bieber and Martens 2011; Breakspear 2012; Baird et al. 2016; Pons 2017; Grey and Morris 2018; Sellar and Lingard 2018; Morgan 2018; Yang and Fan 2019; Hossain 2023).²⁸ These effects are situated on national and federal reforms, policy processes, and adopting PISA as an external measurement tool of educational quality (Breakspear 2012). Studies have also analysed the media’s role and responses to PISA results (Grek 2009; Baird et al. 2016; Grey and Morris 2018; Sellar and Lingard 2018; Morgan 2018). PISA shocks are not only experienced amongst low-performing countries (Grek 2009; Ringarp and Rothland 2010; Sellar and Lingard 2018), however. The policy processes connected to PISA, learning from the best-practices in high performing countries, have been challenged by many studies (Grek 2009; Ringarp and Rothland 2010; Adamson et al. 2017). In some cases, data is ignored or misrepresented, or legitimised policy is contrary to advice from the OECD (Grey and Morris 2018). Convergence with OECD policy recommendations are not uniform across countries (Baird et al. 2016), and the importance and effects of PISA can change over time. Bieber and Martens (2011) stated that the US response to PISA was low because of already established features corresponding to the recommendations from OECD and PISA. Later on, it was claimed that PISA scores in the US needed to improve (‘Race to the Top’), and this focus has created a private educational market where multi-national for-profit companies (e.g., Pearson) are selling targeted educational

²⁶ Therese N. Hopfenbeck, Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J. and Baird, J.-A. 2018. Lessons Learned from PISA: A systematic review of peer-reviewed articles on the Programme for international student assessment. In *Scandinavian Journal of Educational Research*, 62 (3), p. 333. The research presented in the article was supported by ‘Kunnskapssenteret for utdanning’, Norway, which is a government funded unit, linked since 2019 to the University of Stavanger. Originally, ‘Kunnskapssenteret’ was established as a unit in the Norwegian Research Council. Its function is that of a ‘clearing house’, screening and synthesising Norwegian and international research on education to increase quality in the sector. The function of the ‘Kunnskapssenteret’ in providing policy related research may account for why Hopfenbeck et al. address the utility of PISA research for policy in their article.

²⁷ Most articles and findings are inserted from article I.

²⁸ Breakspear represents and worked for the OECD on this publication.

products/packages to American schools (Meyer and Zahedi 2014). In the international literature, the ‘preliminary evidence [is] that PISA is being used and integrated within national/federal policies and practices of assessment and evaluation, curriculum standards and performance targets’ (Breakspear 2012, 27). In Shanghai, an own scientific assessment system has been established based on ‘PISA’s advanced ideas, theories, and techniques of examination evaluation’ measuring green indicators to optimise teaching methods and behaviors (Yang and Fan 2019, 305). However, the Scottish case, throw doubts about PISA’s validity for assessing Scottish education performance. Half of the respondents do not see PISA results reflecting the effects of the Curriculum for Excellence reform. This contradicts the purpose of PISA which is aimed to inform political authorities about effects of reforms and reform development (Hossain 2023).

2.2.3.2 PISA effects in Norway²⁹

In the Norwegian setting, PISA is said to be a very influential test, especially with respect to initiating new assessments and the Knowledge Promotion (LK06) (Breakspear 2012, 19, 24). Moreover, the literature indicates that PISA has had a strong effect in Norway (Mausethagen 2013; Baird et al. 2016; Nortvedt 2018; Sivesind 2019; Thomas 2021). PISA is described as the doctor, Norway the patient, shepherded by the OECD (Thomas 2021). PISA effects are understood when reforms are justified and legitimised with ideas originating from PISA or PISA results. Nortvedt (2018) claims that educational reforms are PISA-driven, in contrast to Baird et al. (2016). All mentioned studies focus on the adoption of a policy, i.e., policy outputs and their justifications, and to some extent the results of policy, i.e., policy outcomes. Some researchers question if PISA reforms threaten equitable and inclusive education (Mausethagen 2013; Nortvedt 2018). PISA is the most frequently mentioned ILSA (International Large-scale Assessment) in Norwegian policy papers and is one of the most common sources used for gaining knowledge (Sivesind 2019). Some researchers claim that processes of decentralization and/or (re)centralization have occurred since or because of PISA (Mausethagen 2013; Baird et al. 2016; Imsen, Blossing, and Moos 2017; Nortvedt 2018). Mausethagen claims that after PISA, a reform in 2006 caused some ‘form of decentralization’ (Mausethagen 2013, 164), in the sense of teachers’ freedom to choose their own teaching

²⁹ Some articles and findings are inserted from article I. Additionally, I have included Imsen, Blossing, and Moos (2017) and Camphuijsen, Møller, and Skedsmo (2021) publications as I considered them relevant for discussions and critique in this chapter.

method. She also mentions that Norway has a ‘highly regulated education system’ (Mausethagen 2013, 162). Baird et al. (2016) claims the educational system is decentralised and that they found no evidence of centralization because of PISA. Imsen, Blossing, and Moos (2017) states that after the millennium the ‘restructuring policy [...] can be described as decentralisation, a strong emphasis on competence aims and students’ learning outcomes, increased assessment and a vast national test system, increased national and local control, and a research-based and expert-based development strategy’ (Imsen, Blossing, and Moos 2017, 573). But Imsen et al. also argues for recentralisation as municipalities ‘report to central authorities about their achievements’ (Imsen, Blossing, and Moos 2017, 574). Thus, the scholars argues that ‘the whole decentralisation system is designed to strengthen state control in an indirect way that results in recentralisation’ (Imsen, Blossing, and Moos 2017, 574). Nortvedt (2018) argues for decentralisation and recentralisation at the same time, after PISA reforms, according to the logic that decentralisation means delegation of local decision-making and recentralisation means accountability measures from the same level. Camphuijsen, Møller, and Skedsmo (2021) argue that test-based accountability (TBA) was welcomed in a ‘highly decentralised education system’ and that PISA results and the ‘scandalisation’ of Norway’s poor PISA results promoted national testing.

2.2.3.2.1 Elaborations on findings of PISA effects in Norwegian studies

The selection of Norwegian studies beneath are included for stressing that the philosophy of science stance and the conceptual framework has consequences for the analysis and discussions. Their absence of the Archerian concept of the educational system with its properties will justify my own research in the Norwegian setting and later provide discussions when comparing different conceptualisations (see chapter 6). I will elaborate on the research that is relevant for the Norwegian educational system, i.e., characteristic claims about the system where PISA effects are situated. Additionally, in relation to article I, I have included Imsen, Blossing, and Moos (2017) and Camphuijsen, Møller, and Skedsmo (2021) publications as I considered them relevant for discussions and critique (see chapter 6).

Mausethagen (2013) has studied shifts in the meaning of the concept ‘competence’ in Norwegian White Papers from 1995 to 2010 through discursive document analysis. She states that the term has changed from emphasising collectivism to a more individualistic conception

of individual performance. She states that this is an effect of adopting the OECDs DeSeCo³⁰ competency definitions to the Norwegian National Curriculum. DeSeCo helped to construct ideas in national policymaking and define what was desirable, necessary, and important (Mausethagen 2013, 174). The author argues that ‘the OECD not only governs through numbers and comparison, but also through what can be described as “governance through concepts”.’ (Mausethagen 2013, 161). Nevertheless, competency was not a new discourse; it was already an established concept within her period of analysis, for instance, in White Paper No. 29 (1994–1995), which was the previous report for L97 (teaching curriculum of 1997). The use of the concept was also present in other White Papers, including those prior to PISA (in teacher education reforms). Interestingly, however, the concept of competence increased from six times in White Paper No. 29 (1994–1995) to over 220 times in White Paper No. 30 (2003–2004), impacting the national curriculum for compulsory school more heavily with time in accordance with DeSeCo’s adaption. Besides the impact of the concept of competence and various conceptualisations of it, Mausethagen also claims that after the introduction of PISA, with a new curriculum reform, the Knowledge Promotion (LK06), some ‘form of decentralization’ has been introduced that entails allowing teachers to choose their method of instruction (Mausethagen 2013, 164). She refers to decentralization as ‘freedom, trust and responsibility’ (p. 164). Previously, she had denoted the Norwegian educational system for being ‘highly regulated’ (Mausethagen 2013, 162).

Baird et al. (2016) examined reactions to the 2009 and 2012 PISA results in Canada, China (Shanghai), England, France, Norway, and Switzerland by contrasting countries that had a high PISA score (Canada and Shanghai-China) with OECD-averages (England, France, Norway and Switzerland). The scholars critically evaluated policy documents, media reports and academic articles in English, French, Mandarin and Norwegian. Through these examinations, the scholars looked at stated reactions to PISA by the governments, which included claims of change of educational policy. Their conclusion is that countries with similar results on PISA had different policy interventions and responses to the PISA results. Baird et al. state that the ‘scandalisation’ or ‘problem pressure’ was evident in four of the six cases as a technique used to motivate change. Five of the six cases showed ‘standards-based reforms’ (used for school accountability systems), and Norway had such responses. In

³⁰ DeSeCo is an abbreviation of ‘Definition and Selection of Key Competences’, the outcome of an OECD project running from 1997 to 2003, ‘Definition and Selection of Competences: Theoretical and Conceptual Foundations’ (Mausethagen 2013, 170).

Norway, the PISA 2000 and 2003 created PISA shocks. Also, the Ministry of Education introduced a series of educational reforms as a response that ended in assessment systems (NKVS/NQAS and skoleporten/school gate) and a new curriculum (LK06) justified by the PISA results. Baird et al. did not find reforms promoting ‘ideal governance’ in Norway, i.e., structural changes introduced to increase school autonomy. Despite these findings, Baird et al. claim that ‘Norway has a decentralised education system, with many of the decisions being made at a local level’ (Baird et al. 2016, 127). Norway has little central intervention supported by the claim of the educational system being decentralised, and thus its assessment system is underdeveloped compared to England and France. Hence, Baird et al. could not find evidence that centralisation was pursued because of PISA results. On the same page, centralisation is understood as ‘policy convergence’ (Baird et al. 2016, 132)

Imsen, Blossing, and Moos (2017) has studied the Nordic Education Model. Their analysis is historical and discursive; they focus on sketches of Denmark, Norway and Sweden’s comprehensive school development. Moreover, they focus on changes in state policy before and after 2000, and on social technologies mentioning shifts in ‘national curriculum plans and learning aims, and control mechanisms such as tests and evaluations systems’ (Imsen, Blossing, and Moos 2017, 570). Lastly, they do a comparison between the three countries on how ‘the basic values of the Nordic model have been affected by recent educational reforms in the three countries’ (Imsen, Blossing, and Moos 2017, 570). In the Norwegian setting, the scholars mentions that PISA results from 2001 created ‘legitimacy’ for educational reforms. They mention that from the millennium the ‘restructuring policy so far can be described as decentralisation’ with focus on ‘competence aims and students’ learning outcomes, increased assessment and a vast national test system, [with] increased national and local control’ (Imsen, Blossing, and Moos 2017, 573). However, school owners were also *given* responsibility for tasks and decision-making at the local level but reporting to central authorities about their achievements. Redirecting that ‘the whole decentralisation system is designed to strengthen state control in an indirect way that results in recentralisation’ (Imsen, Blossing, and Moos 2017, 574). Hence, the scholars present a decentralisation/centralisation paradox.

Nortvedt (2018) studied the policy impact of PISA on Norwegian mathematics education. The scholar considers the impact of PISA on Norwegian educational reforms as strong. She claims that the PISA shock ‘led to’ the implementation of a national quality assessment system and national tests. Moreover, altering of the mathematics curriculum for compulsory school and

mathematics teacher education were also the case, especially with the strengthening of problem-solving, modelling, and mathematical thinking and reasoning. Students in mathematics also show less motivation, lower self-belief, less perseverance, and experience more anxiety than students on average in OECD countries (p. 435). The use of ILSAs for policies and reforming and developing the national quality assessment system and the Knowledge Promotion (LK06) reflects a focus on performance measurement, accountability, decentralisation, and local autonomy. Nortvedt argues that education reforms from 2001 and onwards, after the PISA results were published, have contributed to justify these changes (p. 437). Nortvedt moreover states that PISA may be used to ‘validate existing policy directions’ caused by international trends that are circulating (p. 431). Nortvedt refers to PISA reforms for changes in recentralisation and decentralisation. She mentions that there has been an amendment to the Education Act and a major curriculum reform in 2006, which focused on individualised education and national achievement goals (p. 438). She also mentions national testing and accountability measures. Municipalities and the school leaders had responsibility transferred to the local level with accountability measures. Thus, she argues that decentralisation and recentralisation have occurred simultaneously, however, the reconceptualisation of decentralisation as recentralisation ‘indicates a level of inconsistency’ (Nortvedt 2018, 438, 440).

Camphuijsen, Møller, and Skedsmo (2021) studied how and why test-based accountability (TBA) began to dominate educational debates in Norway in the early 2000s, and how this policy has been operationalised and institutionalised over time. The scholars use a cultural political economy framework and a political sociology-driven approach to policy instruments. The analysis comprehends four white papers and 37 in-depth interviews with top-level politicians, policy-makers, and stakeholders conducted between September 2017 and February 2018. Their findings indicate that the ‘scandalisation’ of Norway’s below-expected PISA results and promotion of standardised testing contributed to national testing in the early 2000s. The scholars found that local actors were pressured to reorient their behaviour due to mechanisms such as increased control and school performance by visibility, benchmarking, and administrative control, that also came with delegated responsibilities. ‘Standardised testing and teacher monitoring and evaluation were once considered controversial and out of step with Norwegian values and traditions’ (Camphuijsen, Møller, and Skedsmo 2021, 625). One the same page, the scholars call this a ‘radical shift in school governance’. The researchers conclude that TBA was a key policy instrument for ensuring equity and quality

standards in the highly decentralised Norwegian education system. However, to a certain extent, equity and quality have been rearticulated as performance indicators based on national and international tests (Camphuijsen, Møller, and Skedsmo 2021, 626). Arguably, these scholars claim that policy-makers now can steer a ‘highly decentralised education system from a distance, by means of outcome measures, visibility, comparison and accountability’ (Camphuijsen, Møller, and Skedsmo 2021, 636).³¹

Below is a table that summarises the scholars possible philosophy of science, stated or indicated, in each publication discussed in this paragraph. The philosophy of science may influence the theories one uses and the definition and conceptualisations of the educational system. The table refers to how either or both decentralization and centralization is understood from the scholars point of view where PISA is contextualised. This is already discussed in the former paragraphs. This table also points to absent Archerian theory which can enrich and correct these scholars’ writings later on in this thesis.

Article	Indicated and stated philosophy of science. Stated is highlighted.	Theory	Concepts and definitions of the educational system in Norwegian context?	Absent
Mausethagen (2013)	Constructivism (p. 162)	OECD, soft governance, national education reform, competence.	Decentralization referred to as ‘freedom, trust, and responsibility’ (p. 164)	Concept of the educational system with its properties
Baird et al. (2016)	Realism (p. 125) and (social) constructivism (p. 122–124).	PISA, policy, scandalisation, standards-based reform, ideal governance	Decentralisation as ‘decisions being made at a local level’ (p. 127), and centralisation as international ‘policy convergence’ (p. 132).	Concept of the educational system with its properties
Imsen, Blossing, and Moos (2017)	Realism and constructivism (p. 569-570).	The Nordic education model, educational equality, neoliberal education policy, educational efficiency	Decentralisation as national and local control - delegating responsibility of tasks and decision-making (p. 573–574). Recentralisation as reporting achievements to central authorities (p. 574).	Concept of the educational system with its properties

³¹ For a slightly different reading of Camphuijsen, Møller, and Skedsmo (2021), see the discussion related to article I in chapter 6.

Nortvedt (2018)	Realism (p. 427).	International large-scale assessment, PISA, policymaking, educational reform, mathematics education	Decentralisation as transfer of responsibilities from the national level (p. 438), and recentralisation as accountability measures (p. 438).	Concept of the educational system with its properties
Camphuijsen, Møller, and Skedsmo (2021)	Critical realism and (social) constructivism (p. 624), and interpretivism (633, 637).	Globalisation, education policy, standardised testing, test-based accountability, New Public Management, policy instrument approach	Decentralisation as either or combined ‘devolution of responsibilities to local education authorities’ (p. 625), and/or steering from distance by outcome measures, visibility, comparison, and accountability (p. 636).	Concept of the educational system with its properties

Table 1: Schematised overview over the macro-level theoretisations related to PISA effects.

2.2.3.2.2 Summing up: the need for a conceptualisation of the educational system

In Archer’s innovative work on the ‘social origins of educational systems’ from 1979, she identified the social form of the educational system consisting of a social structure with causal properties (Archer [1979] 2013). Prominent educational sociologists such as Basil Bernstein and Pierre Bourdieu had been ignoring the educational form and its properties (Archer 1983), a tendency which still lives on within the Norwegian educational field today, seen in my own analysis of a selection of Norwegian studies with PISA. This tendency is also underscored in Skinningsrud (2019). The part of the educational field that is concerned with PISA is characterised by not having a definition of the education system, nor having any conceptual framework built on systemic emergence for analysing internal structures and processes in the system.

A prime example of the ignorance of not defining the educational system is exhibited in Archer’s book review of *Predicting the behavior of the Educational System* by Thomas Green and his associates, where she critiques the authors statement that “the reader will bring to the text an adequate conception of what is meant by “the educational system”” (Archer 1981, 212). Such intuition connected to a structural formation – and even having an adequate understanding of it – is quite controversial, if not bizarre. The question is: is this the same pattern of thinking that the Norwegian scholars have thought for the readers? The ignorance of the definition of the educational system is the impetus of the mainstream conceptualisations of centralization and decentralization, two mainstream concepts that in the Norwegian context has been related to increased decisions at the local level, freedom of methods of instruction, the transfer of responsibility from the central to the local level, and

amendments to the Education Act and the national curriculum, leaving the reader with claims about centralization, decentralization and re-centralization, situated within the same period of time. However, in Archer's terms, centralization and decentralization are traceable to the relative predominance of distinct processes in the system, which questions previous assertions on PISA contributing to either increased centralization or decentralization. The use of mainstream conceptualisations without even caring to provide explicit definitions that relate them to systemic characteristics is problematic because it can make statements of the effects of PISA unsubstantiated and diverge into a multifaceted site of assertions about effects of PISA on (the) Norwegian education (system).

2.3 State of the art in micro-level research related to PISA effects³²

In the international sphere, many secondary Turkish language teachers did not follow PISA results. There is a lack of information and interest with PISA. The reason why some teachers find PISA interesting is due to comparisons with other countries, which some teachers understood as a professional obligation to follow (Dilekçi 2022). Furthermore, mathematics teachers from upper secondary school are reflecting on PISA tasks and why they are difficult for Serbian students. Even teachers have trouble with clearly naming which procedures students need for solving the tasks. Despite that teachers are willing to learn from remodelling PISA tasks, they do not see 'contextualisation of [PISA] tasks as part of their responsibility as mathematics educators' (Radišić and Baucal 2018, 459). In another sphere, PISA tasks are used in constructivist approaches to problem solving following teaching sequences in upper primary school in Ireland (O'Shea and Leavy 2013). PISA has inspired to design PISA-like tasks in Indonesia that are rather contextualised (Nusantara, Zulkardi, and Putri 2021; Aini et al. 2023). PISA tasks has also enabled identification of Spanish pre-service teachers mathematical knowledge for solving PISA tasks with possible implications for teacher training (Sáenz 2009). Interviews with Swedish school coordinators (usually an assistant principal or teacher) on exclusion rates for PISA 2018, which has risen above the exclusion limit of 5 % since 2012, 'indicate that many of them misunderstood the OECD criteria' for exclusion. Neither was exclusions rates followed up by the National Agency for Education (Andersson and Sandgren Massih 2023, 33). Hence, '[a] recalculation of PISA 2018 scores

³² Some articles and findings are inserted from article III.

for Sweden where we assume non-participating students to be low performers show that results are significantly affected' (Andersson and Sandgren Massih 2023, 33).

There are few Norwegian studies that have investigated how PISA had effects on school leaders and teachers in their educational practices. However, there are some PISA studies. For instance, despite risen exclusion rates on PISA tests in Norway (above the limit of 5 % since 2009), 'many Norwegian school leaders see excluding students as positive and beneficial, and are not concerned with its effects on test representativeness and validity' (Aursand and Rutkowski 2021, 16). Hence, school leaders exempt students for minimising feeling of defeat. Also, in Norway, PISA is not used for formative assessment, i.e., assessment for learning in relation to students' progression (Eggen 2010; Bringeland 2015). A lack of method competences has been mentioned as a cause for teachers not further interpreting PISA results (Eggen 2010, 286). Norwegian teachers are also claimed to not follow students up systematically, which in turn creates a 'weak culture' for it (Tveit 2014, 226). There are also difficulties in integrating both formative and summative aspects into one singular test, creating possible unrealistic expectations on certain tests. This also includes National Tests, which teachers have addressed as unsatisfactory for formative evaluations (Tveit 2014). In contrast to the National Tests, PISA is a non-curricular test that is under national control by the Directorate of Education—the test is a part of the NQAS—and further administered through the University of Oslo, to the selected schools, by randomisation.

The OECD, represented by Nusche et al., claims that many schools struggle to use (test) data effectively (Nusche et al. 2011). In international findings, there are indications that data usage amongst school leaders (and teachers) is related to their lack of expertise (Sun, Przybylski, and Johnson 2016; Hornskov, Bjerg, and Høvsgaard 2016). Also, in the US (Sun, Przybylski, and Johnson 2016), and Norwegian context (Skedsmo and Møller 2016), there are expectations that their work should proceed and benefit from evidence-based test results. In the Norwegian context, school leaders and teachers are held accountable for test results (Møller and Ottesen 2011; Skedsmo and Møller 2016; Skedsmo and Mausesthaugen 2017), but some school leaders have seen national tests solely as 'symbolic action', providing little new information (Gunnulfsen and Møller 2017). What about PISA?

Approximately, only 30 percent of the teachers, and only 45 percent of the school leaders, fully or partly agree that PISA measures central aspects of the Norwegian school (Utdanningsforbundet 2008). Despite this, in Norway, PISA creates space for reflections and

discussions amongst school personnel on educational matters (Eggen 2010)—developing their own room for action (Eggen 2011). But fewer teachers and school leaders have looked at PISA tasks (48 %) compared to discussing PISA results (72%) (Utdanningsforbundet 2008). The utility of PISA seems to be two-sided and dependent on the polity level: high level experts close to the government argue that PISA is very influential (Breakspear 2012), while 80 percent of school leaders in the survey by the Union of Education Norway signal that PISA says little about how to improve the quality of Norwegian schools (Utdanningsforbundet 2008). Further, some school leaders feel exploited by the PISA test as they do not receive specific feedback for their school (Eggen 2010), and some school leaders notice resistance from teachers regarding the use of test results (Bringeland 2015; Aas and Brandmo 2018) and their being held accountable for those results (Hopfenbeck et al. 2013; Bringeland 2015). However, Norwegian teachers and school leaders are not rejecting external tests as long as they do not challenge their professional work and values (Skedsmo and Mausethagen 2017).

2.3.1 Other findings on micro-level PISA effects in Norway

Even though all Norwegian studies in the former subchapter has been screened for agency discussions, only a selection of the Norwegian studies will be discussed in this subchapter due to space limitation and for making the point that the philosophy of science stance and the conceptual framework has consequences for the analysis and discussions. Their absences of Archerian agency with its properties will legitimate my own research in the Norwegian setting and later provide discussions when comparing different conceptualisations. Hence, I will present Norwegian micro-level research which has not been discussed to this detail in my articles, i.e., the voices and actions of school personnel in relation to PISA effects. I have beneath also included Bringeland (2015) and Eggen (2011) since I considered these two publications relevant for discussion and later critique (see chapter 6).

Bringeland (2015) has examined if the PISA test functions as a management instrument for and in schools relative to mathematics education. The theoretical foundations for analysing the empirical material were Archer's definition of an education system and her conceptualisations of centralized and decentralized systems. In addition, some of Archer's

culture and agency conceptualisations.³³ However, each of these conceptualisations were not adequately utilised (e.g., their interplay was limited explained). I also used actor-network theory (ANT), formative process models and program-oriented summative results models (Eggen), field and habitus (Bourdieu), structural power (Foucault), communication, language, and discourses (Habermas, Searle, Krüger). The empirical data consisted of OECD and scholars' (experts) discourses surrounding the PISA phenomenon. In addition, four individual semi-structural interviews with school leaders in lower secondary schools were included. Based on these theoretical lenses and data, the findings indicate that for some teachers, PISA functions as a management instrument for and in schools in relation to mathematics education. The PISA test is an instrument for policymakers, but PISA is hard to implement for school leaders in lower secondary schools because of its size (large population) and random sampling of individuals. Three of four school leaders address difficulties concerning who is taking responsibility for the test results. One school leader said it was too big of a project to be used at her school. None of the school leaders complained explicitly about lack of competences for using PISA, but three of the four school leaders answered yes to a question on the need for expert assistance to further apply PISA results. One school leader had attended PISA seminar/courses, and another school leader was interested. Still another reported resistance from teachers for using PISA, and a further three others stated that too little accountability was taken for test results amongst teachers.

Eggen (2010) has, in a group interview study of teachers and school leaders, investigated how they assess the validity of PISA and PISA results for their own teaching practice. Eggen raises the question whether PISA generates a reorientation towards quantitative evaluations or is used in processes that promote learning. Does PISA contribute to summative or formative evaluations in schools? Eggen clarifies that PISA was originally an instrument designed for summative evaluations, namely control and comparisons between countries, and the content of the test is independent from individual schools and national curricula. However, teachers and school leaders use their own judgements in interpreting, valuing, or criticising PISA and

³³ My Master thesis was occupied with other research questions than this thesis. Critical realism brings possibilities of new adjustments to previous perspectives and claims given 'The Holy Trinity' (see chapter 3). Hence, the conceptual framework might be better understood with the possibility to reenter statements for validation or rejection.

PISA results. Moreover, PISA plays a part in local educational discourse just by virtue of figuring in the public domain.

Eggen (2010, 285) study addressed the general problematic of ‘competence for assessment and assessment culture for educational development’. However, the interviews happened to be conducted at a time when the PISA test was receiving much public attention. All the informants in the study mentioned PISA in their interviews. This indicated that PISA is important for school leaders, even when their school is not participating in the test. PISA seems to generate general expectations in the public that teachers and school leaders must face. A teaching inspector (school leader) expressed that we need to have ‘a view on it [PISA], a plan for it, and a strategy for it’ (Eggen 2010, 286). But Eggen herself emphasises that school leaders are not just objects for PISA; rather, they are creative subjects, making room for interpretation, valuing and critique (Eggen 2010, 284–285).

The interviews reported by Eggen show that some school leaders emphasise their loyalty to their educational mandate as specified by the National Curriculum. Many of their tasks are not related to the issues raised by PISA. The respondents considered that much of school life and teacher work cannot be measured, and that certain aspects of school life may be excessively exposed just because it is measurable. They mention that important knowledge is not measurable by tests. PISA augments the importance of certain knowledge areas in ways that occasionally conflict with teachers’ own view of what the school should emphasise (Eggen 2010, 286–287).

One school leader claimed that PISA is not wrong in itself, but the way in which it is used to condemn everything that goes on in the school is disheartening to teachers who are trying to improve their professional skills. It is also painful to witness an incompetent political debate that should instead have been a relevant debate among professionals about how to resolve educational issues (Eggen 2010, 287). Eggen mentions two major effects that the PISA debate has on school leaders. Firstly, since the school is subject to increased public attention, they are forced to legitimise what they do and what they stand for. Secondly, this is demeaning considering the many faceted work that teaching and formation is and the broad mandate from society in terms of education acts, curriculum plans and other regulations they must consider (Eggen 2010, 287–288).

Despite schools having no access to PISA results from their own school (when they are in the national sample), the school leadership is keen to find out what the PISA results mean for them, and they try to get a grip on how they should understand their own organisation in light of the discussion going on in the mass media (Eggen 2010, 288).

However, the teachers and school leaders also mention some positive consequences of the focus on PISA. One of the leaders reports that the emphasis on work plans in connection with reading instruction is an idea that was tested out in their school, and this idea was derived from the PISA test. Another school leader points out that the current emphasis on learning strategies is partly derived from PISA, while others claim that PISA has resulted in increased teacher attention to ‘learning pressure’, as teachers make comments on the current ‘learning pressure’ in their class (Eggen 2010, 289–290).

Some teachers find it unfair that teachers are blamed, since ‘the politicians have got the school they wanted’ (Eggen 2010, 290). They have provided the economic and administrative framework as well as the curriculum plans.

Eggen concludes that teachers confronting PISA-issues in public debates, collegial discussions and encounters with parents are challenged in their professional identities as well as in their professional roles, asking themselves: What are the valid guidelines for professional practice? Is it the content of the National Curriculum, or is it the content of the external test? The teachers’ conceptions of knowledge, learning, and teaching are tied to their local context, which differs from the international and comparative context of largescale international tests, such as PISA. Still, teachers are forced to legitimise their practice and their standing on pedagogical issues in relationship to national results on external tests, such as PISA (Eggen 2010, 294–295).

In another article, Eggen (2011), published an analysis of ethnographic data from three investigations of secondary school principals,³⁴ where she introduces the concept of agency, more specifically ‘teacher agency’. Referring to Wenger’s (1998) definition of agency (specifically for the teaching profession), she defines agency as ‘[...] the opportunity for

³⁴ In this study lower and secondary teachers and a Municipal chief executive officer is also a part of the empirical material of analysis and discussions, but not mentioned in the abstract.

engagement in the social world of learning’ (Eggen 2011, 533), and she focuses particularly on the assessment and evaluation practices of teachers and school leaders.

Based on her research, Eggen considers ‘the pressure for top-down summative purposes and international surveys feeding national educational authorities [as] a challenge [but] [s]ome principals and teachers are developing their own room for action [...] by building up [...] their own toolbox of techniques and tools. The fundamental paradoxes embedded in different purposes and validity concerns are [an ongoing] professional challenge[s] [...]’. She argues that these paradoxes have no general solutions, only particular solutions that depend on each context (Eggen 2011, 541).

Below you will find a table that indicates the possible philosophy of science in each article in the former paragraphs. The philosophy of science may influence the theory usage and the definition and conceptualisations of agency. Moreover, the table refers to what kind of agency is discussed in each publication, which is already mentioned in the former paragraphs. The table pinpoints Archerian absences that can possibly enrich and discuss individuals agency with PISA later on in this thesis.

Article	Indicated and stated philosophy of science. Stated is highlighted.	Theories	Concepts and definitions of agency?	Absent
Bringeland (2015)	Theories indicate realism, critical realism, (social) constructivism (p. 54) , and interpretivism (hermeneutics) (p. 70)	Actor-network theory (ANT). Formative process models and program-oriented summative results models (Eggen. Field and habitus (Bourdieu). Structural power (Foucault). Communication (Habermas), language (Searle), and discourses (Krüger). Archer’s structure and agency concepts. A critical approach.	Concept of primary and corporate agency by Archer (p. 43).	Concept of agency with its properties.
Eggen (2010)	Social constructivism (p. 281), interpretivism (used in the analysis)	Validation, curriculum, and assessment theories within community of practice framework by Etienne Wenger. A critical approach.	How validity is defined could indicate that agency is reduced to ‘judgements’ (p. 282).	Concept of agency with its properties.

Eggen (2011)	(Social) constructivism p. 533–534 , interpretivism (used in the analysis)	Constructivist evaluation models, self-evaluation and the CIPP (Context–Input–Process–Product) model, ethical and democratic aspects of evaluation (within community of practice framework by Etienne Wenger). A critical ethnographic study.	Definition of agency by Wenger (p. 533): ‘[...] the opportunity for engagement in the social world of learning’.	Concept of agency with its properties.
--------------	---	---	--	--

Table 2: Schematised overview over the micro-level theoretisations related to PISA effects.

2.3.2 The quest for a concept of (teacher) agency (with properties)

The results of the Norwegian micro-level studies on PISA with school leaders and teachers reveals that the concept of agency with its properties is absent. Also, in the international literature on measurements and professional (teacher) agency there is a puzzle for conceptualising agency (with properties). For example, the issue about the usefulness of tests in the education process has been extensively addressed and debated in international research. Gert J. J. Biesta, in his ground-breaking work from 2010 ‘Good Education in an Age of Measurement’, argues that the current emphasis on ‘what works’ in education and the idea that educational practice should be ‘evidence based’ is insufficient as a guideline for teaching practice and thereby questions the value of all the internal and external tests that are administered in schools:

If we really want to improve the relationship between research, policy and practice in education, we [...] need an approach in which technical questions [‘what works’] can be addressed in close connection with normative, educational and political questions about what is educationally desirable. (Biesta 2010, 49).

In other words, Biesta emphasises the need to consider teachers’ normative and ethical concerns, bringing forward an idea of a theoretical framework that includes such elements (found in Archer’s reflexive modes theory).

In a more recent study of teacher agency in connection with the implementation of Scotland’s large scale educational reform ‘Curriculum for excellence’, Biesta et al. focus on how teacher beliefs motivate and drive teachers’ action, i.e., their results indicate that ‘a clear and robust

professional vision of the purposes of education' play an important role in teachers' work.³⁵ However, the development and activation of teacher agency does not just rely on individual beliefs that teachers bring to their practice, it also requires collective development and consideration (Biesta, Priestley, and Robinson 2015, 624).

Biesta and his associates claim that there is a tension worldwide in educational policy between the two opposite ideas of reducing opportunities for teachers to exert judgement and control over their own work and the contrary view of seeking to promote teachers' own judgements. Countries differ in deciding between these seemingly contradictory views. Some see teacher agency as a weakness within the operation of schools and seek to replace it with evidence-based and data-driven approaches, whereas others argue that due to the complexities of situated educational practices, teacher agency is an indispensable element of good and meaningful education. Agreement with the latter point of view will make it 'important to understand the dynamics of *teacher agency* and the factors that contribute to its promotion and enhancement' (Biesta, Priestley, and Robinson 2015, 624, my italics). Despite the worldwide quest for 'what works' in education, Biesta et al. claim that in the UK and elsewhere, there is an emerging tendency in curriculum policy to acknowledge the importance of teachers' agency – that is, teachers' active contribution to shaping their work and its conditions.

Others have followed Biesta's lead, for example, a group of Finnish educational researchers (Eteläpelto et al. 2013). Contributing to the discussion on 'teacher agency', they note that despite increasing attention being paid to the notion of 'agency' there is disagreement on how the concept should be defined. The aim for these researchers' is to work out a viable definition of 'teachers' professional agency'. They examine the major research traditions' use of the concept, looking for elements that may inform their own definition of it, while paying special attention to ontological premises and how the relationship between the individual and the social is defined. They examine four traditions: i) the social science tradition, ii) the post-structural tradition, iii) the sociocultural approach and iv) the identity and life course approach.

³⁵ There is a similarity between Biesta and Archer's emphasis on aims, ends, and values. But although Biesta and Archer agree on the importance of considering people's (including teachers') ethical concerns, Archer does not to the same extent as Biesta emphasise the role of language in expressing these concerns. Archer primarily investigates how such concerns are expressed in actions and activities, but also how these are justified.

Discussing the fruitfulness of the various previous approaches to agency in understanding ‘professional agency’, Eteläpelto et al. conclude that ‘[...] social science discussion[s] on agency [...] can significantly contribute to our understanding of the contextual constraints, structures, and resources for agency in working life’. However, ‘they are not in themselves sufficient’ [...], since they do not ‘provide conceptual tools for understanding and developing professional identities and subjectivities[,] which – in the broad sense – are central for professional learning; this is especially the case in domains such as education, health care, and creative work, where employees need to act as whole human persons, containing emotions and ethical commitments’ (Eteläpelto et al. 2013, 51).

Eteläpelto et al. emphasise the specificities of ‘professional agency’, mentioning professional and work-related identities. Their distinction between the general concept of agency and the conception of ‘professional agency’ implies the assumption that agency differs in different social positions and roles.

This group of researchers list seven propositions about how ‘professional agency’ should be understood. Among these are the criteria that ‘professional agency is always exercised for certain purposes and within certain (historically formed) socio-cultural and material circumstances’; the practice of professional agency is closely intertwined with professional subjects’ work-related identities comprising their professional and ethical commitments, ideals, motivations, interests, and goals; and professional subjects’ unique work experiences, knowledge and competences function as individual developmental affordances and individual resources for the practice of professional agency at work. Moreover, professional subjects have discursive, practical, and natural (embodied) relations to their work; and professional agency is especially needed for developing one’s work and work communities and taking creative initiatives. It is also needed for professional learning and the renegotiation of work-related identities in (changing) workplaces (Eteläpelto et al. 2013, 62).

Most of Eteläpelto et al.’s criteria are uncontroversial, but one of them is not, which is the criterium that ‘individuals [professional agents] and social entities are analytically [separable] but mutually constitutive of each other’ (Eteläpelto et al. 2013, 62). This is a central point to Margaret Archer in her discussion on how individuals and social entities are related. In her view, and in critical realism generally (further elaborated in chapter 3), individuals and social entities belong to different emergent strata in social reality and possess different properties and powers. Individuals and social entities (social structures) are therefore not mutually

constitutive. Conceiving entities at these different levels as mutually constitutive makes it impossible to study them independently and investigate their separate contributions to events: to what extent have individual persons' values and commitment contributed to the event, and to what extent have social entities, i.e., social structures, constrained or enabled the events in question?

Several authors, among them Richard Sennett, whom Eteläpelto et al. refer to, have raised the discussion concerning whether fundamental changes in working life, including short term employment and the demands for flexible, entrepreneurial, self-responsible, and creative employees, have consequences for employees' personal values and commitments. Sennett argues that increasingly insecure working conditions leads to 'the corrosion of character' (Sennett 1998), i.e., a fundamental change in individuals' normative commitments. Against this assertion, others have claimed that understanding current adaptations to working conditions will be inadequate without considering the worker's personhood and subjectivity, which are fundamental and enduring (Eteläpelto et al. 2013, 62). Empirical studies investigating lifelong learning have, for example, shown that individuals, although placed in work situations constructed by others, still preserve an enduring sense of self, and moreover, try to 'be themselves' both in their work and in their lives outside work. They do not continually reconstitute their selves in accordance with changing environments. Rather than being captive to external forces, people strive 'towards securing a "sense of self" and "being themselves"' (Billett and Pavlova 2005, 199–200).

Seeing individuals as bearers of relatively enduring orientations and values, that are also expressed in their work context, *may* reduce the need for a separate concept of 'professional' or 'occupational' agency. Since the considerations of agency, in a general sense, will include the person's valuing of his/her work, as well as his/her prioritising leisure or family and friends, to designate a separate professional or occupational agency may seem theoretically unnecessary, or even counterproductive, by partitioning domains that should be treated as constituting a totality. Moreover, if, as Archer suggests, personal identity (as well as self-consciousness, thought and emotionality) are prior to and more basic than our social identities (Archer 2000, first page), there are even more reasons to study 'agency in general' among teachers, rather than 'professional agency' or any other aspects or types of agency.

2.4 Summary

The main problem in previous research on PISA effects seems to be a lack of concepts that are adequate for studying both micro-level and macro-level effects of PISA in domestic contexts aggregated through a philosophical scientific paradigm. Concepts for analysing macro effects should go beyond the level of how PISA is used to justify policy, that is, how it impacts policy outputs. They should encompass whether and how these policies have effects on the system itself, in other words, whether and how reforms and interventions legitimised by PISA, lead to anticipated changes in the system. But without defining the system and conceptualising its internal processes it is difficult to assess whether and what kinds of changes have occurred. The most frequently used terms for characterising systems in previous research are educational centralization and decentralization, but these terms are, unfortunately, seldom defined explicitly, and their implicit meanings diverge.

At the micro-level there seems to be an even more serious lack of theoretical and conceptual tools for assessing how PISA affects agents (school personnel) in their work. It has been documented that PISA is a topic for discussions among school personnel and they are in some instances held accountable for the national PISA results by parents and in their local community. Teachers and school leaders have opinions and reactions to this. Some think it is unfair to be held accountable for Norwegian PISA results. However, how PISA makes an impact through for example the National Quality Assessment System and how PISA results are communicated internally in the system and has effects on how work tasks are dealt with, have not been systematically studied and analysed through the concept of (reflexive) agency.

Therefore, embarking on a journey to develop the PISA research and ‘PISA effects’ in a frame of a philosophical scientific paradigm such as critical realism and further utilising Archer’s conceptualisations of macro and micro phenomena within the same conceptual framework to the study of PISA effects seems to hold promises that deserve to be explored.

3. Meta-theoretical underpinnings: critical realism

3.1 Critical realism

The philosophy of science concerns ontological, epistemological, methodological, and axiological questions and claims (Bhaskar and Danermark 2006). It concerns the basic assumptions that guide our research. In the social sciences, there are several research traditions on offer based on diverging ontological and epistemological assumptions. The most common traditions are positivism and constructivism.³⁶ Critical realism is a post-positivistic meta-theory that includes other meta-theories (Bhaskar and Danermark 2006). It is also a philosophy of emancipatory change (Price and Martin 2018).

This meta-theory was selected as the basic framework for this thesis. One crucial reason for this choice was that Xavier Pons, in a review article on ‘Fifteen Years of Research on PISA’, concludes that there are three types of challenges facing researchers on PISA effects: theoretical, epistemological, and methodological. The theoretical challenge concerns the notion of PISA effects itself. Pons asks the rhetorical question: What is a PISA effect, actually? The epistemological challenge, on the other hand, concerns the need to preserve uncertainty about PISA effects and not take for granted a series of such effects. It entails adopting ‘an epistemology of uncertainty’. It allows perspectivism and considers knowledge as fallible, i.e., knowledge is continuously revised, falsified and developed. The methodological challenge’s purpose is to ‘normalise’ the research on PISA and to produce research that is more systematic and cumulative (Pons 2017, 140–141). The meta-theory critical realism seemed to fit the bill presented by Pons. It is a theory of causation that may be applied to refine the notion of PISA effects. It is dedicated to the cumulative production of scientifically sound knowledge, acknowledging that scientific progress can be made by discerning between worse and better theories based on rational judgements.

In 1975 Roy Bhaskar published the book *A realist Theory of Science* (Bhaskar [1975] 2008), which developed a new position within the philosophy of science termed ‘transcendental realism’. Bhaskar’s new position resolved some issues in classical empiricism and rationalism. Essentially, transcendental realism revindicated ontology as a legitimate

³⁶ Moderate constructivism (i.e., humans create society) is compatible with critical realism, while radical constructivism (i.e., everything is constructed and there is no objective reality) is incompatible with critical realism (Sayer 2000a; Al-Amoudi and Willmott 2011).

metaphysical issue for the philosophy of science. In 1979 Bhaskar issued another book on the philosophy of the human sciences *The possibility of Naturalism*, which addressed the ‘old problem of naturalism’, i.e., whether social objects can be studied in the same way as natural ones, that is, ‘scientifically’. In *The possibility of Naturalism* Bhaskar sets out to refute the currently dominant positivist and hermeneutical traditions in the philosophies of the human sciences and puts forward an alternative philosophy of ‘critical naturalism’, where, to put it simply, meaning, understanding, and interpretations are considered as essential human properties that a philosophy of the human sciences must incorporate, while also preserving the essential qualities of being sciences. Over time, the term ‘critical realism’ was increasingly used as a common designation of ‘transcendental realism’ and ‘critical naturalism’, and Bhaskar – after the fact – agreed to use this as a common name for his two related philosophies of the natural and human sciences.

The basic principles of critical realism that differentiate it from other current philosophies of science may be summarised as follows: *ontological realism*, *epistemological relativism*, and *judgemental rationality* (humorously called ‘The Holy Trinity’ of critical realism’) (Bhaskar 2016). These are concepts that concisely convey critical realism’s stance in relation to basic positions in philosophies of science. *Ontological realism* entails the conception of a real world that exists independently of us, our consciousness, experiences, thoughts, and language. *Epistemological relativism* entails the conception of knowledge (of the real world) as a social product produced by human activity. As a human product, our knowledge is fallible and subject to subsequent corrections. It entails perspectival relativity in the sense that we always view the world from some historically transient epistemic framework. *Judgemental rationality* entails that not all knowledge, including scientific theories, is seen as equally valid. It is possible to discern between knowledge that is more or less valid and theories that have more or less explanatory power (Bhaskar 2011). If there are no agreed standards on which to judge the validity of opposing or divergent claims, all scientific arguments become meaningless. Judgemental rationality is contrary to seeing different theoretical paradigms as incommensurable.

Critical realism is different from other branches of realism as it contains an ontology that understands reality as *stratified*, *differentiated*, and *open*. A reality that is stratified consists of different strata with different structures, generative mechanisms, and tendencies. The notion of a stratified reality is based upon the concept of emergence: over time, new strata emerge with their own properties (Bhaskar 2016, 32–33). The idea that new strata with their own

structures and mechanisms emerge diachronically over time, is opposed to reductionistic theorising, which thinks that all complex entities can be reduced to more elementary and basic units. In the 1930s, a novel idea was that all social sciences could in principle be reduced to physics. That reality is differentiated implies that the domain of ‘the actual’ (events) and the domain of ‘the empirical’ (experiences/observations) are sub-domains of the real, which also contains generative mechanisms at a deeper level that are not observed or may not even be observable. The unobserved or unobservable mechanisms, which are part of a deeper layer of reality are, however, fundamental to understanding and explaining change (See Table 1 for further clarification).

Finally, critical realism holds that social reality is an open system in the sense that many different powers exist and are exercised simultaneously. In open systems, universal regularities are not possible because mechanisms other than the one under study may interfere. That is why lab experiments are conducted as closed systems studying one causal factor or a variable (mechanism) at a time. If society was governed by universal regularities, social change through history would have been impossible.

Lab experiments are aimed at discovering and activating regularities and universal laws that must also be valid outside the closed system of the laboratory. Otherwise, technology would be inconceivable (Bhaskar [1975] 2008, 33ff). For various reasons, practical as well as ethical, it is very difficult, if not impossible, to conduct experiments that operate as closed systems on human beings. Therefore, other methods are developed in the social sciences. The general principle is that research methods must be adapted to the properties of the objects being studied.

Critical realist philosophy in educational science is not widely known at Norwegian universities (Nyhus, Annamo, and Jakobsen 2019), and international education research is dominated by quantitative methodologies (Scott 2014). However, for the discovery and investigation of mechanisms, qualitative data are generally better suited than quantitative data. Thus, this research mainly relies on qualitative data from different levels of social reality.

3.2 Intransitive and transitive dimensions of reality

Bhaskerian realism distinguishes between two dimensions of reality (Bhaskar [1975] 2008, 21–23). The intransitive dimension refers to ontology, things that exist, that have tendencies

regardless of our knowledge of them, as well as structures and mechanisms that we often cannot directly observe; hence, they exist and act independently of our beliefs about them. The transitive dimension refers to epistemology: our knowledge about the things that exist. The transitive dimension comprises materials of science, such as established facts and theories that are produced by social and conceptual means. The intransitive and transitive dimensions are thus analytically divided aspects of reality, as both dimensions, the intransitive, and the transitive, are real. The distinction between them makes it understandable why we sometimes are mistaken about reality; our knowledge is fallible.

3.3 Three domains of reality

Bhaskar ([1975] 2008, 56) advocates for three ontological domains of reality: the empirical, the actual, and the real. The empirical domain comprises our experiences and observations; the actual domain contains the phenomena that occur, regardless of whether they are experienced/observed; and the real domain, in addition, consists of structures and mechanisms that create actual events. The inclusion of structures and mechanisms in the domain of the real means that reality is more than what we can observe. To investigate this, we need to study our empirical findings and use retroductive reasoning.

In critical realism, the primary aim of scientific investigations is the identification of structures and their generative mechanisms, which tendentially generate change (Buch-Hansen and Nielsen 2005). Bhaskar’s model of different domains of reality is an inclusive ontological model, i.e., all the ‘objects’ (mechanisms, events, and experiences) are real, but the mechanisms belong to a layer that is deeper (unobservable) compared to that of events and experiences. Bhaskar’s model of the domains of reality belongs to the intransitive dimension and the transitive dimension of reality (Bhaskar [1975] 2008, 242).

	The real domain	The actual domain	The empirical domain
Mechanisms	X		
Events	X	X	
Experiences	X	X	X

Table 3: The Three Domains of Reality. Source: Bhaskar ([1975] 2008, 13).

3.4 Generative mechanisms

Bhaskar asserts that '[t]he real basis of laws is provided by the generative mechanisms of nature. Such generative mechanisms are, it is argued, nothing other than the ways of acting of things' (Bhaskar [1975] 2008, 14). Generative mechanisms are properties of objects and how they act. Structure, agency and culture consist of generative mechanisms that are transfactual, i.e., they are latent in social reality and have causal powers, regardless of whether they are activated—what I term 'built-in capacities'. Objects with their mechanisms possess tendential capacities. These can be found at different levels of scale, i.e., macro-, meso- and micro-levels, in social reality. Moreover, the social reality is open and differentiated, thus consisting of several causal powers. Generative mechanisms are inherent in structures, have possible tendential effects and explanatory power. If there were no tendential effects of objects in different levels of reality, there would be no reason to conduct emancipatory scientific work towards a) social transformation and b) changing the mechanisms/properties and thereby the potential causal power/tendency of the (generative) mechanism.

3.5 Tendencies, causality, and effects in open systems

Bhaskar ([1975] 2008, 14) states that '[t]endencies may be regarded as powers or liabilities of a thing which may be exercised without being manifest in any particular outcome.' Structures of objects possess tendencies per se because of their mechanisms/properties and these can lie dormant as dispositions, i.e., they are not always exercised or actualised but are indeed activated when some sort of agential mediation begins. In critical realism the aim of scientific investigations is to identify generative mechanisms, which possess causal powers.

Mechanisms in open systems co-determine action/events and are always activated by agency. In open systems, such as society, there are several mechanisms with causal powers operating at the same time, which may annul, reinforce, or modify each other. Therefore, according to critical realism, the Humean tradition, which understands causality in open systems as the constant conjunction of events, is a mistake and inadequate.³⁷ The Humean way of thinking

³⁷ In critical realism, sequences of events, where A (PISA) has direct and constant effect upon B (a reform), is impossible, because A requires activation to even have such an effect (multi-causality), and A (PISA) might not be used for justification/reason all the time. The absence of social stratification and causal necessities makes causal theorising flat. What we can ask is; what are the necessary connections that are needed for the reform to be existing. These are the deep causes we want to 'reveal' and that can explain causal necessities: structures and their processes ongoing in the system. In other words, critical realism seeks deeper causes (structures and processes) creating (latent) tendencies, whether they are observed or not, co-creating (an) effect(s), and questioning if (the time-series of) events within the system would have been different. In short, critical realism sees causality as necessities and not constant sequences of events (causal connections/associations).

about causality does not make the distinction between open and closed systems (Bhaskar [1975] 2008, 14).

What causality is, then, is a philosophical debate between different schools of thought. Groff (2004) has summarised the differences between the critical realist and two other major philosophical traditions' concept of causality by pointing out that critical realism considers causality as a feature of the real world (the intransitive dimension), which is contrary to both the Humean and the Kantian tradition. Hume did not consider causality as inherent in the external world but regarded it as a subjective expectation about the constant conjunction of two events, persisting over time, while Kant considered causality to be a category of the mind, which was imposed on external reality by the human consciousness.

Groff (2013) points out that arguments for an alternative to Hume's empiricist ontology and concept of causality, namely a powers- and dispositions-based ontology and theory of causation, started to emerge in the late 1960s and early to mid-1970s. Rom Harrè and E. H. Madden's book from 1975, *Causal Powers: A Theory of Natural Necessity*, was published during the same year as Bhaskar's *A Realist Theory of Science*. Harrè had been Bhaskar's PhD-supervisor at Oxford. However, in the mainstream of analytical philosophy, this position on ontology and the theory of causation gained prominence only as late as the 1990s, when a second wave of dispositions- and powers-based realism emerged, represented by among others by Nancy Cartwright, Stephen Mumford, Brian Ellis, John Heil and Anjan Chakravartty. In recent years, however, interest in ontologies based in dispositions and powers has intensified, indicated by the growing number of conferences addressing such questions as well as special issues of academic journals debating this topic. Groff summarises the main tenets of powers-based ontologies as follows: i) belief in irreducibly dispositional properties, ii) realism about causality, and, for some, iii) essentialism (Groff 2013, 7).

In agreement with critical realism, Mumford and Anjum (2010, 143) challenge the conception of causation as invariable sequences of events claiming that still '[m]any contemporary treatments of causation follow from Hume [...] as he was [...] understood prior to the "New Hume" debate.' Against Hume's understanding of constant conjunctions of events based on observation, which is an atomistic understanding of it—and also found in correlation and regression analysis based on numeric variables—Mumford and Anjum argue for causal dispositionalism. This is a theory of causation based on a meta-physics of real causal powers, or dispositions, and is non-Humean (associated with Aristoteles and Aquinas).

The three ontological entities of structure, culture, and agency (which are the basic entities in Archer's social ontology) are seen to possess real causal powers. Moreover, in social reality causal powers are pulling and pushing on each other. This does not mean that effects do not occur in the open system. The fundamental perspective here is that in open systems there are continuously several causal powers at work (multi-causality), which may counteract or reinforce each other. An effect is an outcome, 'produced by many powers acting together' (Mumford and Anjum 2013, 221).

3.6 Power1 and power2 relationships

Critical realism holds that relations between positions and institutions have effects on people occupying institutional positions and to some extent regulate their behaviour (Archer 1995; Donati and Archer 2015; Bhaskar 2016). Bhaskar sees the primary object of the social sciences not as individuals or groups of individuals but as relations. Power differentials are an element in most social relations. Relations of power are social mechanisms that make a difference to the 'possible actions between two persons.' (Al-Amoudi 2007, 562). But there is a difference between Power1 and Power2. Power1 has transformative capacity, while power2 is the capacity to dominate and oppress. Power2 is modelled on Hegel's master-slave relationship, but is by Bhaskar generalised to include relations based on gender, ethnicity, age, disabilities and social positions (Bhaskar 2016, 55, 131). Both types of power could be present in both formal or informal relations within organisations (such as school leader–teacher, manager–worker or husband–wife relations). Power1 and power2 can be present in the same relations as structural enablements and obstructions, respectively. '[E]mancipation from power2 relations will in general depend on an augmentation of the transformative capacity or power1 of the oppressed' (Bhaskar 2016, 55). This will partly consist 'in knowledge of the power2 relation, that is, of the explanatory structures and mechanisms that account for power2', and under which conditions they can be transformed (Bhaskar 2016, 55). The goal of emancipation from any type of master-slave relations is not 'primarily the liberation of those who are slaves, but the overthrow of the master-slave relation itself' (Bhaskar 2016, 56). Power-relations of both kinds are found in the educational system.

3.7 Reasons as causes

Critical realism considers reasons as causes in the sense that reasons for subsequent action may be considered as causes for these actions. Conceiving reasons as causes entails bridging the gap between naturalism and hermeneutics (interpretivism) in the social sciences, claiming that the concept of causation applies not only to the natural but also to the human sciences.

The tradition of hermeneutics denies that causation is a valid concept in the human sciences. Bhaskar emphasises that all social sciences must include people's own understanding of what they do when they do it. This is a basic step in any social inquiry (Bhaskar [1979] 1998, 154). He also considers social structures as concept-dependent (Bhaskar [1979] 1998, 38). Both Bhaskar and Margaret Archer emphasise that causal power of social forms [structures] are realised through human agency (Archer 1995). Bhaskar claims that 'social forms are a necessary condition for any intentional act' (Bhaskar [1979] 1998, 25), and Archer further elaborates this point by suggesting that the generative powers of social structure and culture 'are mediated to people by shaping the situations in which they find themselves', thus providing reasons for actions and 'directional guidance' (Archer 1995, 196). Archer focuses on the importance of situational logics for providing reasons for action, together with personal commitments and reflexive modes, and that retroduction may be applied for identifying deeper causes, such as beliefs, myths, ideology, political positioning, pedagogical pledging, power relationships, health, economy, resources, systems, etc.

3.8 Abduction and retroduction

Abduction and retroduction are modes of inference that link individual observations to general theoretical concepts. These modes of inference are supplements to the more commonly used modes of inference, e.g., deduction and induction (Danermark et al. 2002, 79ff). The American pragmatist Charles S. Peirce suggested abduction as a mode of inference and defined its logical form, which is comparable to deduction and induction, where the first premise is logically true, but the second premise is only probable. Abduction has in more recent years been conceived as 'redescription or recontextualization', which is compatible with Peirce's use of the term, since he considered abduction as more than strictly logical derivations but also as a way of reasoning in a wider sense (Danermark et al. 2002, 89). Abduction is used in this thesis in the sense of redescriptions and recontextualisation of data and prior findings. 'By means of abduction we recontextualize and reinterpret something as something else, understanding it within the frame of a totally different context.' (Danermark et al. 2002, 96).

Critical realism is committed to 'transcendental arguments'. This is a type of argument initiated by Immanuel Kant, which addresses the presuppositions of human practices. A transcendental argument asks: '[W]hat must be the case for some feature of our experience to be possible, or more generally[,] what must the world be like for some social practice (as conceptualised in our experience) to be possible[?]' (Bhaskar 2016, 3). Bhaskar considers this

to be a sub-category of a more encompassing category of questions, which play an important role in science, which he calls *retroductive arguments*. ‘A retroductive argument asks what would, if it were real, bring about, produce, cause, or explain a phenomenon; and retroduction is the imaginative activity in science by which the scientist thinks up causes, or [...] generative mechanisms which, if they were real, would explain the phenomenon in question.’ (Bhaskar 2016, 3). The phase in the research process when retroductive arguments are used is only one of several phases. Another phase is when alternative explanations are eliminated.³⁸ However, ‘science moves continually from the description of phenomena to the retroduction of possible explanatory causal mechanisms for them’ (Bhaskar 2016, 7).

Blom and Morén (2011), who have studied social work practice with the aim of discovering causal mechanisms, exemplify retroductive questions as follows: ‘What is fundamentally constitutive for the structures and relations (X) that are studied? How is X possible? What properties must exist for X to be what X is? What causal mechanisms are related to X?’ (Blom and Morén 2011, 70). Andrew Sayer (2000b) suggests asking counterfactual questions instead of the common practice of associational thinking in order to discover causal relations, such as the question: can one thing exist without another?

3.9 Summary

My choice of critical realism as the meta-theoretical foundation for this thesis is supported by Xavier Pons’ quest, featured in his review of ‘Fifteen Years of Research on PISA effects’. It comprises i) a better definition of PISA effects, ii) not taking for granted that there are specific PISA effects, and iii) the ‘normalization’ of PISA research, making it more systematic and cumulative. Critical realism answers the requirements raised by Pons: it has an explicitly worked out and well-argued theory of causation; it advocates scientific principles that include the rational assessment of the relative merits of different explanations of phenomena and therefore enables cumulative research.

Critical realism, which is a post-positivist philosophy of both the natural and the human sciences, adheres to three basic principles (The Holy Trinity of critical realism), which are the following: *ontological realism*, *epistemological relativism*, and *judgemental rationality*. These principles entail that i) reality exists independently of our knowledge of it, ii) knowledge is a human product, which is fallible and therefore corrected and improved over

³⁸ See Bhaskar (2016, 79) for a full description of the whole research process.

time, and iii) it is possible to discern between better and worse theories and explanations based on rational criteria, such as the explanatory power of theories.

A basic feature of critical realism is its vindication of ontology (conceptions of what the world is like) as a legitimate concern for philosophies of science. This standpoint repudiates the view that philosophies of science should only be preoccupied with epistemological questions, i.e., how knowledge is produced. Bhaskar's argument is that ontological presuppositions are always present in philosophical and scientific works, even when they are not explicitly addressed. Therefore, critical realism distinguishes between the transitive and the intransitive dimensions, in reality and in science. The intransitive dimension refers to reality that exists independently of human knowledge and is the object of science. The transitive dimension refers to the social processes that produce knowledge – epistemology.

Based on an analysis of the classical scientific lab experiment, Bhaskar developed an ontology that claims that reality is structured, differentiated and open. That reality is structured means that it is layered in hierarchical strata, based on the principle of emergence. Each stratum is defined by specific structures and mechanisms. That it is differentiated means that there are different but overlapping domains of reality: the domain of the real, which is all encompassing, incorporates the domains of the actual (events that occur, but are not necessarily experienced) and the domain of the empirical (measurements and experiences).

The operations of generative mechanisms must be understood in terms of the exercise of tendencies and causal powers. 'Tendencies may be possessed unexercised, exercised unrealized, and realized unperceived (or undetected) by men.' (Bhaskar [1975] 2008, 184). A feature of critical realism that is of particular relevance in the context of the present thesis is its theory of causality. A predominant conception of causality since Hume has been to see it as based on sequences of events. Repeated observations of a regular succession of events, plus expectations of such sequences' reoccurrence, are seen as the basis for our conception of causality in much of contemporary mainstream educational and social research. With this conception, to cause is to be that which invariably comes first of two consecutive events. Opposed to this conception, making an alternative, is a powers-based theory of causality associated with dispositions and critical realism, which conceives causation as the expression of the power of things, possessing dispositions that may or may not be activated. Effects are caused by several powers operating, creating an outcome (multi-caused/co-caused).

Two other distinctive features of critical realism are that reasons should be considered as causes, i.e., reasons are causes for individuals' actions and may have material effects, e.g., the martyr who gives his life for a cause is one extreme example. Another feature is the distinction between two types of power, Power1 and Power2, where Power1 often is set to enable action and Power2 constraining action. Social relationships where Power2 predominates are akin to oppressive master-slave relationships.

Central to scientific work informed by critical realism are the two methodological procedures, or modes of inference, abduction and retrodution, which link observations (data) to theoretical concepts. Abduction entails the redescription or recontextualisation of observations (data or phenomena), i.e., conceptualising something as something else. Retrodution is another mode of inference that seeks to identify causes or mechanisms that can explain observed data or events and may also involve counterfactual reasoning.

4. The key conceptual framework

4.1 A system theory and the causal agent

The choice of the theoretical framework for this thesis can be justified by the need for overarching theoretical conceptualisations in the ‘PISA effects’ field that could grasp possible systemic effects on the parts and people. Thus, there was a need for an educational system theory that incorporates the parts and the people. Archer’s general theory on structure, agency, and culture (SAC) seems to fit the bill due to autonomy, emergent properties (generative mechanisms), and causal power of each entity, where agency is the causal agent for change. For example, Luhmann’s theory of systems without causal agents is difficult to accept. Further, Pons made a request for systematic cumulative PISA research that emerges from a certain paradigm. This often entails comparative perspectives, which would be a side effect of embarking on such studies. The comparative perspectives may not seem obvious, since its data are restricted to one country, Norway. However, studies of one country may be of value in a comparative perspective if they employ concepts that are applicable across countries; comparisons require a common ‘yardstick’ which social realist conceptualisations offers (Skinningsrud 2006). Thus, Archer’s domain theories on educational systems and reflexivity are included. These theories conceptualise commonalities as well as uniqueness in national educational systems and on (reflexive) agency. The theoretical concepts that are used in my investigations (art. I, in particular) are developed in connection with comparative studies of four European countries (England, France, Russia and Denmark) (Archer [1979] 2013, 1984, 1995). They capture structures and processes that are found in all four countries, while also allowing for differences between them. The same logic of comparison applies for the reflexive modes, but between individuals.

4.1.1 Morphogenetic cycles

The process of social morphogenesis or reproduction is illustrated in Archer’s morphogenetic cycles which applies to the examination of structure, culture, and agency. Morphogenesis is contrasted to morphostasis, which refers to those processes in a complex system that tend to preserve the form, structure or state unchanged (Archer 1995, 75). The structural outcome at the end of a cycle is the result of interactions between agents who are differently positioned and promote different ideas and interests, struggling either to change or preserve the status quo. This approach entails that social effects are produced through agency, although often resulting in unintended side effects. Each morphogenetic cycle consists of three phases: T1,

T2-T3 and T4 (see Archer 1995, 193f). Archer stresses that each phase involves structures with properties that create situational logics that provide guidance for action (see figure 1). T1 involves the structural contexts/conditions within which individuals are positioned and act: for example, a centralized education system. Here, pre-existing properties of the system influence and divide the population ‘into social groups working for the maintenance versus the change of a given property, because the property itself (e.g., distribution of wealth, enfranchisement, educational control) distributes different objective vested interests to them at T2’ (Archer 1984, 8f). Social interaction, when initiated at T2, is the response to these properties that already exist. T2-T3 designates the social *interactions* of individuals and groups (agency) that have interests and ideas and who strive for outcomes that favour those interests and ideas. Individuals’ and groups’ interests and ideas regarding retaining or changing the established properties become activated, as do their objective capacities for doing so. T4 designates a new formation of structural/cultural (*elaboration*) of the T1 with new social conditioning (morphogenesis), or that the original structure at T1 has been reproduced (morphostasis)—both outcomes may constrain or facilitate the individuals and groups involved (Archer 1984, 8–9; 1995, 192ff).

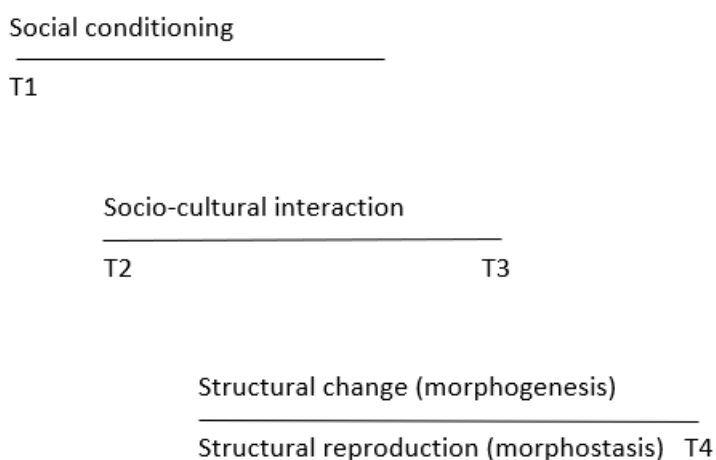


Figure 1: The Basic Morphogenetic/Static Cycle with its Three Phases. Source: Archer (1995, 157).

4.1.2 The concept of the educational system

Archer defines a state educational system as a structural entity, ‘a nationwide and differentiated collection of institutions devoted to formal education, whose overall control and supervision is at least partly governmental, and whose component parts and processes are related to one another.’ (Archer [1979] 2013, 54). In general, the education system, its structures and processes, are viewed as a result of ‘what people have wanted of it and have

been able to do it.’ (Archer [1979] 2013, 2). Moreover, educational policies throughout history have been based in part on the educational needs and interests of various social groups (i.e., education towards solving specific tasks, such as knowledge of navigation as world trade evolved) and educational ideas (such as the folk high school movement in Denmark inspired by Grundtvig’s ideas).

As reported in ‘Social Origins of Educational Systems’, first published in 1979, Archer decided to study four countries—England, France, Denmark and Russia—for comparison. These countries were selected to bring forward historical developments of education and socio-cultural contexts and increase the validity of the general theory she developed in order to enable generalisation beyond these four cases (Archer [1979] 2013, 42). Moreover, Archer’s theory is historically specific and is limited to education in countries that have not gone through territorial redistribution or colonisation (Skinningsrud 2012). As such, her educational theory applies to societies that are institutionally differentiated, and where educational development has taken place as an endogenous process. The first limitation excludes monolithic empires and the second excludes colonised countries from the theory (Skinningsrud 2012, 17–18).

State educational systems in European countries emerged at specific times in history as social structures (Archer [1979] 2013; Skinningsrud 2012). Moreover, they emerged in different socioeconomic and political contexts, characterised by an uneven development in historical time and geographical contexts. A state education system emerged in France in the 1840s, in Russia in the 1880s, and in Denmark and England between the 18th and the 19th centuries (Archer 1984). Archer found that the French and Russian systems were centralized, emerging from educational interaction where a policy of restriction through state legislation and control predominated. The English and the Danish systems were, however, decentralized, emerging from educational interaction characterised by competition among several interest groups that controlled different educational networks.

Archer is a methodological collectivist in the sense that the social form of the educational system pre-dates agency i.e., the structure of the educational system pre-dates the actions of the agents (Archer [1979] 2013). Apart from that, she stresses analytical dualism, that capture change both on structure and agency and for understanding their interplay. Culture can’t be left out, because some ideas will be compatible (complementarities) or incompatible (contradictions) with the agents concerns.

4.2 Educational systems: centralized and decentralized, structures and processes

Educational systems contain internal structures and processes that promote unification, systematization, differentiation, and specialization along a continuum. Centralized and decentralized educational systems differ regarding the relative dominance of mechanisms of change. For my analysis, this often entailed a qualitative and quantitative aspect, as all education systems are characterised by unification, systematization, differentiation, and specialization, to some extent. Nevertheless, the mechanisms are weighted differently in the two systems, as shown below in Figure 2 where unification and systematization are the predominant mechanisms, with less (weakened) differentiation and specialization. This is symbolised³⁹ with addition and subtraction signs. This figure can also be used to illustrate decentralized systems, but in the opposite order. Educational systems are not absolute; rather, they are understood along a continuum. In centralized systems, unification and systematization predominate, while in decentralized systems, differentiation and specialization predominate. These structures and processes are further elaborated below.

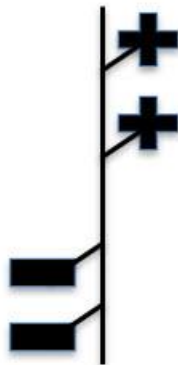


Figure 2: Illustration of the Internal Systemic Structure of Centralized Educational Systems. Source: Author's illustration.

4.2.1 Unification

Unification refers to national uniformity, or having the same standards throughout the nation, i.e., the same admission criteria, content and final competences. The school structure is

³⁹ Additional examples for addressing the differences between one or the other can be found in programming and biology (among others). In programming, binary numbers are an example of either one or the other number being operative for certain activities. In biology, one distinguishes between men and women, and different gender systems possess different characteristics, such as the different degrees of testosterone and oestrogen, despite both being classified as a homo sapiens. This observation is necessarily followed by deep properties, such as the presence of X and Y chromosomes.

decided at the political centre and the educational programmes are the same everywhere. The structure of the education system is a result of political decisions, which are based on compromises between different political groups and parties that result in unintended effects of the political decisions. Unification has two dimensions: *intensive unification* and *extensive unification*. Intensive unification entails that the state control of the education system is strong and effective. State educational policy is effectively implemented regionally and locally. This is ensured through management hierarchies and clear lines of command—and nowadays, increasingly also through goal-and-result management or input and output measurements. Extensive unification means that the policy is implemented on a national scale. There are few or no regions, localities or school types that are free from the state policy (Archer [1979] 2013, 174–176). There are degrees of both intensive and extensive unification. In contrast to decentralized systems, in centralized systems there is a continuous process towards greater unification.

4.2.2 Systematization

Systematization refers to the interconnected relations that contribute to the coherence of the system. It represents a transition from several essentially different and incoherent structures or independent networks to a unified and consolidated system. Systematization involves strengthening previously existing relationships between the parts and developing relationships that have not previously been in relation to each other (Archer 1984, 74–75). Archer notes three arrangements that contribute to the systematization process: a) the use of national exams or exams that provide nationally recognised competences; b) teacher recruitment, education and certification that are valid for the various educational levels; and c) the development of a variety of roles, services, establishments and trained personnel required to complete the linking and coordination process throughout the system. Strong systematization reduces bottlenecks in the system. In Norway, the systematization process started during the second half of the 19th century (Thuen 2010; Skinningsrud 2012).

4.2.3 Differentiation

Differentiation refers to education as comprising ‘a specialized educational collectivity, occupying a distinctive educational role structure, transmitting definitions of instruction [learning content] which are not co-terminous with the knowledge or beliefs of any single social institution’ (Archer 1984, 75). Differentiation is obtained by education being integrated with a multitude of social institutions (also called multi-integration), which entails that the

system must cater to the whole range of competences, knowledge and training that society needs. Nevertheless, all singular demands cannot be fully met, as no school learning will successfully meet all of society's needs: interests are always negotiated, and compromises must be reached and temporally agreed upon (Archer 1984, 176). As different groups and agents have different interests that they want to promote, 'the conjunction of these different interests means that each [interest group] acts as a watchdog to prevent the re-establishment of exclusive links between education and another party' (Archer 1984, 76).

Institutional differentiation of the educational system may be strong or weak. *Strong differentiation* characterises decentralized education systems that have relatively high autonomy. Internal processes are less controlled by central authorities, and there is no single national curriculum. This means that the actors, i.e., the teacher profession, can themselves decide on the internal work operations within the system. Correspondingly, *weak differentiation* characterises centralized education systems, which have less autonomy compared to decentralized systems. The educational activities are controlled to a greater extent by political authorities in centralized systems than decentralized systems (Archer [1979] 2013, 256). This entails that professionals in centralized systems receive more orders from the political centre, possess less authority and take fewer initiatives. Moreover, the teacher associations tend to be reduced to purely interest groups, with no educational policy goals, educational visions or ideals, instead focusing on safeguarding the profession's working conditions. The teacher profession is thus forced to bring their demands to the political negotiation table just like any other interest group. Additionally, and mentioned by Archer in her original comparative education work, 'there may be sections of the elite which prefer a low level of [or weak] differentiation, with an intermingling of political and educational roles and activities, [where] trained teachers represent loyal cadres' (Archer 1984, 76).

Weak differentiation entails little autonomy for the teaching profession. It means that the schools cannot conduct transactions with groups outside the education system concerning changes in educational provisions, or the establishment of new ones. They are subject to strong political control, which prevents them from developing specific changes that may be desired by local external groups. The teacher profession cannot change curricula, exam forms or admission criteria, as these are centrally determined. In centralized systems, which have weak differentiation, the professions can usually only initiate internal changes that are compatible with state policy and the existing organisation of the system. In systems with

weak differentiation (i.e., centralized systems), there is: a) little control by teacher professionals over ‘internal operations’ i.e., interference from the political authorities regarding how the teachers perform their tasks; and b) little involvement by the teaching profession in educational policy processes (‘insertion’) (Skinningsrud 2019).

4.2.4 Specialization

Specialization entails that ‘intake, processes and outputs [...] meet demands whose diversity is incompatible with unitary procedures.’ (Archer [1979] 2013, 181). Specialization can result in new school programmes, new content in existing programmes, the delineation of new roles, new forms of recruitment and training, increased complexity in intake policies, the development of alternative educational courses, combinations of previously separated education programmes, various professional specializations, exam forms and types of qualifications, new special equipment and teaching material (Archer [1979] 2013, 182). Specialization and unification (extensive unification) are opposite processes in an educational system, as an increase in unification will weaken specializations and vice versa. In centralized education systems, specialized provisions will undergo a process that reduces the degree of specialization and creates greater uniformity.

In centralized education systems (see Figure 3) where the negotiation process is brought upwards to the central government, the biggest problem is to integrate the wide variety of educational demands. Uniformity and standardisation will facilitate the effective implementation of state educational policy but not ensure that various groups will receive the specific educational services they want. There is dissatisfaction with the system as a whole among groups in society whose educational interests are not satisfied (Archer [1979] 2013, 254–255). In centralized education systems, as well as decentralized education systems, the political elite are influenced by what professional groups (such as the teacher profession) and external interest groups consider important. The results of the interaction process and negotiation often end in compromises that favour the most dominant and resourceful groups, including the political elite. Moreover, in centralized educational systems, there is a strong hierarchical structure of governance. The administrative levels are tightly linked, with the education minister at the top. This gives little freedom and room for individual choice of action at the lower levels. In fact, Archer states that ‘if the aims of the political elite are to be satisfied, unification must be intense and extensive.’ (Archer [1979] 2013, 200).

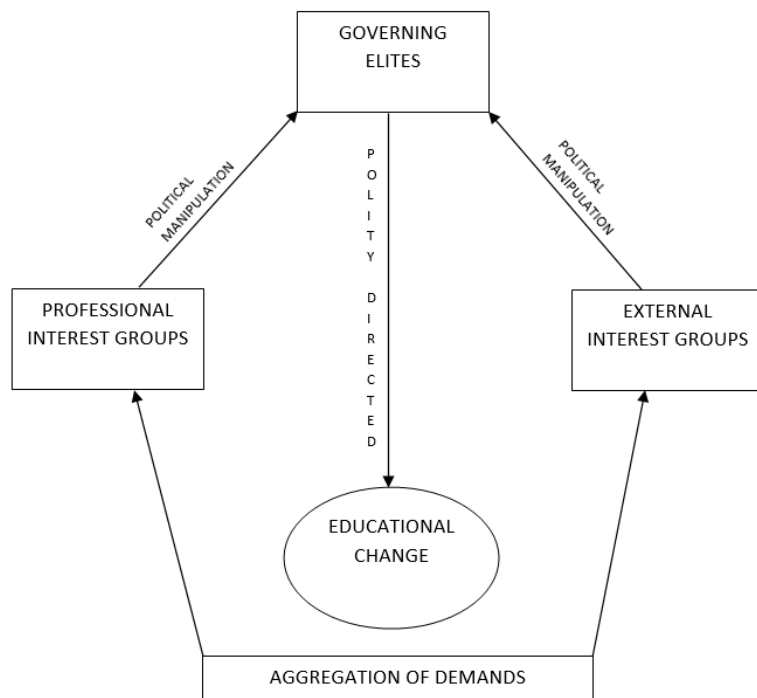


Figure 3: The Structural Conditioning of Educational Interaction in the Centralized System. Source: Archer ([1979] 2013, 266, adjusted illustration by the author).

In decentralized systems (see Figure 4), government control is continuously challenged. In such systems, changes are initiated in many places when specific changes are desired, and local agreements are made with schools regarding the implementation of specialized educational programmes. Decentralized systems will entail locally initiated changes, which can create problems for the implementation of state education policy. A direct interaction between the professional groups and external interest groups, where state education policy is neglected or challenged, threatens the uniformity and systematization of the system. This is crucial, as in decentralized systems the teaching profession, through established structures, may have the power to determine pathways for education themselves by transacting with external groups. In decentralized systems, different groups struggle to promote their educational interests without using formal political channels, which can lead to a greater diversity of provisions but simultaneously create instability and disorder in the system.

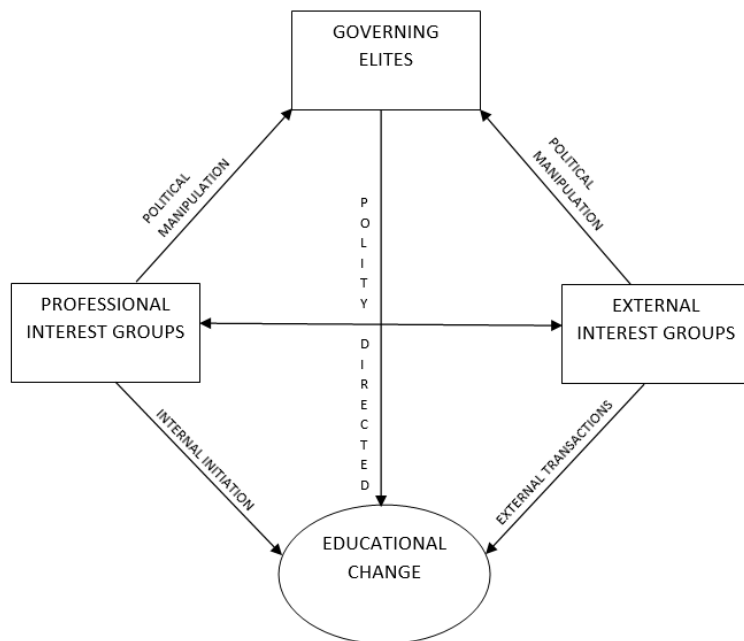


Figure 4: The Structural Conditioning of Educational Interaction in the Decentralized System. Source: (Archer [1979] 2013, 267, adjusted illustration by the author).

Archer states that centralized education systems tend to persist, regardless of government and constitutional changes. This was the case for the education system in France, where shifts in the governing political elite did not alter the system. No governing elite voluntarily relinquishes central governance, which Archer believes is a substantial finding in comparative educational research (Archer 1984, 200). Moreover, she claims that in centralized systems, to ensure that the educational policy of the political elite is implemented, the administrative framework is designed to promote strong unification—intensive and extensive—in the system (Archer [1979] 2013, 200).

4.3 The concept of agency: personal and social identity

Archer's conception of agency is elaborated in several of her works, for example, in 'Realist social theory: the morphogenetic approach' (1995), 'Being Human' (2000), and 'Structure, Agency and the Internal Conversation' (2003).

4.3.1 The human 'self' and 'agency' in different theoretical traditions

In the book 'Being Human', Archer presents her conceptualisations of the human 'self' and 'agency' in comparison with how these notions are treated in other schools of thought, in particular: i) the tradition of the Enlightenment, which adheres to the notion of 'rational Man'; ii) the tradition of social constructivism, which considers the self and agency as derived

entirely from the social; and iii) post-modernism, which tends to dissolve the notions of self and agency altogether. Pointing out the weaknesses of these other positions, she argues that the notion of self is a prime condition for society to exist at all, and that the agency of individuals as well as incorporated groups have causal power in effecting social change.

Archer argues that post-modernist theories – mentioning Michel Foucault as a leading representative of this school of thought – tend to ‘dissolve [the self] into discursive structures.’ (Archer 2000, 3). Against this Archer argues that ‘our sense of selfhood is independent of language’ since it emerges from our practical activity in the world (Archer 2000, 3). She points out that even those who consider discursive structures as more fundamental than the human ‘self’ admit that agents are not entirely passive. Foucault, for example, suggests that there is a human capacity for resistance, and Richard Rorty insisted on the human ability for self-enrichment. But as Archer points out, ‘neither resistance nor enrichment could be coherent without a human self who engaged in them.’ (Archer 2000, 3).

Against the theoretical *tradition stemming from the Enlightenment*, which tends to grant human beings ‘one property alone, that of rationality’, Archer argues that it overlooks how society shapes human beings, because human rationality is considered as a pre-given (Archer 2000, 3–4). Furthermore, when the model of ‘rational Man’ was reinvented as ‘*homo economicus*’, it could not incorporate human normativity and emotionality (Archer 2000, 4). ‘[T]he lone, atomistic, and opportunistic bargain-hunter stood as impoverished model of “man”’, ‘who knows the price of everything and the value of nothing’ (Archer 2000, 4). This model of human beings, based on rationality as their only property, could not cope with the ‘human capacity to transcend instrumental rationality and to have “ultimate concerns”’ (Archer 2000, 4). ‘Ultimate concerns’ is by Archer defined as commitments that are ‘constitutive of who we are’, expressing our personal identities. They are not a means to something else (Archer 2000, 4).

Against the theoretical tradition of *social constructivism*, Archer argues that one of its major theoretical difficulties is its neglect of the fact that humans are embodied. Social constructivism claims that ‘human properties and powers, beyond our biological constitution’ are given by society, i.e., Archer names this conception ‘Society’s Being’, i.e., humans are considered as entirely a product of the social (Archer 2000, 4). ‘We are nothing beyond what society makes us, and it makes us what we are through our joining society’s conversation’ (Archer 2000, 4). Thus, social constructivism sees the distinctively human powers of

‘selfhood, reflexivity, thought, memory and emotionality’ as derived from ‘society’s discourse’ (Archer 2000, 4). Against this Archer argues that embodied practice is the non-linguistic source of the enduring sense of self. ‘Bodies have properties and powers of their own and are active in their environment, which is much broader than “society’s conversation”’ (Archer 2000, 4). In Archer’s view the social constructivist conception merges ‘the concept of self’ with ‘the sense of self’ (Archer 2000, 4). She points out that seeing human beings ‘as purely cultural artifacts is to neglect the vital significance of our embodied practice in the world.’ (Archer 2000, 4).

4.3.2 Agency, personal identity, and ultimate concerns

Archer’s account of how agency evolves in the course of each individual’s development from childhood to maturity starts with the emergence of a continuous *sense of self*, which is acquired early in life, and culminates in ‘the active acquisition of a *personal identity* at maturity.’ (Archer 2000, 9). Every individual’s personal identity is unique, but it is still constituted by what happens to that person during his or her life. Humans are born into circumstances that are not of their own choosing and do not have full control of all circumstances during their life course. Still, their personal identity emerges in an active interplay with their environment, and the development of their personal identities is based on their emergent ability to *reflect* on the reality in which they are situated and engage with.

Archer distinguishes between three orders of reality that humans must relate to. These are the natural order, the practical order, and the social order. Different kinds of concerns, related to our emotional reactions in encountering these different orders, are physical well-being, related to the natural order; performative achievement in the practical order; and self-worth in the social order. Emotions are conceived, by Archer, as commentaries on our experienced welfare in these three orders of reality, thus on our physical wellbeing, our performative achievements, and our self-worth. Humans’ unique personal identities consist of how they prioritise between these basic concerns, among which none can be completely ignored. ‘Which precise balance we strike, and what exactly features as our ultimate concerns is what gives us our strict identity as particular persons – our personal identity.’ (Archer 2000, 10).

4.3.3 Reflexivity – the internal conversation

Archer defines reflexivity as ‘the regular exercise of mental ability, shared by all normal people, to consider themselves in relation to their (social) contexts and vice versa.’ (Archer

2007, 4). In Archer's conception, our personal identity, or concrete singularity, is generated by our inner conversations:

The internal dialogue entails disengaging our ultimate concerns from our subordinate ones and then involves elaborating the constellation of commitments with which each one of us feels we can live. The 'inner conversation' is about exploring the terms of a liveable degree of solidarity for the self in its commitments, and the unique *modus vivendi* to emerge is what defines the uniqueness of personal identity. Whereas self-identity, the possession of a continuous sense of self, [is] held to be universal to human beings, personal identity is an achievement. It comes only at maturity, but it is not attained by all: it can be lost, yet re-established. (Archer 2000, 10).

In short, we are who we are because of what we care about: in delineating our ultimate concerns and accommodating our subordinate ones, we also define ourselves. We give a shape to our lives, which constitutes our internal personal integrity, and this pattern is recognisable by others as our concrete singularity. (Archer 2000, 10).

Archer considers the 'rich inner life of reflection upon reality' as the generative mechanism of the person (Archer 2000, 10). Neither rational Man/Homo economicus nor the human being as a cultural artifact created entirely by society play an 'active role in who they are' [...] they are not 'allowed to play a major part in the making of their own lives.' (Archer 2000, 10).

While personal identity is derived from all the orders of reality, the natural, the practical and the social, our social identities are defined in only social terms, although the two are intertwined. 'Strict social identity is achieved by assuming a role(s) and personifying it, by investing oneself in it and executing it in a singular manner.' (Archer 2000, 11). The notion of a personal identity enables us to explain the individually unique performance of a role. Social identity is a subset of a much broader personal identity. Since personal identities are based in humans' relationship with the three orders of reality, the natural, the practical and the social, 'it is ultimately *the person* who determines where the self-worth, that he or she derives from the social roles, stands in relation to their other commitments in the world as a whole. It is also the person who arbitrates upon the relative importance of their multiple social roles and between their greedy demands.' (Archer 2000, 12). The *person* strikes the balance between various social and personal concerns. This requires prioritising, which will determine how

much effort and how much of themselves as persons they will invest in their various social identities.

In the person's internal conversation, reviews of commitments are made, and what kind of social actors one has become, are continually carried out. Prime concerns and commitments may be re-endorsed or rejected at any time, as may social identities, based on whether they are still in agreement with the priorities of one's personal identity:

'Personal identity is an accomplishment, but it has to be reconstituted from day to day by a re-affirmation and renewal of our concerns. Such active continuity makes us recognisable to others in our concrete singularity and consistent as social Actors through the consistency of our personified conduct in our social positions.' (Archer 2000, 12).

'Society enters into us, but we can reflect upon it, just as we can reflect upon nature and upon practice.' (Archer 2000, 13).

'Making our way through the world'⁴⁰ necessitates discernment, deliberation, and dedication (DDD)⁴¹, which can change on the way, as we occupy new social roles and conditionings.

4.3.4 Modes of reflexivity

Reflexive deliberations constitute the mediatory process between 'structure and agency' and 'culture and agency', which becomes the interplay of causal powers between social objects (Archer 2003, 130). All humans practice internal conversations, i.e., reflexivity, but their way of conducting internal conversations vary. Archer developed her concepts about modes of reflexivity through in-depth interviews among a sample of 20 participants. She identified three modes of reflexivity: communicative reflexivity, autonomous reflexivity, and meta-reflexivity (Archer 2003)⁴². In her original sample, 15 participants were clustered into the three modes. The remaining participants were, for various reasons, unable to sustain internal conversations about themselves and society (Archer 2003, 298); their reflexivity was

⁴⁰ Book title of Archer (2007).

⁴¹ If not DDD, most likely then fractured reflexive.

⁴² Archer acknowledges that there may exist other forms of reflexivity, which may not have been manifested in her sample due to the lack of relevant social conditions.

fractured. What distinguished the participants whose reflexivity was fractured was that they did not have clear projects that they were able to pursue. Fractured reflexives might have concerns, but they are too disoriented to effectively pursue those concerns.

4.3.4.1 Communicative reflexives

Communicative reflexives are characterised by internal conversations that invite competition and confirmation from others before they lead to courses of action. A communicative reflexive is oriented towards reaching consensus. These are people who initiate internal dialogues in the privacy of their own minds, but that is not where they complete them. Instead, their pattern is one of ‘thought and talk’. When they raise an issue intrapersonally, they seek to resolve it interpersonally. For this reflexive type, friends and family are their ultimate concern. They are conformists and have difficulties taking a stance in relation to a cause (Archer 2003, 342).

4.3.4.2 Autonomous reflexives

For autonomous reflexives, work and performative achievement are their ultimate concern (Archer 2003, 265). Autonomous reflexives are characterised by self-contained internal conversations that lead directly to action. This reflexive mode is often strategically oriented to performances, outcome, quality, accountability and ‘best practices’ (Archer 2007, 321–322). This type is an individualist who takes responsibility for his or her own actions rather than being the victim of circumstances. Autonomous reflexives attempt to be strategists in their own lives and the social conditions in which they are situated—that is, to be active agents who make things happen rather than passive agents to whom things happen. However, strategic action is conducted in an open system and is always at the mercy of unforeseen and unforeseeable contingencies (Archer 2003, 251). Archer has elaborated about autonomous reflexivity in (high) modernity as the dominant mode of supporting globalised ideas (capitalistic in nature) and as possibly dominant amongst people that retain social positions⁴³ (Archer 2007, 2012).

4.3.4.3 Meta-reflexives

Meta-reflexives are characterised as critical subjects towards the prospects of effective action in society and towards their own internal conversations. These reflexives are often oriented to

⁴³ Archer has predicted that meta-reflexivity in high modernity is especially connected to young educated people, and possibly being dominant in society (Archer 2012, i).

values and ethical deliberations. Meta-reflexives are idealists and experience a constant tension between structure and culture. The meta-reflexive has difficulties in completing the sequence of concerns → projects → practices, to his/her satisfaction (Archer 2003, 258–259). Moreover, they are social critics of a society that does not fulfil their ideal, but also self-critical of themselves and the lives they live (Archer 2003, 258). They are subversive and can also be volatile.

4.3.4.4 Fractured reflexives

Fractured reflexives are characterised as distressed and disorientated, as their internal conversations lead to no purposeful courses of action. Fractured reflexives are non-oriented or disoriented. They are passive agents (Archer 2003, 300). ‘Passive agents are the opposite of those taking a social “stance”; they are people to whom things happen rather than people who exercise some governance over their lives by making things happen’ (Archer 2003, 343). These individuals are unable to pursue individual projects.

4.4 The causal agent with ideas and materials

Two principles that are crucial in applying the concepts of social structure, culture, and agency in the study of PISA effects are that the “causal power of social forms is mediated through social agency” (Bhaskar 1989, 25-26, quoted in Archer 1995, 195), which means that ‘agents are the only efficient causes in social life’ (Archer 1995, 195). This entails that any ‘PISA effect’ generated by social structures requires mediation by agency. The other principle concerns the study of PISA in the cultural sphere, which starts with ‘ideas which at any given time have holders’ (Archer [1988] 1996, xxi). These two principles assert that agency is crucial in the study of both structural and cultural PISA effects (for example on social forms). As part of culture, PISA results and ideas could be studied as items in the cultural system and how they are used by individual and collective agency in socio-cultural interaction. Archer defines the cultural system as specifying the logical relationship between the components of culture residing in contradictions and complementarities (Archer [1988] 1996, xviii). Socio-cultural interaction refers to how ideas and propositions are used in interactions, where people try to influence each other’s thoughts. Interpersonal influence may include persuasions, force, censorship, argumentation, legitimation and mystification (Archer [1988] 1996, xviii-xix; 1995, 179). Sometimes ideas are manifested as (physical) materials (includes structural entities) such as institutions (social positions and power-relations), jurisdictions, money, booklets, which have possible effects on the structural relation itself.

Archer (1995, 185) distinguishes between two types of agents: primary and corporate. Primary agents are agents that are gathered as a group or collective that have not stated their demands publicly and are unorganised regarding their pursuit. Primary agents therefore reproduce established macro structures and often struggle to live within them. Nevertheless, they can become a member of a group, and so be organised to drive change through different means. Corporate agents are organised into interest groups that have publicly articulated common goals based on their common interests, e.g., teachers' professional associations. They have interests in re-modelling the system. However, the corporate agents' capacities to successfully achieve transformation—the morphogenesis of structures—are dependent on the conditional influence of other systemic structures, e.g., the corporate interests of other groups.

4.5 Research questions: The Norwegian case

Research questions are meant to narrow the gap between what we know and what we would like to know more about; however, time and resource limitations have an impact on the extensiveness of the research. The overarching research questions were previously justified and posed in chapter 1 where a specific conceptual framework was missing in the PISA effects field with another way of understanding causality. This prompted a response to Pons featuring many of his requests where the Norwegian case was of interest. The subsidiary research questions are justified by prior Norwegian PISA studies that lacked conceptualisations of the educational system and a model of structural change. A conceptualisation of agency was also missing. More specific, their emergent properties (generative mechanisms) and their interplay were absent. Thus, this research will use the morphogenetic model for discussing structural change of the educational system before and after the introduction of PISA and reflexive modes to capture the dominant reflexive modes when dealing with PISA amongst a mathematics teacher and three school leaders.

In general the overarching research questions and subsidiary research questions are supported by requests for research of PISA effects on national educational policy (Sjøberg 2014a, c; Pons 2017; Stray and Wood 2020), and schools (Sjøberg 2014a; Pons 2017; Hopfenbeck et al. 2018; Stray and Wood 2020; Hossain 2023). Additionally with conceptualisations that capture normative, ethical, and collective concerns (Biesta 2010; Biesta, Priestley, and Robinson 2015) and that includes Pons's model of 'reception, uses and effects' in socio-cultural interaction. All these requirements are fulfilled with Archer's conceptual framework prompting a scientific version of systematic and cumulative research.

Overarching research questions:

- i. How does Archer's theoretical approach enable a reconceptualisation of alleged 'PISA-effects' on the Norwegian educational system?
- ii. How can Archer's concept of 'reflexivity' and 'reflexive modes' increase our understanding of how school personnel (teachers and school leaders) react to the idea of using PISA-test results in their own educational setting?

Subsidiary research questions for article I concerning the Norwegian educational system:

- After the introduction of PISA, the following public debate and the implementation of reforms justified by PISA, was there a change in which kinds of structures and processes that predominated in the [Norwegian] system?
 1. What kinds of structures and processes predominated in the Norwegian educational system in the decade before the introduction of PISA?
 2. Who were the protagonists in the public debate after the release of the first PISA results, and what were their concerns and projects?
 3. Which systemic structures and processes were strengthened and which were weakened by the reforms legitimized by PISA results?

These questions concern whether the Norwegian system is centralized or decentralized, in Archer's definition of these terms, and whether educational development took 'a new turn', i.e., whether fundamental characteristics of the system were changed after the implementation of reforms justified by PISA, for example, the introduction of NQAS and the Knowledge Promotion Reform.

Subsidiary research questions for article II concerning a mathematics teacher's reflexivity and reflexives modes:

1. Which modes of reflexivity are activated in a mathematics teacher by the PISA test?
Which mode(s) of reflexivity predominate(s)?

Subsidiary research questions for article III concerning school leaders' reflexivity and reflexive modes:

1. Which reflexive modes are activated in the three school leaders when they engage in internal conversations about PISA?
2. Which is the dominant reflexive mode of each school leader (regarding PISA)?



Figure 5: The Hierarchy of the Overarching and Subsidiary Research Questions. Source: Author's illustration.

4.6 Summary

In this chapter I have argued and presented a social realist conceptual framework for analysing individual countries' educational systems and modes of reflexivity among incumbents of the system. In other words, these conceptualisations are suggestively a common social 'yardstick' that enables comparisons of the PISA phenomenon in different contexts (countries). The common 'yardstick' advocated in this chapter implies the use of real definitions of structure and agency. These are employed in the study of social diachronic change (morphogenesis) through the analysis of successive morphogenetic cycles. But to explain change within a system implies defining the system. In critical realism and social realist theory, structure pre-dates and conditions agential action. The structures of educational systems have distinct properties (mechanisms). These are unification, systematization, differentiation, and specialization. Unification and systematization predominate in centralized systems, while differentiation and specialization are more pronounced in decentralized systems.

In addition, structural changes or reproduction in the educational system are dependent on agency. Reflexivity therefore becomes vital, since it filters the environment through reflexive

processes and deliberations that determine the actions we take. The agential, i.e., agency has its own structures and properties (mechanisms) related to the internal structure of it, i.e., reflexivity (the internal conversation) and the interplay between the four reflexive modes, which determine the development and social trajectories of individuals. Human change, or agential change, is connected to Archer's three orders of reality—the natural order, affecting our concern for physical well-being; the practical order, affecting our concern about performance; and the social order, affecting our concern for self-worth. These arenas trigger emotional commentaries and concerns that humans deliberate upon, intermingling with the 'sense of self', which affects our internal conversation and the (re)making (of) our ultimate concerns.

When studying PISA effects on concepts such as structure, culture and agency, it is vital to remember that the “causal power of social forms is mediated through social agency” (Bhaskar 1989, 25-26, quoted in Archer 1995, 195). This entails that ‘agents are the only efficient causes in social life’ (Archer 1995, 195). Applying Archer's ontological distinction between social structure, culture, and agency, due to their distinctive emergent properties and causal powers, entails studying PISA as a set of ideas (that also materialise into different material) and how they are used by individual and collective agency. This entails if PISA is compatible or incompatible with stakeholders own concerns.

Based on Archer's conceptual framework and understanding, I have at the end of this chapter presented the subsequent research questions, which represent a further development of the former research questions posed in the beginning of this thesis. All levels of research questions are informed by this conceptual framework which is underpinned by critical realism.

5. A case study framework

5.1 Case study research

Case studies are suitable for examining ‘a contemporary phenomenon within some real-life context’, where the boundaries between the phenomenon and context might not be clearly evident (Yin 1994, 1). Case studies are a suitable choice when one wishes to know more about the relationship between a phenomenon, the context, and the people, such as the Norwegian case of PISA effects (on (pre-dated) structures) in this project. Case studies often include how or why questions and that they often have more variables than data points (Yin 1994, 13), which suits with the intention and scope of this study. Additionally, former case studies have not fully succeeded in distinguishing key variables and factors for how and why PISA effects occur or not (Pons 2017), which also justifies this contribution with fundamental units (factors and variables). A case study framework invites the use of multiple methodologies and methods, which also is enabled in Archer’s conceptual framework. Case studies can either be grounded in constructivist epistemology (i.e., what we know is socially constructed) (Merriam 1988; Stake 1995) or positivistic epistemology (i.e., what we know is real and exists) (Yin 1994). This case study research was grounded in the latter due to a belief that the natural world is real and not solely socially constructed, and that humans are not solely a social construct.

5.2 Design and context

The research design in this case study followed the logic that the collected data must be connected to the ‘initial questions of a study’ (Yin 1994, 18), and the selected units must be situated within a theoretical explanatory framework to guide the analysis. Here, the focus concerned ‘PISA effects’ on systemic structures, with three units (factor and variables) represented by five cases discussed and explained within a theoretical context. Each case is represented by one unit, and each unit is represented by its own case study design. In article I, presenting case 1, the single case is seen as a single holistic critical case with which to identify ‘PISA effects’ on the internal structure of the Norwegian education system. Article II, with case 2, is devoted as a single holistic critical case on one unit. Article III involves cases 3, 4 and 5, which are seen as multiple holistic critical cases corresponding to the same units. Both article II and III have the rationale of uncovering, informing, and change the direction of the PISA research by producing research from the micro-level. All cases are devoted as ‘holistic’ because they are based on macro contexts that are shared and the data is analysed with universal categories.

Thus, the thesis can be understood as an overall multiple holistic case study, as the project comprises several single cases within a broader context. It is imperative to distinguish between the internal and external context to prevent a flat ontology and conduct a more coherent analysis. A counterfactual analytical question that enables the external context is, how must reality be for that to happen? This question solicits explanations beyond the internal context in which the cases are represented: these largely follow an analytical rationale between the internal and external context of the cases, corresponding to an extended social reality outside the cases. This can be viewed in the distinction between abstract and concrete research, where the former identifies necessary relationships, terms, and properties in the open reality by abstraction; and the latter examines actual events and complex phenomena which have multiple causes, but where some are more dominant determinants than others. Therefore, identifying mechanisms in relation to different levels of social reality constitute key elements in the establishment of explanations (Danermark et al. 2003, 123–132; Blom and Morén 2011).

Article I describe characteristics of the Norwegian education system before and after the introduction of the PISA test to examine possible effects of PISA on the characterising structure(s) of the system. The case is empirically situated between 1990 and 2010, as the first PISA test was conducted in 2000. The aim was to reveal existing structures, interactions, and structural elaboration before and after the introduction of the PISA test to critically determine whether the test has had any possible effects.

Articles II and III present findings about the effects of PISA on school personnel who are situated in two lower secondary schools, especially focusing on mathematics as a teaching subject due to my background as a mathematics teacher and the fact that mathematics is one of the domains of the PISA test. Here, interviews were needed to understand and identify possible effects on school personnel and in relation to mathematics as subject. A mathematics teacher and three school leaders were selected as units. The choice of these numbers of respondents was aimed at narrowing the scope of the empirical data, identifying generative mechanisms, and presenting examples, not as a basis for statistical generalisations. Article II utilised a criteria-based approach to sampling but was also strategically and theoretically informed in accordance with the resources and availability of respondents. Together, article I (one country) and article II (one mathematics teacher) and III (three school leaders) offers vertical and horizontal comparative perspectives.

5.3 Evidence and analysis

The evidence should sufficiently answer the overall research questions in this thesis and the sub-questions presented in each article. Following Yin (1994), generic evidence that can be used in case studies are documents, archival records, interviews, observations, and physical artefacts, but are not limited to these sources. Case studies can be based on qualitative or quantitative methodologies.

Article I, presenting case 1 and any possible effects of PISA on the characteristics of the educational system, centres around a historical examination, prompting research questions that capture changes. The case study was based on document analysis and the research literature on 'PISA effects'. Overall, data from educational literature, green papers, white papers, a blog, and an autobiography were used to answer the research questions. In this case, a morphogenetic model/approach to analysing a limited morphogenetic cycle was used to examine whether there have been transformation(s) or reproduction(s) in macro structures following PISA's introduction. A *chronological time series analysis* was conducted, mainly because the analysis is foremost ordered and presented in time series of events (Yin 1994). A supportive thought on agency from *program-logic models analysis* were also included, as multiple causes create effect(s), where dependent (agency) and independent (e.g., PISA) mechanisms (variables) are vital contributors for creating (systemic) effects, especially a reminder for article II and III. Moreover, *explanation-building analysis* was utilised in all articles to explain events and phenomena (Yin 1994).

Chain-searching was used to identify relevant empirical material, largely because this approach facilitated effective and efficient searches targeting the sub-questions. All the documents⁴⁴ and extracted data were organised using the EndNote reference management program, facilitating later access for closer examination. Three principles guided the data collection: relevance, authenticity and credibility (Thagaard 2018, 119). The data stemmed from both primary and secondary analyses, where the former refers to the author's analysis of raw empirical material and the latter to already published data. All empirical data were

⁴⁴ The author acknowledges that the use of NVivo for maintaining an electronic overview and storage of the documents included for articles I and III could have been beneficial for having easier access to these documents; this would also increase the reliability if one wanted to do a quick screening and track the analyses and systematisations already undertaken. Nevertheless, as mentioned, the EndNote reference management software proved sufficient for tracking the included documents, even if it did not grant transparency regarding the analysis. Therefore, NVivo might be a good solution for further research, as it was for article II.

systematised according to theoretical categories (nodes) that guided the codes and extraction of relevant data.

Article II argues for the use of reflexivity and reflexive modes when analysing 'PISA effects' on the micro-level compared to several other conflated theories. Moreover, article II, presenting case 2, focuses on the effects of PISA on a mathematics teacher as a unit, which functions as an illustrative example but still enables the identification of generative mechanisms. A semi-structured interview was conducted, in line with the conceptualisations and theoretical categories used in article III. Pattern-matching was used implicitly as an analytical strategy; the prominent reflexive mode was anticipated from case 1 to case 2 with retrodictive thinking connected to the internal structure of the education system. It was also supported by the research from Skinningsrud (2019) after 2010 on the Norwegian educational system being centralized, based on the same conceptualisations as my article I. The empirical descriptions are mostly situated in identifying the effects of PISA on the mathematics teacher's reflexivity and work, but they also briefly touch upon a necessary explanation for understanding the actions and thoughts of the respondents, which is the specific education system, and this also has effects on his reflexivity.

Article III, presenting cases 3, 4 and 5, focuses on the effects that PISA has on three school leaders in the Norwegian education system. Semi-structured interviews from the master's thesis were re-analysed with more suitable theoretical categories, within a more adequate conceptual framework. The analysis was connected to the concept of reflexivity and reflexive modes as categories (Archer 2000, 2003, 2007, 2012). The research method used was a secondary analysis of qualitative data (Heaton 2008, 35), with a manual systematization of codes according to the reflexive mode categories. New research questions made it possible to re-analyse the data from the master's thesis, as the theme of the research was the same. The master's thesis consisted of respondents' knowledge about PISA and its purpose, and whether PISA had effects on their school and their practical work in teaching mathematics (including formative evaluations). The choice to re-analyse data from the Master thesis was grounded in the philosophical justification of the 'Holy Trinity' of critical realism. Here, the justification was informed by a realist ontology, a relative epistemology, in combination with the rational judgement that there might exist better and more comprehensive explanations for the phenomenon under study (i.e., better theories). The analytical strategy utilised pattern-matching between the three cases; findings were implicitly and explicitly anticipated from previous cases, and were implicitly connected to the findings and argumentation found in

article I and II (case 1 and 2), concerning the internal structure of the education system – especially processes of strong unification (a centralized system), which article III also addresses by its paragraph ‘Introduction: Educational testing for accountability and learning’⁴⁵ – maintaining implicitly that the system is centralized in Archer’s terms.

Generative mechanisms, identified in article I, II and III, are understood as pre-requisites, which raises questions about causal linkages between PISA and the educational system and PISA and agency (critical realism sees causality as a feature of the real world - the intransitive dimension). Furthermore, article I highlights a centralized educational system, with processes of unification and systematization between 1990 and 2010 being dominant. Both article II and article III also pinpoint mechanisms that are prevalent in a centralized system in Archer’s terms. I mention in article II confirmation of one national curriculum, and in article III there was indication of the lack of power to terminate PISA, which means that the teaching profession does not have that (locus of) authority in a centralized system.

5.4 Data collection and piloting

For article I, the data collection was mainly substantiated and collected during autumn 2018 and spring/autumn 2019. As mentioned earlier, data were collected via chain-searching (Rienecker et al. 2013, 119), in which one gathers relevant data, starting the process with an initial and essential reading of the secondary literature. As also noted above, the data were not collected and grouped in a coding program but systematised manually using the Microsoft Word program, with the logic of the morphogenetic cycles model of social morphogenesis/morphostasis (M/M). This meant that there was no rigorous process of analysing the data beyond systematising the data within a historical sequence of events.

For article II, the interview guide was tested with two pilots. Initially, I sought to recruit personal and professional acquaintances to participate in the pilot studies, but to no large avail. The pilot studies were carried out with two doctoral students: one was a former lower secondary teacher, and the other was a former upper secondary teacher. This enabled feedback on the interview guide, based on these two former teachers’ locus within the

⁴⁵ One curriculum (LK06), one Education Act, NQAS, all these mechanisms are indicators of strong unification in Archer’s terms, characterising a centralized system.

Norwegian education system. The pilot was aimed at removing errors and misunderstandings in relation to the theme and questions in the interview guide. Moreover, it also provided feedback on the chronology in the interview, and the researcher's interview style. A few questions were reformulated, some were removed, and some were added, as a natural consequence of pilot studies. The predefined Likert scale—the Internal Conversation Indicator (ICONI), from Margaret Archer's (2007) work on reflexivity and reflexive modes—was translated to Norwegian by two doctoral colleagues.

Before starting the collection of data, I checked the translated ICONI-version thoroughly for grammatical errors and I personally piloted it on myself. Before embarking the interview with the mathematics teachers, the research was notified and approved by Norwegian Centre for Research Data (NSD). The specific data for article II were collected in April 2020 from one mathematics teacher in a lower secondary school in an urban municipality in Norway. This single case study entailed method triangulation: a research approach that uses two methodological traditions to capture the same phenomenon (i.e., the reflexive mode) (Robson and McCartan 2016). In this approach, the qualitative method captures the reflexive mode that corresponds to the specific topic, while the quantitative one captures the number of each reflexive mode beyond a specific topic. More specifically, I conducted a semi-structured interview via Skype, which was recorded and later systematically coded in NVivo in accordance with the four reflexive modes (or 'nodes'). Lastly, I used the seven-point⁴⁶ ICONI Likert scale mentioned above. The ICONI scale has 13 questions, where groups of questions are connected to four categories of reflexive modes that are indicative of types of concerns and actions. The interview guide consisted of four themes and questions about background, environment, professionalism, and PISA, aiming to capture a better understanding of the PISA phenomenon contextualised in situ. The interview transcripts were translated into English. The mathematics teacher's age, years of work experience and number of students in that school were not included in the analysis to preserve the teacher's anonymity due to his enrolment for this study.

For article III, the interview guide was piloted with three teachers in 2014. The aim of piloting is to validate and judge the reliability of the interview questions and whether they correspond

⁴⁶ Extreme cases might occur, as they did for Archer (2007, 335). Nevertheless, all four reflexive modes, as well as their absence, were registered with a mean score somewhere between one and seven.

to the research questions and include themes that will enable operationalisations. The recruitment process of school leaders as participants in the pilot study proved challenging. As such, three teachers were ultimately selected, due to their accessibility and their experience with lower secondary schools. The piloting process promoted reflections and changes to the interview guide. Nevertheless, piloting the interview with school leaders may have mapped out other key aspects, and possibly better operationalised the scope of the study. It may also have led to a more structured and standardised interview, which in turn could have impacted the respondents' space of mobility.

As mentioned earlier, the master's thesis was centered around respondents' knowledge about PISA and its purpose, and whether PISA had effects on their school and their practical work in teaching mathematics (including formative evaluations). The interviews were conducted between March and April 2015 with three⁴⁷ school leaders in three different lower secondary schools in one urban Norwegian municipality. Each of the schools had between 400 and 600 students. The interviews were characterised as semi-structured and were digitally recorded. Participants were e-mailed for information about their age and educational background after the interview (since this was relevant to the thesis's discussion). The interview guide was made available to the school leaders beforehand, and they had the opportunity to consult with their staff (mathematics teachers). There were four topics in the interview guide: school leaders' understanding of PISA, PISA in formative work, PISA's importance for mathematics education and educational practices. All three interviews were conducted in Norwegian. For this doctoral thesis, the interviews were translated into English with new pseudonyms. The identification of reflexive modes was not the purpose of my original investigation. My master's thesis project, initiated in 2014/15, did first seek approval from the Norwegian Centre for Research Data (NSD) and was classified as 'non-notifiable'. Later on, I decided to collect some background information through e-mail. Due to the type of data being collected, the data collection strategy, the storage of data and how it was presented, it was not communicated further to the NSD. It followed the research guidelines at that time, and the General Data Protection Regulation (GDPR) was not yet implemented.

⁴⁷ Initially, there were four respondents, but one school leader withdrew his consent to making the empirical data available for future research.

5.5 Research ethics

NESH (2021) discusses internal relations in the research community as well as external relations to external parties. In this context, I focus on respect for the respondent, as I consider this to be the most crucial element. Normative guidelines for the researchers' relationship with respondents are to maintain their dignity and privacy, keep them informed, obtain their consent, treat data as confidential information, respect time limits for the reuse of data and store personal data safely (NESH 2021). In my research, it was important to maintain the respondents' anonymity, which influenced the scope of the contextual information I have provided about each case. Although one generally wishes to provide thick case descriptions, a compromise was to shift the thicker case description towards the Norwegian education system, as it is fundamental for understanding each case (and argumentation) and related events through a retroductive lens (i.e., considering the structures that condition actions).

NESH (2021) stipulates that researchers must protect personal integrity, safety and welfare. This entails respecting respondents' freedom and self-determination, as well as safeguarding them against harm. Thus, it is crucial to evaluate the research process before, during and after. This includes the research topic, one's relationship with the respondents, the methods used, the results, and the publication and dissemination of the results. This can be upheld by openness, transparency and evaluation, and also maintained through a reputable third party. In this research project, the NSD (Norwegian Centre of Research data) evaluated the project and interview guide with the mathematics teacher in article II. For article III, the NSD was not further formally involved after the project was classified as 'non-notifiable', despite emailing the school leaders for some background information later on. If the project was completely re-done today, the project would have become 'notifiable', especially due to the General Data Protection Regulation (GDPR). Despite this fact, the data was handled with care, and ethical considerations were mirrored in how the data were collected, stored and presented.

Clockwise, maintaining the personal integrity of the respondent entails transcribing the data in a manner that accurately conveys what the respondent has expressed. Finding the right words can be difficult when translating the data into a second language (in this case, English). One must critically evaluate whether the translations truly represent what was originally expressed. One example is when asking respondents about PISA and if you have understood them correctly. The expression of negative opinions could be perceived as a lack of loyalty to the institution (the school and the Norwegian educational system); moreover, criticism of PISA might jeopardise the job of the respondent, especially if he/she holds a position of leadership

in the school. The anonymity of the respondents in this research was therefore carefully attended to, and a minimum of personal identity markers have been presented to protect the respondent(s). Minimal information was obtained about the mathematics teacher (art. II).

Confidentiality is the process of handling data discretely and the data must be de-identified, i.e., the research material must be anonymised. Here, the juggling of different concerns might create a dispute between confidentiality and the process of being open (NESH 2021, 13). In this research I have valued respondents' confidentiality and anonymity as the prime concern. I have also used pseudonyms and been restrictive on information given for the readers. In this process, I have, for example, given the characters fictional names, and in article II, the respondent's age, education, work experience and school size were not included, neither were they 'manipulated' in the presentation of the case, to avoid any identity markers. This is quite a different strategy compared to the school leaders in article III, due to the enrolment process. The empirical material from the respondents in this project is hidden to maintain the safety and confidentiality of the respondents and preclude access from others apart from the researcher.

The concept of privacy in research entails deliberation on participants' autonomy, integrity, freedom and co-determination (NESH 2021). Therefore, information about the research was given to the respondents concerning the research topic, the purpose, who can access the information, the use of the results, and the consequences of participation. I gave information to the respondents through combined information and consent, amongst other methods, outlining their right to withdraw from the study without any reason or consequences. The combined information and consent letter also provided information about the data storage approach. The respondents received personal contact information for me, my supervisor and the university privacy officer (*personvernombud*). Three respondents in the master's thesis had given their consent for reuse of the data in further research.

Dalland (2011) and Heaton (2008) claim that data from qualitative studies are rarely used for secondary analyses. They argue that previously collected data can answer new research questions or be reanalysed for validating results. Critical realism justifies and supports the reuse of data through retroductive arguments and abduction, as new research questions and deeper theories might be better suited to explaining the objects (in this case, the units) under study, potentially offering new perspectives and understandings of a phenomenon. This

necessitates an understanding of ontology prior to epistemology. However, the reuse of data must be approved by the respondents, which was the case here.

5.6 Validation and generalisation

This research has suggestively maintained a stronger objective outlook guided by the theoretical propositions promoted by realist theory (Yin 1994). It has arguably used suitable methods and instruments for answering the overarching research questions and the subsidiary research questions. In addition, this research enhances trustworthiness by being transparent about the research process.

This multiple case study research project has used case study strategies presented and schematised by Yin (1994, 33), to enhance validity (i.e., using multiple sources of evidence, establishing a chain of events), to increase internal and external validity (i.e., using different analytical strategies (pattern-matching, explanation-building, time-series analysis)).⁴⁸

According to Yin (1994, 33), external validity is also connected to replication logic. With regard to theorising about generative mechanisms, one could replicate similar results by using the same methods and sources, i.e., the prominence of the same generative mechanisms would emerge from the data. Nevertheless, in other cases, the dominant reflexive modes could vary.

Applying mixed methods, in this case qualitative interview and quantitative ICONI-data, allows reliability and validity checks. However, the qualitative interview provided a richer and more nuanced understanding of the agents' subjectivities, the way they related to PISA in their internal conversations and their reasons for action. To enable its use across cultures the ICONI scale is non-referential, which means it only measures the prevalence of personal concerns and commitments in general (reflexive modes), without relating these concerns to concrete contexts and action. The lack of mix-method design in article III is thus not considered vital for the validity of the order of the reflexive modes that were identified.

This research did not draw on an explicitly rigorous case study protocol (Yin 1994, 64–65), due to a lack of training with this approach; relatedly, Yin (1994, 76) has stated that 'expertise

⁴⁸ Internal validity can increase by using analysis with same conceptualisations of centralized and decentralized systems, such as those found in Skinningsrud (2019), which indicate that the Norwegian educational system is centralized beyond 2010. This can be further corroborated by Archer's theory on centralized and decentralized systems as some structures are more prevalent in the Norwegian educational system today, making it hard to argue that the system is otherwise because fundamental characteristics of a decentralized system in Archer's terms are weak or absent.

with [...] these activities [...] will improve with practice'. A case study protocol was followed more loosely. One may argue that this might weaken the reliability of the study, as it could make it more difficult to replicate the process, but the transparency of the conceptual framework, references, interview guide and general openness weigh against such claims, even if the protocol was not written. In the interviews, the respondents were able to add additional information and clarify any misunderstandings, which is a benefit of semi-structural interviews. All codes abstracted from the empirical material were translated into English with the aid of Google Translate. I undertook the data collection and analysis myself, leaving less room for other interpretations. I did not validate the reflexive modes with the respondents, as that would possibly increase the chance of social desirability bias. All these strategies and processes have contributed to stronger corroboration of the results in this research i.e., the characterising educational system and the order of the two prominent reflexive modes.

My own tacit knowledge from being a human myself, working as a substitute schoolteacher, and the involvement with the PISA phenomenon and Archer's theories, have influenced this research and propositions, focusing on school personnel and mathematics as a teaching subject. In scholarly terms this is known as naturalistic generalisation i.e., experiences and expectations in tacit or articulated forms that shapes the research process and even affects the results (Stake 1978). Another form of generalisation that was applied in this thesis is analytical generalisation. I have used a theoretical template which contextualise and address mechanisms, where own findings are supported with this theoretical template. This enables transferability to other cases (Yin 1994, 30–31). Each article is an example of analytical generalisation, given that the generalisation was based on theoretical categories used in this thesis, with abduction as a reasoning logic, moving back and forth between the theoretical template applied and the specific research conducted. This means that each article includes mechanisms that are operating between cases, such as properties of the educational system, reflexivity and reflexive modes, structural power-relations, and agential mediations and justifications. Statistical generalisation cannot be applied in this research, which is primarily based on qualitative data and using ICONI with only one respondent. However, one could argue that the survey from the Union of Education Norway used in this project has these properties and might signal the effects of PISA on school leaders and teachers. However, it should be noted that this survey is from 2008.

5.7 Critical remarks on the theory, design, and methods

Archer concept of fractured reflexivity can be discussed due to some restrained applicability for qualitative research on PISA. Fractured reflexives are noted as passive agents that are unable to have projects. I argue that we are all fractured reflexive to some extent in life when filtering some structural or cultural phenomena. In fact, Archer's ICONI measures all four reflexive modes, but in a more general term. Hence, it embraces this 'passiveness'. The question is when will the qualitative interviews indicate fractured reflexivity among individuals that hold social positions where they already are familiarised with the PISA phenomenon? Fractured reflexivity can be expected when school personnel are not known to PISA, or understand it, or have no project with PISA. The question is if we ever can expect such a direction to be identified amongst school personnel in qualitative interviews if we focus upon them having no project with PISA, instead of focusing on them being disorientated. If we focus on disorientation rather than passive agents, we can address absences in individuals' knowledge base or thwarted structures that may need correction. Such a focus could inform leaders of knowledge gaps that could be turned into seminars and courses. A question is whether it will be ethically correct to identify and address such knowledge gaps amongst the respondents in relation to PISA. In this thesis I have not been preoccupied with such identifications as I have followed Archer's use of the concept rightfully, which meant that I did not expect the respondents to be passive agents with PISA. These are the reasons for not incorporating fractured reflexivity further in the qualitative analysis of the two articles on reflexive modes.

Another experienced restraint with this research is the case study framework. Case studies focus on contextualisation and theorising on mechanisms than the focus on data points (interview data, scale-numbers, statistics). This shift from data points to contextualisation can easily prove right in case studies where the extent of data points might become reduced in favour of the explanatory mechanisms (e.g., theories). Hence, less time, less respondents, less data. Despite this, this research has succeeded to put forward hypothesis about the educational system and reflexive modes amongst school personnel which can further be investigated. That said, it is also important to discuss changes in the educational system i.e., doing historical analysis as structure pre-dates agency and the fact that it gives more information about the status quo of Norwegian education altogether by encompassing both structure and agency.

There are also limitations in the interview guide in the qualitative interview with the mathematics teacher. Asking for contextualised data diminished the number of PISA

questions. Also, the use of a semi-structured interview strategy had consequences for the conversation; it shifted focus now and then. Moreover, questions on OECD's intentional work on making progress in the world through comparisons was not an explicit topic with the respondents, which could possibly bias the respondents towards becoming more pro-PISA, making them less meta-reflexive. However, it is unlikely that meta-reflexivity and even autonomous reflexivity would have changed as respectively the second and first prevalent mode for this specific reason due to the respondents way of operating with PISA.

Nevertheless, these historical perspectives on OECD's PISA can be integrated in the PISA research and interview guides because they provoke questions about 'combinability with other governing complexes and political agendas', 'implications [on] the human condition' and 'the future of education as an institution in society', questions which are put forward by Ydesen (2019a, 300–301). Furthermore, the claim that PISA is a cultural neutral and non-curricular test and entangled with edu-business PISA have not been discussed with the respondents due to late realisation of relevance. However, these topics can be included in further research on PISA. These topics can challenge respondents autonomous reflexivity and meta-reflexivity; however, it is not likely the order of them being changed due to PISA being grounded in law.

PISA spin-off products such as TALIS (PISA for teachers and school leaders working conditions and learning environments at their schools), SSES (Study of Social and Emotional Skills) (PISA for 10- and 15-year-olds), PISA for Schools (PISA local school test on demand) and PISA4U (online learning modules and collaborative activities on PISA data for teachers), and the Future of Knowledge and Skills 2030 project for curriculum standards were neither asked about. The main PISA test was the focus of this thesis. The inclusion of spin-off products in the interviews could have augmented or diminished manifestations of all reflexive modes. Since spin-off products were existing at the time the semi-structured interviews were conducted, although in different numbers, the respondents had the opportunity to discuss them. This was not the case. As this research is informed to a certain extent by Sjøberg's writings, this affected the architecture of this thesis and research strategies with the consequences of focusing on the main PISA test. Personally, the interest was also in the main PISA test. To solve the problem with spin-off products, if they are even relevant at the specific school, one can ask whether these are vital products for how the respondents characterise their relationship with the PISA phenomenon.

5.8 Summary

This research is based on a case study framework. Case studies are suitable when you are studying the relationship between the phenomenon, the context, and the people. The case study consists of three units: the internal structure of the educational system, a mathematics teacher and three school leaders, constituting a total of three articles. The three articles consist of five multiple holistic case studies, where each case in the articles is defined as a case in itself, thereby, cumulating five cases in total. Each article has been situated in its own context.

Article I, which is a historical examination, uses data from educational literature, green papers, white papers, a blog, and an autobiography to answer the research questions. It utilised a chronological time series analysis and explanation-building analysis. Article II and III use semi-structural interviews with a mathematics teacher and three school leaders in lower secondary schools, where the interview with the mathematics teacher is in combination with ICONI (Internal Conversation Indicator). All qualitative interviews in this project were piloted. Article I and III used manual analysis, and article II used NVivo software for the analysis. Both articles II and III utilised explanation-building analysis and thoughts from program-logic model analysis.

I have valued respondents' confidentiality and anonymity as the prime concern for this research and provided few identity markers from the respondents. The respondents have been given information about the research and been able to pursue further information about this project. I have argued that the research has maintained a stronger objective outlook by using realist theory propositions. I have also used suitable methods and instruments that (has) enable(d) answering the overarching research questions and subsidiary research questions posed. I argue the results in this thesis has been stronger corroborated through the research strategies and processes in this research.

I have used naturalistic generalisation, as my background such as my tacit and articulated knowledge has influenced this research. This research also enables analytical generalisation of mechanisms between cases, which can be replicated due to transparency of the research process. Statistical generalisation is not used as the research contains qualitative data and the ICONI (Internal Conversation Indicator) has only been used for one respondent. I have noted some difficulties with Archer's fractured reflexive concept in the qualitative analysis, which had consequences for the further use of the concept. These are possible challenges that might occur using her reflexive conceptual framework with PISA. Moreover, a few critical remarks

on the design and methods have also been addressed such as the focus on contextualisation in case studies, the neglect of some themes about PISA found in the introduction chapter of this thesis and absence of discussing spin-off products in the interviews.

6. Results and analytical discussions

6.1 A specific reply

In critical realism, the meta-theory, which is underpinned in Archer's theoretical work, entail rational arbitration on the merits of different theories by assessing, for example, their relative explanatory power. In addition, that all pervious knowledge can be fallible and open for reexamination and thereof possibilities of 'reclaiming reality' (Bhaskar 2011). Whether this reclaim of reality is a success depends on the evidence and argumentation. This is especially relevant for article I, but could also be understood when holding effects solely on policy and governance, or in schools, leaving out effects on the actors themselves, which is counteracted by my two articles in this thesis by effects on reflexivity (art. II and III). The overarching conceptualisations which have enabled this 'reclaim' through reconceptualisation of PISA effects is the most fundamental concepts in critical realism: ontological realism, epistemological relativism, and judgemental rationality. This thesis is a specific reply to (Pons 2017) critical review on PISA effects where case studies within this field 'rarely conceptualise the PISA effects themselves and do not always distinguish the key variables or factors that can explain why and how such effects occur or not' (Pons 2017, 138), where he propose the use of overarching theoretical conceptualisations. In this thesis, the key factors and variables are the educational system and reflexive agency with their mechanisms. Pons also recommends using the model 'reception, uses and effects' with this survey. Pons arranged model suits the logic of the analysis of the three articles: from socio-cultural interaction (amongst scholars and stakeholders) with PISA following the introduction of PISA, but also relevant for scholars alleged PISA effects on the Norwegian educational system (art. I). Moreover, the model also fit with the socio-cultural interaction with each of the respondents on PISA (art. II and III). Thence, making claims of effects on the Norwegian educational system and on agency given their activated mechanisms i.e., the place of situ of deep effects of PISA.

6.2 The overarching research question I

- i. The first theoretical research question was as follows: How does Archer's theoretical approach enable a reconceptualisation of alleged 'PISA-effects' on the Norwegian educational system?

Archer's theoretical approach enables a reconceptualisation of alleged 'PISA-effects' on the Norwegian educational system because she provides a system theory for discussing such

effects on the internal structure. Thus, alleged PISA effects are discussed through mapping the educational system characteristics before and after the introduction of PISA, where the latter was initially followed by a ‘PISA shock’, causing a claimed turn for Norwegian education. The systemic essence Archer provides is absent in the field of PISA effects amongst scholars claiming the educational system is decentralized, or simultaneously decentralized and recentralized after PISA reforms. This flat theorising occurs because scholars do not provide a real definition of the system and conceptualise the (two types of) educational system(s), which is based on emergence, causal powers, and autonomy. This means that there are different gateways to centralized and decentralized claims based on different theoretical premises. As Pons (2017) has pointed out, one of the difficulties in summarising the various findings on ‘PISA effects’ is their divergent theoretical frameworks. The same theoretical terms are defined differently within each framework, almost prohibiting comparisons and making it difficult to relate new research to previous research. The field of ‘PISA effects’ is ‘fuelled by many individual contributions from various disciplines and academic traditions, or by some specific groups of scholars whose works are rarely confronted in a dialogical and cumulative way’ (Pons 2017, 133).

6.2.1 Discussing findings about structures and processes in the Norwegian educational system before and after the ‘PISA shock’

Mausethagen (2013), whose work has been discussed in chapter 2 in this thesis, studies the role that concepts, or conceptualisations, play ‘in soft governance’. She describes her theoretical framework as ‘the constructivist theoretical paradigm’ from the field of ‘international relations’. Her methodology is inspired by discourse analysis, and is applied in analysing Norwegian policy documents, that is, White Papers issued during the period 1995-2010. Her object of study is the influence of international organisations (specifically the OECD) on national educational reforms. Mausethagen’s theoretical perspective could be broadly categorised as ‘neo-institutionalism’, since she is interested in ‘ideas that travel’, and that globally circulated ‘master-ideas’ not only provide solutions to problems but also define which problems that needs to be addressed (Pettersen and Røvik 2014).

Mausethagen’s focus on how concepts (specifically the notion of competence) influence educational ideas among partners in discourse, contributes to our understanding of how policy agents at the national level are subtly persuaded to adopt OECD’s policy ideas. Influencing policy agents at the national level to accept and adopt OECD policy is a precondition for

subsequent PISA-effects at the level of the national educational system. Mauseithagen discovered, in her study of White Papers, that when competence was discussed, the defined ‘problem’ was a general lack of competence in Norwegian schools, far below the level needed in today’s knowledge society. Solutions to this problem were ‘often framed with an emphasis on the need for more competence’ (Mauseithagen 2013, 174).

Mauseithagen’s approach, discourse analysis applied to the study of policy documents (White Papers), entails the study of ideas, how they are spread and adopted. Her focus is on culture, less on social structure and agency. Her presentation of the Norwegian educational system⁴⁹ emphasises ‘Norwegian educational policy’⁵⁰, centring on ‘policy *ideas*’ and *ideas* about the system. Mentioning compulsory comprehensive schooling, social inclusion, and egalitarian ideas, as well as the predominance of public education in Norway, where 97 % of the students are enrolled, she focuses on the *justifications* for keeping private provisions at a minimum, i.e., the promotion of equality and democracy.

Presenting Norwegian educational policy (or the educational system?), Mauseithagen draws attention to a ‘form of decentralization’ associated with the Knowledge Promotion Reform of 2006, which, again at the level of ideas, is stated in a policy document (Ministry of Education 2004). This document emphasises that teachers, principals, and municipalities will maintain the flexibility and discretion to make decisions about pedagogical practices, to achieve the competence aims in the National Curriculum. This ‘form of decentralization’ was by the Ministry described as embodying “freedom, trust and responsibility”. Referring to Karseth and Sivesind (2010) and Skedsmo (2009), Mauseithagen claims, however, that ‘the state remains a strong actor working toward the goal of systemic change’ (Mauseithagen 2013, 164). Concerning how the OECD exerts influence on Norwegian education, Mauseithagen states that ‘[t]he broad approval for the new educational reform, *The Knowledge Promotion*, in 2006, would likely not have been possible without the OECD’s assessment studies and country reports’ (Mauseithagen 2013, 165).

At the structural level, Mauseithagen describes Norwegian education as characterised by comprehensive schooling and social inclusion, and generally as a ‘highly regulated education

⁴⁹ (Mauseithagen 2013, 162).

⁵⁰ (Mauseithagen 2013, 163).

system' (Mausethagen 2013, 162). When she mentions that some 'form of decentralization'⁵¹ was implemented in connection with the Knowledge Promotion Reform, she is however, referring to how this was justified at the level of ideas. Social and systemic structures are not central to Mausethagen's analysis. However, she mentions structural characteristics, such as the predominance of public sector schools, comprehensive schooling, the principle of inclusivity, and some 'form of decentralization'. An interpretation of her conclusion, as it relates to PISA, could be that PISA together with other inputs from the OECD contributed to introduce some (form of) decentralization in the Norwegian system by ensuring the unanimous passing of the Knowledge Promotion reform in the Norwegian Parliament. Mausethagen's implicit notion of decentralization, which is widely accepted among Norwegian educational researchers, seems to be delegation of decision-making to lower levels of authority. Her theoretical framework, which is cultural, lacks more detailed conceptions of internal structures and processes in educational systems.

Likewise, Baird et al. appear to equate educational decentralization with the delegation of decision-making, claiming that 'Norway has a decentralised education system, with many of the decisions being made at a local level' (Baird et al. 2016, 127). Moreover, these authors refer to the Norwegian system as 'devolved', underlining that a devolved system entails local diversity, including 'a wide range of assessment practices' (Baird et al. 2016, 127). Furthermore, Baird et al. underlines that in terms of direct central interventions, devolved systems, like the Norwegian one, have only a limited amount of such interventions. Despite their description of the 'devolved [Norwegian] system' as a system with few central interventions, they claim that 'the PISA shock led to a series of reforms of both curriculum and assessment', and that 'evidence from PISA was a large part of the justification for change' (Baird et al. 2016, 128).

Among the 'central interventions' mentioned by Baird et al., which were justified by PISA results were The National Quality Assessment System (NQAS), introduced in 2004, into which both PISA and national tests were incorporated. Influence from PISA was also at play in the development of national tests in reading, which were inspired by the PISA framework, the test developers being members of the Norwegian PISA Team. The Knowledge Promotion Reform is also mentioned as justified by PISA results.

⁵¹ (Mausethagen 2013, 164).

A widely accepted reading of Baird et al.'s article is that increased centralization in the educational system after the PISA reforms was not to be found. However, a closer reading of their results makes it less obvious that Baird et al.'s findings warrant this conclusion. One of their findings was that the centrally initiated Norwegian reforms after the PISA shock were 'standards-based', counteracting local diversity (Baird et al. 2016, 132). In Archer's terminology reducing local diversity in favour of national standards would entail 'increased unification' in the system and indicate that centralization was being strengthened. However, Baird et al. make the proviso that 'we do not have firm evidence that centralisation was pursued *because* of PISA results, as standards-based reform and the centralisation that it entails has a longer history' (Baird et al. 2016, 132). Here, apparently, the authors are contradicting their previous claim that the Norwegian system is [or has been] a devolved system.

Mitigating circumstances to Baird et al.'s lack of clarity concerning centralization and decentralization in the Norwegian system may be that their main concern is to show that PISA-effects (in terms of reforms justified by PISA results) are not promoting global policy convergence. The main reason being that PISA results are interpreted differently in accordance with different national histories and traditions.

Nortvedt (2018) contributions to previous PISA research have been extensively discussed in chapter 2. However, it is enlightening to examine one of her central references in discussing possible PISA effects. For descriptions of the Norwegian educational system and the current educational policy, Nortvedt relies heavily on Imsen, Blossing and Moos' (2017) account.

Imsen, Blossing, and Moos (2017) consider the turn of the millennium as a starting point for a restructuring policy in Norwegian education. Their question is: How have the basic values of the Nordic Model in education been affected by recent educational reforms? They present a long list of the essential features of The Nordic Model, which is the yardstick against which the new millennium reforms are measured by the authors. The Nordic Model is their base line, so to speak. Identifying central traits of the Nordic Model, they list: 'equal access to education, a common core of subjects, social community, democratic student cooperation, no segregation with regard to ability, gender, or social class (i.e., no organisational streaming), differentiation within mixed-ability classes, and individualisation adapted to students' prerequisites in order to provide a meaningful learning environment for all. Flexible national curriculum plans and open-ended learning objectives are important conditions needed to

achieve this, along with trust in individual schools and professional teachers as the main resources in the construction of educational practice' (Imsen, Blossing, and Moos 2017, 570). This is the yardstick against which the new reforms at the turn of the millennium are measured.

Describing the restructuring of Norwegian education at the turn of the millennium, Imsen et al. use some key phrases: decentralisation, a strong emphasis on competence aims and learning outcomes, increased emphasis on assessment and the introduction of a vast national test system, increased national and local control, and research- and expert-based school development. PISA results were used to justify educational reforms. New responsibilities were transferred to the municipalities. As school owners they were given the responsibility for quality control, having to establish their own quality control system. They were obliged to construct their own local curriculum plans, to set up their teachers' in-service training programs, and to carry out wage negotiations with teacher unions. Crucially, the municipalities were also required to report to central authorities about their 'achievements'.

The crucial element in these 'decentralizing interventions' of delegating tasks and decision-making was the obligation of the municipalities to report upwards in the system about their 'achievements'. Imsen et al. consider the mentioned delegated functions to have had the opposite effect of decentralisation, namely, to strengthen state control, and they call it *recentralisation*. The new design meant to decentralize the system instead, and in an indirect way, it strengthened state control resulting in 'recentralisation' (Imsen, Blossing, and Moos 2017, 574). Consequences of these 'decentralizing' interventions, were a growing bureaucracy, which had to deal with written reports at all levels of the system, and that teachers' time was diverted from teaching to 'paperwork' (Imsen, Blossing, and Moos 2017, 574).

Nortvedt (2018, 438), referring to Imsen et al, is not quite clear on whether she agrees with the reconceptualisation of decentralization as 'recentralisation', describing it as 'a level of inconsistency.' Imsen et al. mention a number of paradoxes that are experienced among Norwegian teachers, among the new ones is the centralization/decentralization paradox.

Centralization and decentralization also feature as a central theme in the recent article by Camphuijsen, Møller, and Skedsmo (2021) (discussed in chapter 2) writing about Test-Based Accountability (TBA) in Norwegian education from 2003 to 2016. The authors' intent is to

identify the ‘drivers’ and rationales for adopting, developing, and retaining TBA in Norway. Centralization or decentralization in the Norwegian system is not part of any of their research questions. However, the notion of decentralization becomes relevant, as their article repeatedly refers to ‘the highly decentralised Norwegian educational system’ (Camphuijsen, Møller, and Skedsmo 2021, 626, 636). Moreover, the allegedly decentralized Norwegian system is associated with one of the detected ‘drivers’ for adopting TBA, namely the value of equity.

Identifying the Norwegian educational system as ‘highly decentralised’, Camphuijsen, Møller, and Skedsmo (2021, 627) refer to Christensen and Læg Reid (2011), who discuss the influence of New Public Management (NPM) ideas, introduced in the late 1980s, with consequences for subsequent public sector reforms. Decentralization was one of the influential NPM ideas. In a further description of the ‘Norwegian educational context’ the authors also mention that the system promotes values associated with equity, solidarity, social justice, and democracy. Moreover, the comprehensive school model is intended to promote equal opportunity.

Presenting the ‘educational context’, the authors also state that although 82 % of the Norwegian population live in urban areas, ‘many municipalities and schools are small[,] [s]chool choice is limited, especially for compulsory education’ (Camphuijsen, Møller, and Skedsmo 2021, 627). This demographic information may project an image of huge geographical distances and dispersed settlements as one dimension of ‘the highly decentralised’ Norwegian system - which is not entirely untrue. However, the authors also draw attention to another type of decentralization, which is mentioned in the 1988 OECD’s ‘Country Review’ of Norwegian education. This report claimed that the Norwegian system was too decentralized in the sense of needing a stronger role for the state (Camphuijsen, Møller, and Skedsmo 2021, 627–628).

The authors further substantiate their claim about a high degree of decentralization in the Norwegian system by referring to several responsibilities being devolved to local education authorities and individual schools in the early 2000s. The delegation of responsibilities was, however, tempered by a simultaneous introduction of national testing (where PISA results was a cause for that implementation), teacher monitoring and evaluation, as well as the introduction of an outcome-based curriculum. Thus, measures promoting decentralization were introduced in parallel with measures promoting centralization (Camphuijsen, Møller,

and Skedsmo 2021, 625). This account corresponds closely to Imsen et al.'s description of 'decentralizing' reforms. However, Camphuijsen et al. do not explicitly draw the same conclusion as Imsen and her associates, namely, that the 'decentralizing' reforms did not promote decentralization.

According to Camphuijsen et al. the high degree of decentralization in the Norwegian educational system was one of the factors that made TBA attractive to Norwegian politicians. TBA would ensure that national standards were upheld in a system with a high degree of regional and local diversity. TBA was initially adopted to ensure equity and quality standards in 'the highly decentralised Norwegian educational system'. The authors admit, however, that by now, 'to some degree [...] equity and equality has been rearticulated to performance indicators' (Camphuijsen, Møller, and Skedsmo 2021, 636–637). The legitimization of TBA in Norway, however, differed from other countries that were 'early adopters' of TBA. In these countries, TBA was used to justify market-based reforms, while in Norway equity (and quality) were drivers (Camphuijsen, Møller, and Skedsmo 2021, 637).

Summing up the merits of TBA, Camphuijsen et al. conclude:

Our analysis portrays that TBA formed a key policy instrument to modernise and raise the performance and equity of the Norwegian education system. TBA replaced a steering tradition based on prescription and intervention, by allowing government officials to steer a highly decentralized education system from a distance, by means of outcome measures, visibility, comparison and accountability. (Camphuijsen, Møller, and Skedsmo 2021, 636).

This statement does not clarify whether the institutionalisation of TBA contributed to making the system less decentralized. One would assume that a system where there is 'steering at a distance' tends towards centralization, i.e., minimising local diversity. In Archer's theoretical reasoning, the prime characteristic of centralized system is that 'the state is the leading part'.

Camphuijsen et al. are, however, optimistic about the prospects for TBA, since it functions as 'an empty vessel', which can be 'filled with' different content and promote different values, depending on local circumstances. However, in the Norwegian context, these authors see a paradox in the National Curriculum Guidelines providing 'a broad framework allowing autonomy [for] local schools, [while this] scope is narrowed by the municipal use of national standardised tests' (Camphuijsen, Møller, and Skedsmo 2021, 638). At the same page, they

also concede that standardised tests foreground certain aspects of teaching and learning, while constraining others.

One of Camphuijsen and her associates' basic premises in describing how TBA step by step was institutionalised in the Norwegian educational system is that the system, when TBA was first introduced, was 'highly decentralised'. TBA was politically attractive because it could contribute to 'uniformity', i.e., upholding a common national standard in a system plagued with divergence, and regional variation. Moreover, TBA was seen to promote equity and quality, which resonated with traditional values in the Norwegian system.

Comparing Imsen et al.'s and Camphuijsen et al.'s analyses, which cover roughly the same time period - Imsen et al.'s study from 1990 to 2015 and Camphuijsen et al.'s study from 2003 to 2018 - they basically agree on what they have observed. However, they analyse it in different ways, using different concepts and different terminology. Imsen et al.'s conceptualisation is 'recentralisation', while Camphuijsen et al. call it 'steering from a distance'. However, they seem to agree that towards the end of the period they cover, the state has gained more influence on the system and its leading part. Both 'recentralisation' and 'steering from a distance' entail a more important role for the state. The state has gained power, in Camphuijsen et al.'s words by 'steer[ing] [...] from a distance' and in Imsen et al.'s words by 'recentralisation'.

A difference between their analyses is, however, that Imsen et al. see the values being promoted by state policy during this period as a threat to the traditional value of equity in the Nordic Educational Model, while Camphuijsen et al. underline that the new policy, on the contrary, has increased equity, as well as raised performance levels.

What is of particular interest in the context of this thesis is that both studies may be seen to describe and try to conceptualise an increasingly centralized system. This resonates with my analysis in article I, where I used Archer's concepts of centralization and decentralization and her concepts of internal processes in state systems. I claimed that the Norwegian system became increasingly centralized after the PISA shock and interventions justified by the PISA results.

However, my examination of Imsen et al.'s and Camphuijsen et al.'s studies also illustrate the point made by Pons (2017), that studies of PISA effects (and effects of other ILSAs) are informed by different research traditions and different schools of thought. The diversity of

theoretical frameworks makes it difficult to establish a cumulative body of knowledge both about PISA effects, and of causal effects in general, in the broader field of educational development. Hence, alleged PISA effects need to be contextualised with other processes in the education system to follow scholars argumentation on centralization and decentralization. It requires considerable skills in the interpretation and translation of concepts and terminology to find out whether different researchers basically agree or seriously diverge in their accounts of the same social reality.

6.2.2 Subsidiary research questions - article I: The Impact of PISA on Education in Norway: A Morphogenetic Perspective on Structural Elaboration in an Education System

Author: Terje André Bringeland

The overall research question for article I:

- After the introduction of PISA, the following public debate and the implementation of reforms justified by PISA, was there a change in which kinds of structures and processes that predominated in the [Norwegian] system?

The subsidiary research questions for article I:

1. What kinds of structures and processes predominated in the Norwegian educational system in the decade before the introduction of PISA?
2. Who were the protagonists in the public debate after the release of the first PISA results, and what were their concerns and projects?
3. Which systemic structures and processes were strengthened, and which were weakened by the reforms legitimized by PISA results?

This article, which is based on document studies, examines the possible effects of the PISA test on the structure of the Norwegian education system. The article employs Margaret Archer's conceptions of morphogenetic cycles and educational systems with their inherent systemic mechanisms of unification, systematization, differentiation, and specialization. The analysis of the structures and processes taking place in the system before and after the 'PISA shock' and debate indicate that the same types of systemic structures and processes continued to predominate after the introduction of the PISA test. The centralized system continued

despite new reforms and other interventions justified by the PISA results. A change in the fundamental characteristics of the centralized Norwegian education system did not happen.

The introduction of the PISA test was enabled by the introduction of New Public Management (NPM) principles in the Norwegian public sector and the educational system. The new principle, management by objectives (MbO), was introduced at about the same time as the 1988/89 report from the OECD, which emphasised that Norway lacked statistical evidence of educational outcomes, alongside a warning that the system was becoming too decentralized. The Labour government in 1996 made the decision for Norway to participate in PISA. However, 'it was a Minister from the Conservative Party, in a coalition government, who found PISA useful in legitimising her party's educational policy' (Bringeland 2022b, online abstract). The PISA results were a 'flying start' for conservative educational policy (Bergesen 2006, 40).

Among the protagonists in the PISA debate were education Minister Clemet, some representatives from the various political parties, teacher union leaders, and a university professor of education. Their reactions to the test results were communicated through the mass media. Clemet wanted a knowledge-based school. The Socialist Left Party focused on *Bildung* in a broader sense (Bergesen 2006, 82). The Norwegian social-liberal newspaper *Dagbladet* contrasted the results with Norway's top performance in the Winter Olympics. Teachers were not fans of the test (Bergesen 2006, 43). 'We are good enough', the school must have self-confidence and hold on to its mission, said the leader of the Union of Education Norway, Helga Hjetland⁵² (Bergesen 2006, 41). 'This is too stupid, it is groundless and unscientific to say that the Norwegian school gets too much and gives too little. Well-being is important for long-term learning' said Anders Folkestad, the leader of the Teacher Union (Bergesen 2006, 41). PISA says nothing about quality of Norway's schools, said Professor Stefan Hopmann at NTNU (Norwegian University of Science and Technology) (Telhaug and Mediås 2003, 327).

The question is whether PISA would have been introduced by another education minister from another political party, and if the response from politicians and the wider public to the test results would have been significantly different under another education minister. It is

⁵² Before the merger with the Union of Education Norway, she was the leader of the Norwegian Teacher Union (Norsk lærerlag).

likely that PISA would have been introduced by another education minister from another political party since Norway's participation was already decided in Parliament in 1996; secondly, the response to the test results could possibly have been a bit different depending on the political party affiliation of the minister, but it is vital to remember that interests and processes that already were in motion in the education system prior to PISA continued after the introduction to PISA. It is not likely that another educational minister at that time would have rejected PISA, or the results. One question is central: would NQAS have emerged without PISA? This is a counterfactual question that aims to find the causal effect of PISA—for example, whether the introduction of the Norwegian NQAS was a 'PISA effect'. By contrast, an associational question would merely ask whether NQAS came after PISA. Most likely, the NQAS would have been implemented without PISA, as there was no necessary connection between these two elements. 'PISA was [just] one of many possible [justifications: the] NQAS could have been introduced for many other reasons, such as to enable the effective implementation of national policy, ensuring the uniformity and quality of provisions, etc. Hence, the NQAS could have come into being without PISA' (Bringeland 2022a, 172).

The Norwegian education system was already centralized prior to PISA; therefore, alleged PISA effects (for example, the NQAS and the K06) did not alter the system significantly. Djupedal (2022) state that extra hours in mathematics, science and reading came as PISA effects. These changes did not alter the characterising educational system. Instead, these changes strengthened processes that were already prominent, i.e., unification processes in the educational system. The prominence of unification and systematization processes in the educational system entails a weakening of other inherent mechanisms that is needed for stronger professional autonomy in schools, i.e., differentiation and specialization, contrary to what has been suggested by Østerud (2016, 32). In Archer's terms, the educational development in Norway did not take a fundamentally new turn with the introduction of PISA and reforms legitimised by PISA. The Norwegian educational system was centralized before and after PISA (Bringeland 2022a), and also continued being centralized after 2010 (Skinningsrud 2019). Still, to date, PISA is a part of NQAS (Sjøberg 2023), i.e., under the unification mechanism. This has side effects on other mechanisms in the educational system.

6.2.3 A short comment

The finding that the system remained essentially the same after the PISA shock resonates to a certain extent with the two PISA researchers at ILS and CEMO at the University of Oslo who argue that ‘it is obviously not possible to establish any clear causal link between Norwegian participation in international assessments and the changes that have taken place [in Norwegian education] during the same period’ (Björnsson and Olsen 2018a, 20). Basically, the PISA test did not cause a fundamental change of the Norwegian education system. On the same page, Björnsson and Olsen, referring to Nortvedt (2018) claim that ‘the development of policy takes place in a complex cultural context where recommendations for Norwegian schools from the international organisations are not necessarily followed up. Instead results and analyses from international studies are used to justify already existing policy trends, a finding that resonates with analyses made in other countries, e.g., by Takayama (2008)’. One can discuss to what extent PISA has had damaging effects on education worldwide since the Norwegian case reveal that the educational system already was centralized before and after the introduction of PISA, and that management by objectives and NPM was normative establishments. Maybe teachers and school leaders reflexive modes can tell us something about that in the next subchapters. Hence, it should not be undermined that PISA is a part of the unification mechanism in Archer’s theory. This has side effects on other mechanisms of the internal structure of the educational system such as differentiation (teacher profession) and specialization (teaching content).

6.3 The overarching research questions II

- ii. The second theoretical research question was the following: How can Archer’s concept of ‘reflexivity’ and ‘reflexive modes’ increase our understanding of how school personnel (teachers and school leaders) react to the idea of using PISA-test results in their own educational setting?

A side effect of using a systemic theory that encompasses agency and causality, is that a reconceptualisation of PISA effects was also enabled on agency. Agency has emergent properties, autonomy and causal powers and is a necessary connection for filtering the environment. Hence, the focus of PISA effects on school leaders and teachers were redirected to reflexivity and reflexive modes. Rather than seeing school personnel’s reactions to the PISA test as inexplicably uniform or inexplicably diverse, or settling for categorising them as binary, ‘for’ or ‘against’, the concept of reflexivity and reflexive modes enable a deeper

understanding of why persons respond and react in different ways to phenomena in their environment, such as PISA. The concept reflexivity draws attention to the fact that school personnel, and persons in general, conduct internal conversations, i.e., deliberate in their minds, on how to deal with, or reconcile, external demands made on them, and their own deeper concerns, personal projects, and what matters most to them as persons. Thus, the person is no longer a 'black box' between in-put and out-put. We get a better insight in what is going on inside the 'black box' and get an understanding of how structural constraints (and affordances) are mediated by the agent.

The assumption is, which has also been substantiated by Archer's empirical research, that persons have concerns, i.e., certain things matter to people. These might be family and friends, work, or ethical issues. The 'things that matter' are not mutually exclusive, but given external constraints, most people must prioritise between their concerns. This is done in their internal conversations, where they develop a *modus vivendi*, a prioritisation of personal concerns that they can live with, at least for some time – between revisions. Archer's distinction between three reflexive modes reflects the relative prevalence of distinct concerns: communicative reflexivity is centred on 'friends and family' as primary concerns, autonomous reflexivity is centred on work results, and meta-reflexivity is concerned with ethical issues, critically examining both the external environment and themselves. Fractured reflexive agents are agents that are unable to have projects, they are passive agents. This type of reflexivity is not focused upon in article II and III as school personnel is considered to have projects with PISA.

The causal power of PISA, as an element in the educational social structure (as part of NQAS), is mediated by agents' (persons') reflexive modes. The activation of reflexive modes filters the impact of the environment. Further details and explications of how reflexive modes function as filter of environmental impacts and reveal deeper layers of agency is presented in articles II and III.

Before entering the two next subchapters, it can be beneficial for the reader to be reminded of the introduction of this thesis: what is human? and what is schooling about? and what role will reflexivity and reflexive modes play in this (and on emotional commentaries and further actions)? These questions, I will leave to the reader.

6.3.1 Subsidiary research questions - article II: PISA and Teachers' Reflexivities. A Mixed Methods Case Study

First Author: Terje André Bringeland

Second author: Tone Skinningsrud

The research question for article II is presented below:

1. Which modes of reflexivity are activated in a mathematics teacher by the PISA test?
Which mode(s) of reflexivity predominate(s)?

This article outlines three theoretical approaches that were applied in the study of teachers' reactions to the introduction of NPM reforms in education: labour process theory (Ozga and Lawn 1988; Reid 2003), which is a variant of structural Marxism; post-structuralism (Ball 2003, 2016), relying primarily on Michel Foucault's conceptualisations; and the theory of professions (Troman 1996; Svensson 2006; Evetts 2011; Adams and Sawchuk 2020). Margaret Archer's theory of structure, culture and reflexive agency is suggested as an alternative to these approaches since her conceptions of various types of reflexivity can account for individual differences in reactions to the same structural and cultural conditions.

This article is a further attempt to start filling the gap in our knowledge about PISA effects at the micro-level. Thus, the results from the single case study of a lower secondary school mathematics teacher (art. II) present the effects of PISA on the reflexivity of one mathematics teacher. The empirical data consist of a semi-structured interview and a Likert scale developed by Archer measuring modes of reflexivity. Archer's conceptualisations of reflexivity and reflexive modes were used as analytical categories. The results confirm that the mathematics teacher had internal conversations involving PISA. In his internal conversations about the test the autonomous reflexive mode predominated, which indicated that his primary concern was success in his work. Aiming to be a successful teacher, he adopted a strategic stance to his environment. Trying to familiarise his students with test-situations, he implemented a strategy of 'teaching to the test' by using items from previous tests to create 'trial' test situations. Thus, he tried to optimise his students' future test result.

Although autonomous reflexivity was the teacher's predominant reflexive mode, he also engaged in the meta-reflexive mode by in addition to criticising the test, deliberated on why he himself was so critical of the test. He emphasised that he was appointed to administer the

test rather than choosing it himself. The respondent, to some extent, also engaged in communicative reflexivity, i.e., concerning his collaboration with colleagues and discussing the PISA test with them. The identification of all three reflexive modes and fractured reflexivity and their order of prominence was validated by the ICONI-indicator.

Having established, in article 1, that the Norwegian system is a centralized system, where major decisions concerning the system are made at the central level of policy making, enable the hypothesis which is put forward in the article. This suggests that the combination of autonomous and meta-reflexivity, that is, being critical but at the same time strategically adapting to current states of affairs, may tend to predominate among teachers in centralized systems. The mathematics teacher expressed that exertion of power was the reason why PISA was not terminated. Thus, admitting that the structures of centralized systems are difficult to challenge at the school level since decisions about such matters are made elsewhere.

6.3.2 Subsidiary research questions - article III: School Leaders' Reflexive Mode in their Internal Conversations on PISA

Author: Terje André Bringeland

The research questions for article III are presented below:

1. Which reflexive modes are activated in the three school leaders when they engage in internal conversations about PISA?
2. Which is the dominant reflexive mode of each school leader (regarding PISA)?

This article is one of few studies in the field of education that aims to fill the gap in our knowledge about 'PISA effects' at the micro-level (Bringeland 2022c). The theoretical toolbox used is Margaret Archer's concepts of reflexivity and reflexive modes when identifying similarities and differences of lower secondary school leaders under the same structural circumstances. The study of three cases, which were based on semi-structured interviews, explores how PISA was engaged in three Norwegian school leaders' internal conversations about their work with the aim of identifying their reflexive modes. The findings suggest that both autonomous and meta-reflexive modes were activated when the three school leaders engaged in conversations about PISA. The dominant reflexive mode amongst the three school leaders was autonomous reflexivity; they dealt with their environment, emphasising leadership efficacy and tasks to be completed. The communicative mode was not

detected among the three leaders. In their predominant autonomous reflexive mode, the school leaders made autonomous decisions about how to use the test. Their major concern was to generate better results, mentioning national and local tests as relevant material when preparing for tests ('teaching to the test'). Only one school leader considered PISA useful for his school. The others emphasised that PISA does not give specific data back to individual participating schools, which makes it hard to use and justify. In the meta-reflexive mode, the school leaders discussed the (limited) range of usage, which evoked ethical and critical deliberations. However, two of these school leaders had either attended or were interested in attending PISA seminars/courses.⁵³ The three school leaders acknowledged that the PISA test is relevant in the Norwegian context for other purposes, i.e., comparison with other nations and input for policy changes.

In this article, the NQAS is one example of unification in Archer's terms, where educational testing is aimed for accountability (subject to centralized control) and learning. As written in Bringeland (2022c, 182): 'A central feature of the Norwegian Quality Assessment regime is that schools and teachers are trusted to use test results for the improvement of learning without introducing incentives or sanctions. This high level of trust, however, contains a paradox in the sense that at the same time as the improvement of individual learning from assessments is emphasized, the national control regime is tightened through uniform national guidelines for school and student assessments (Skedsmo and Mausethagen 2017, 176)'.

6.3.3 Discussing previous Norwegian micro-level findings with PISA before and after the use of reflexivity and reflexive modes

Previous Norwegian studies on PISA with school personnel have concentrated upon validity and the use of the test (Eggen 2010), and the perception and use of the test in connection to mathematics as a school subject (Bringeland 2015). Previous studies on 'PISA effects' have not used Archer's concept of reflexivity and reflexive modes, which might make it harder to analyse and understand the (longitudinal) trajectories of the agent, and institutional battles and changes, as the theoretical toolbox pinpoints agential orientations and directions. However, despite not using these conceptualisations, previous studies do inform about thoughts, decisions, and actions with the test, but clearly have other focuses in the analysis. But they

⁵³ More information about PISA seminars/courses are found in Aursand (2018) and Aursand and Rutkowski (2021).

have not conceptualised reflexive agency (different from a definition) which can make it hard to do systematic and cumulative research.

The commonness in studies on PISA amongst school leaders is that leaders need to have a strategy with PISA (Eggen 2010; Bringeland 2015). PISA triggers discussions in schools about the extent of the test's usage. PISA also challenge school leaders work, questioning whether it is the PISA test or the national curriculum that should be of priority. Despite the fact that many day-to-day tasks of school leaders are not related to the issues raised by PISA, the PISA test has effects upon school leaders' identity by increasing input-out processes substantiated in learning pressure (Sjøberg 2014c). However, an important note by Eggen is that the school personnel are participating subjects with the PISA test, not only objects. Both Bringeland's and Eggen's studies acknowledge that PISA is a challenge for school personnel beyond the intended script. There is no particular solution to paradoxes that come with such tests (Eggen 2011).

One of the limitations in Bringeland's (2015) study, is that he does not define reflexive agency although mentioning primary and corporate agency. Hence, school personnel's real projects with PISA and their professional orientation becomes hidden. Eggen's definition of agency in her works on international tests is reduced to 'judgements' (Eggen 2010, 282), and 'the opportunity for engagement in the social world of learning' (Eggen 2011, 533). Eggen is either focused on evaluative judgments with PISA or focused on building agency for 'knowledge construction' and for 'democratisation' with tests. Both Bringeland's and Eggen's reductive conceptualisation of agency has left questions about agents properties and their ultimate concern(s) behind. In that sense, the orientation they take with PISA can be obscured and hard to identify. For instance, claiming that PISA is not a topic at school, having no focus on it, while attending PISA courses might seem like a contradiction (Bringeland 2015). And what does critical (research) (Eggen 2010, 2011), and resistance (Eggen 2011), mean compared to Archer's modes of reflexivity, and especially, meta-reflexivity; being subversive can be quite different from these two concepts.

In my own recent reanalysis of PISA, autonomous reflexivity was the dominant mode of school personnel, i.e., they are focused on results and outcomes through managing tasks. They are characterised as being individualistic individuals, rather than conformists. Secondly, meta-reflexivity was the next mode of prominence, i.e., having critical and ethical issues with the test. A prime example of disinterest in PISA was the mathematics teacher in article II who

stated that he and his colleagues could interpret PISA results, but not all colleagues cared about PISA. Professional values restrained to some extent the further use of the test due to its design. However, orientation to values should be noted as a second concern for the school personnel, not as the primary concern, which was orientation to performing tasks, indicated by conducting the test using the standard script. Hence, PISA had effects on agency and informed about school personnel's' ultimate concern(s) in relation to PISA. There could be argumentation suggesting that, until we see some social movements against PISA in schools, there might be little possibility that autonomous reflexivity will change as the dominant mode in relation to the PISA phenomenon.

In the two articles, reflexivity and reflexive modes contributed to clarify the orientation and concerns of the school personnel with PISA. The order of the reflexive modes in this thesis, autonomous and meta-reflexivity, may be prompted by the Norwegian educational system characteristics. This is because a pre-dated educational system with its mechanisms triggers different concerns and ways of operating in school that can be of more supremacy than others. Another hypothesis could be that autonomous and meta-reflexivity are the dominant modes with PISA. However, research do indicate that PISA tasks, PISA data and PISA frameworks are used in education (Giberti and Maffia 2020). This can challenge such a hypothesis.

6.4 The key factors and variables

The key factors and variables in this thesis were structure and agency and their emergent properties. However, culture can't be left out of the equation, as ideas and materials are what the agent mediates. Some ideas and materials will be more dominating than others when producing an effect due to various of reasons. The same can apply for established effects when they revisit, but they can be challenged by the agent itself or by others, making new internal conversations, causing possible change in the factors' variables. For instance, PISA being a part of the Norwegian educational structure has effects on agents filtering and acting, they can for example subscribe to PISA as intended, reject PISA or be innovative with PISA. Such operations will have effects on their reflexive modes. The key factor agency; translated to a teacher and school leaders, can be critiqued for little relevance for this thesis due to little political impact. Primary agency, such as teachers and school leaders do not have the causal powers to negotiate changes at the macro-level. They need to address their aims to union representatives. However, they can be creative subjects with PISA (Eggen 2010), that can include opposition or fellowship with the test.

6.5 Summary

This research was a specific reply to Pons (2017) where case studies were missing key factors and variables that could explain why and how PISA effects occur or not. This thesis has used Archer's social realist work, which is underpinned by critical realism, as a gateway for further enabling a reconceptualisation of PISA effects.

The reason for a reconceptualisation on the educational system was the denotation 'PISA effects', which is used by numerous researchers studying the role played by PISA ideas and results in producing educational development and change, professed as more decentralization or recentralization. Previous studies on 'PISA effects' understand decisions on the local level and distributed responsibility transferred to the local level as decentralization. Some sees reporting achievements as recentralization. This is the decentralization/centralization paradox. Hence, decentralization and recentralization have been advocated simultaneously within the same period. The PISA researchers illuminated in this thesis and their discussed claims about 'PISA effects', lack a systemic theory for understanding the relative predominance of the various internal processes in the educational system and thereby mapping its character as centralized or decentralized. Archer's theoretical approach as enabled this through a systemic theory of the educational system, that enabled discussing changes in the internal structure of the educational system before and after introduction of PISA. In that way, previous scholarly claims about the system and 'PISA effects' become less substantiated because their claims are not built on systemic emergence, which is a fundamental premise in Archer's theorising on the educational system. Based on my own research, using the systemic theory of Archer, I have argued that PISA has strengthened processes of centralization by promoting, or strengthening unification. The test activates already established macro structures and processes. Hence, the Norwegian educational system was already centralized prior to and after the introduction of PISA.

As a side effect of theorising on an educational system which includes agency, is that a reconceptualisation on PISA effects on incumbents also happened through this thesis. PISA effects on the micro-level is mainly devoted to effects on school leaders and teachers work. But incorporating causality as a necessary connection entailed a reconceptualisation of this understanding. Mainly, there was an awakening of a necessary connection between PISA and reflexivity. To realise such a connection, one need to avoid conflation of agency and ask counterfactual questions. The same applies for structure. This entails acknowledging the

autonomy and causal power of structure and agency. Thus, previous studies on school personnel's understanding and use of PISA have contributed with findings and discussion related to their thoughts, decisions, and action with the PISA test. The limitation of these previous studies is that they lack reflexivity as a concept, i.e., the internal systemic structure of the agent, which precludes investigating how further actions with the test fit in with the individual's ultimate concerns which will influence their further actions with the test. Applying reflexivity and reflexive modes in this thesis, entailed the identification of autonomous reflexivity as the predominant reflexive mode. Meta-reflexivity was the second dominant mode. For instance, attending PISA seminars/courses – which was discussed in article III by two school leaders – is an indication of curiosity and the desire for 'know-how' about how to use the test. This means seeing the test more as a task which should be successfully accomplished than as an item that should be critically evaluated. Suggestively, the order of these two dominant reflexive modes may be prompted by the Norwegian educational system being centralized.

PISA has had effects, although activated by agency, on the structure of the Norwegian educational system; the mechanism (variable) unification. It also had effects on reflexive agency: the mechanisms (variables) autonomous reflexivity and meta-reflexivity. These are the deep effects on the internal structure of the Norwegian educational system and on agency in this thesis. These findings can develop hypotheses for further research that either can be validated or rejected. For instance, that autonomous and meta-reflexivity are the two dominant modes with PISA.

7. The thesis' contributions to the PISA research

7.1 Reconceptualising 'PISA effects'

The reconceptualisation of PISA effects, that is where effects are situated, had effects on two levels, due to how causality is understood in critical realism and the work of Margaret Archer. PISA effects can therefore be established on the macro-level and the micro-level of the educational system. But 'PISA effects' at the macro- and micro-levels of education are entirely different things. Thus, a reconceptualisation of PISA effects, must include a clear distinction between effects at the macro and micro-levels. Causes and effects at the macro (systemic) level are different in nature from causes and effects at the micro-level of teaching staff.

7.1.1 Macro-level 'PISA effects'

In Norway, PISA is a part of the National Quality Assessment System (NQAS), which is under the mechanism unification in the Norwegian educational system. PISA effects at the macro-level concern how PISA results are used in decision-making at the national level which may have further consequences for the social form of the educational system. International research tells us that PISA results are used to legitimise current educational policy, as well as justifying new policy. Such findings are abundant in international research on 'PISA effects', and some of them are presented in my article I.

Another feature of previous studies of 'PISA effects' is that some researchers refer to the sequence of events when they claim that PISA has caused macro-level policy, such as national reforms. Thus, policies initiated after the public announcement of PISA results is seen as a 'PISA-effect'. However, to be fair, in such cases the succession of events is not the only criterion for attributing 'PISA effects', another criterion is the nature of the policy change, whether it concerns issues that have figured in the public PISA debate.

At the macro-level, PISA results are information inputs for policy decisions, which could be the decision to implement a reform meant to raise student achievement in mathematics, or it could be a decision to continue current educational policies. It should also be mentioned that among the studied *causes* of macro-level 'PISA effects' are not just the plain figures of PISA results or ranking in the international league table, but also the way these results have been received in the public domain, especially mass-media producing 'PISA shocks'.

7.1.2 Micro-level ‘PISA effects’

At the micro-level of teachers and school leaders, PISA impacts are more indirect. Decisions made at the central level, which have been informed by PISA results, have consequences for agents’ structural context at the micro-level. Central level decisions may for example, like in Norway, have included PISA in the National Quality Assessment System (NQAS), or made it part of some other national accountability scheme. Such decisions have consequences at the micro-level, conditioning agents’ actions, and even the way they see themselves.

The PISA test being incorporated in the NQAS in Norway means that PISA scores are seen as indicators of quality of the system. It also entails that teachers must administer the test if their students are in the national random PISA sample; through their school, teachers and school leaders receive information and feedback about national PISA results; and they are expected to learn from the test results to improve their teaching practice (assessment for learning). This means that PISA in various ways has become part of the internal structures of the educational system. The school staff is to some extent forced to relate to PISA and PISA results.

However, at the individual level the structural impact of PISA on school staff is filtered through each individual’s reflexive modes, as suggested by Archer, and demonstrated in articles II and III, making effects on their reflexivity.

The Norwegian researcher Astrid Birgitte Eggen (2010, 290), reporting from discussions among leaders on how public debates on PISA results affect them, render that leaders ask themselves: ‘are we measuring up?’ Another example, speaking on behalf of her students, one teacher claims ‘I think it is unfair, that the school pretends they will not be tested by PISA! We teachers must teach the kids the stuff in which they are tested on their exams. You cannot give an exam which tests something different from what they have been taught at school.’ (Eggen 2010, 291). As will be recalled, PISA aims to be culturally neutral and is therefore a non-curricular test (in reading, mathematics, and science).

7.1.3 Macro-level and micro-level ‘PISA effects’

Macro-level ‘PISA effects’ differ from micro-level effects. Macro-level effects consist in decisions to change or maintain the direction of national educational policy with consequences for the structure of the educational system. Micro-level effects consist in teachers and school leaders concerns and actions with PISA and relating them to the micro-level effects of macro-level decisions. In addition, comes the pressures from the local school

environment associated with the general public's interest in PISA. The environmental pressure may, as Eggen (2010) has shown, result in ponderings over what consequences the PISA results should have for teaching practice, or school leadership, and even how teachers and school leaders see themselves as professionals. A reconceptualisation of 'PISA effects' must therefore include a clear distinction between macro and micro-level PISA effects. Effects at the systemic level are different in nature from effects at the micro-level, though the two types of effect are connected in the sense that macro effects may in turn produce micro effects.

Fortunately, Archer's conceptual framework includes both macro and micro phenomena and these different levels of theorising are coherent, i.e., rests on the same fundamental principles. Her approach and perspective can therefore accommodate the study of PISA effects both at the micro-level and macro-level, as I have tried to demonstrate in my articles.

7.2 Empirical contributions

In Norway, 'PISA effects' as a research topic has a voluminous literature, especially related to the macro-level of national policy and the educational system. Far fewer studies focus on the micro-level of the school staff, and none to the meso-level of the municipalities, who in Norway are the school owners.⁵⁴ The present project has studied 'PISA effects' both at the macro and micro-level, which is an example of analytical dualism.

This thesis' empirical contribution to the study of 'PISA effects' at the macro-level of the Norwegian educational system was guided by Archer's definition of internal processes in educational systems. My finding was that processes which predominated before the Norwegian 'PISA shock' in the early 2000 continued to predominate after the 'shock' (see article I). According to Archer's definition of centralized systems, the Norwegian system was centralized both before and after the 'PISA shock'. Østerud (2016, 16) claims the 'PISA shock' have been 'a turning point for Norwegian educational policy'. In Archer's terms not a

⁵⁴ I have not been able to find any Norwegian research on the reception and use of PISA results at the municipal level, i.e., among the municipalities, the 'school owners'. This is surprising, since the implicit recipients of the Norwegian National Curriculum, since 2006, has been shifted from 'the individual teacher' (in the 1974 National curriculum) and 'the teaching profession in general' (in the 1987 curriculum) to 'the municipal school owners' (in the 2006 curriculum) (Engelsen 2008). These shifts indicate that, at present, the meso-level of the municipalities is considered the locus of accountability for educational outcomes.

turning point for the outcomes of policy in terms of systemic change. However, PISA has had effects on the social form of educational system i.e., unification, with side effects on other internal mechanisms.

My second empirical contribution to PISA research is that I have given voice to a lower secondary teacher and school leaders—at the grade level that is targeted by the PISA test. In previous PISA research, teachers' and school leaders' voices are underrepresented.

A third empirical finding and contribution – at the micro-level – is that the PISA test did appear to have effects on agents (teachers) in the sense that PISA was (and probably is) an object of internal conversations and deliberations among teachers and school leaders. The empirical data shed light on how agents perceive PISA as either compatible or incompatible with their personal commitments. This entails that PISA has effects on individuals' reflexivity and reflexive modes. For this specific research, PISA had especially effects on autonomous and meta-reflexivity.

7.3 Methodological contributions

In the present project one of the articles (art. I) is based on documents studies which are accommodated into a limited morphogenetic cycle, analysed with social realist conceptualisations in the Norwegian context. Two of the studies are based on qualitative semi-interview structured data (article II and III). The study reported in article II employs mixed methods; reflexive modes are both investigated through the analysis of qualitative interview data and through the administration of a short questionnaire with Likert scale items, ICONI (Internal Conversation Indicator). ICONI measures the relative predominance of different modes of reflexivity by counting frequencies of occurrence and calculating mean scores, thus, profiles of individuals' reflexive modes may be constructed (most individuals practice several reflexive modes but to various extents). However, conducting interviews are not new to the PISA effects field, nor is having school leaders and teachers as respondents. But interviews with a mathematics teacher in combination with social realist⁵⁵ ICONI is new to the PISA effects field. Also, reanalysis of qualitative interviews with school leaders on PISA with social realist conceptualisations is new to the field. Thus, the methodological

⁵⁵ Archer has preferred to call her own theoretical approach 'social realist', although she also considered herself a critical realist. She particularly appreciated Roy Bhaskar's book 'The possibility of naturalism', which she considers one of the major contributions to social theory during the 20th century.

contribution to the PISA effects field is found in all three articles, which representatively have used the social realist morphogenetic cycle with PISA in the Norwegian context, reanalysis of qualitative interview with school leaders on PISA with social realist conceptualisations in the Norwegian context, and a mathematics teacher interview in combination with the social realist ICONI in the Norwegian context.

Another methodological contribution to the PISA effects field is posing counterfactual questions. This is seldom a part of the training offered PhD-students and other researchers attending courses in research methods. But asking such questions is a basic feature of critical realist research, fueled by an interest in the preconditions of phenomena and asking questions such as: what must be the case for X to exist? Discussing my findings in article I, I employ counterfactual questions concerning the introduction of PISA in Norway and our continued participation in the test, dispelling possible blame that could be addressed to the Minister of Education at the time or other persons who made the decision on Norway's initial and continued participation in the PISA test. Counterfactual questions could perhaps have been asked more frequently in this project. But it is particularly apt when searching for structures which (co)-determine action. What would happen if one element, a structure in society was eliminated (in thought)? I will argue that through my research I have at least given my readers a flavour of the power that lies in asking counterfactual questions, thus practicing retrodution (tracing generative structures), which is part of the repertoire of methods that critical realism particularly endorses.

7.4 Theoretical contributions

Previous PISA research within the field of 'PISA effects' lacks overarching conceptualisations and has not been established as a 'normal science', enabling cumulative research (Pons 2017). Pons points out that PISA research has been scattered across different research environments with different research traditions, representing different scientific disciplines and theoretical preferences. Pons wants a common/cohesive/coordinated research environment that could gather around common theoretical approaches so that the research could be cumulative: i.e., new research would build on and further develop previous research. Pons also critiques previous PISA research for having too much focus on the 'PISA shock' in different parts of the world. This thesis is a start for the cumulative research tradition that Pons recommends.

Critical realism is almost absent in the PISA effects field, with one study mentioning Roy Bhaskar (Nash 2005)—but not with qualitative analysis. Thus, I consider the application of critical realism in this thesis as a contribution to the PISA effects field. Archer's theoretical framework is a contribution to the PISA effects field since it is absent. In the present project on 'PISA effects', Archer's conceptual framework has enabled the distinction between four levels of theory: the meta level (critical realism), the general level (SAC⁵⁶), the level of domains (educational systems and reflexive agency) and the specific level (in the present context, the Norwegian educational system and school staff).

Archer's methodology *analytical dualism* has contributed to a better understanding of the relationship between structure and agency, and culture and agency. Analytical dualism recognises that the agent and social structure/culture belong to different strata in social reality, they carry distinctive mechanisms and powers. If structure and agency are conflated, i.e., amalgamated, it will be impossible to study how they are impacted by each other. Therefore, in all empirical investigations, individuals, social structures, and culture must be studied separately and not treated as one unit, that is, conflated (Archer 1995, 165ff).

In previous chapters of this text and in my articles, I have outlined and documented the contribution that critical realism and Archer's various conceptual frameworks can make to improve research on the reception and use of the PISA test. I have presented and applied the morphogenetic approach; the model of morphogenetic cycles and conceptions of social structure, agency, and culture (SAC), as well as mediating mechanisms between structure and agency, that is agents' reflexivity, and reflexive modes. Archer has rooted her theoretical conceptions and models in a philosophy of science, critical realism, which means that formulating a sustainable critique requires addressing her basic premises. Her conceptual framework *applies to all social domains*, not just education.

7.5 Strengths and limitations of the research

I have succeeded in pinpointing macro mechanisms in the one educational system I studied, and I was also able to identify reflexive modes among the school personnel. However, the historical analysis presented in article I is limited to a specific period so further investigations on the educational system characterising structures should be prompted. I have used Skinningsrud (2019) for support for the characterising education system after 2010, but this

⁵⁶ SAC stands for structure, agency, and culture.

publication is limited in scoping of the educational system too. The analysis of reflexive modes in this thesis are limited to a few numbers of cases, so further investigations on reflexive modes are needed with PISA to establish if these modes identified are the ones that are mostly dominant beyond these cases. In general the number of respondents between social strata i.e., from different levels of the educational system can also be increased in further studies. One circumstance which limited my number of cases was that I had difficulties recruiting respondents. There was little interest in PISA. I could possibly have provided perspectives from a Norwegian subject teacher or Norwegian science teacher, as these subjects are PISA domains for testing, but I did not want to embrace all aspects with PISA at once, for comparisons. This project was foremost intended as a start for doing systematic and cumulative research within a paradigm. Neither was time a resource for making further interviews.

Using other research designs and methods was suggested to me when I planned the project. Conducting a survey, which was suggested by some, was, however, not an option. Doing a survey, I would be unable to investigate how reflexive modes were activated in concrete contexts. Surveys from the micro-level (and macro-level) with the use of PISA are scarce. Ethnographic observational studies are also scarce in the PISA effects field. Doing an ethnographic study using participant observation could involve discovery of other PISA effects which the respondents could have forgotten mentioning in the interviews. However, I wanted to be focused on what was being articulated and not being distracted by the surroundings. I also had limited time. For these reasons, I conducted the interview with the mathematics teacher reported in article II via Skype without the video function on. However, with interviews, I have no guarantee that the social-desirability bias affected the way the respondents reported their relationship with PISA. Focus group interviews were not an option as it would be hard to identify each reflexive modes. However, a group of respondents, gathered in focus groups, could have validated statements and pinpointed other PISA effects that operate in schools. However, in focus groups there is a risk that one respondent becomes too dominant in the conversation, meaning that the others become passive which could have consequences for the analysis. However, focus groups interviews in the PISA effects field is scarce.

This research has not focused on OECD's historical perspectives with the respondents, nor that PISA is a neutral non-curricular test and entangled with edu-business. These are design limitations that further research can include. Initiatives such as TALIS (PISA for teachers and

school leaders working conditions and learning environments at their schools), SSES (Study of Social and Emotional Skills - PISA for 10- and 15-year-olds), PISA for Schools (PISA local school test on demand) and PISA4U (online learning modules and collaborative activities on PISA data for teachers), and the Future of Knowledge and Skills 2030 project for curriculum standards has neither been included in the interview guide. However, this inclusion would depend on whether it's the main PISA test you are going to focus on or not. If data on these arise in the interview, one can ask whether these products have changed the view on the main PISA test, for comparison.

Since it has been difficult to grasp Archer's theories from an early start, this has impacted the research process. I have not been able to do a survey as her conceptualisations needed to be understood correctly before I embark on such a method. Since Archer's conceptual framework is new to the PISA effect field there are multiple of ways to do research with her conceptualisations. However, systematic and cumulative research do imply that the methods used in this thesis should accumulate more similar research. The focus and limitation in this thesis has had implications for the mentioned contribution of this research. As this research has investigated a historical period of the characteristics of the educational system where PISA is situated and pinpointed some indications and hypothesis of reflexive modes, this can be viewed as a start to validate or reject further hypothesis with the PISA phenomenon, for example replication (re-checking) for does who might disagree, new time-period of investigation for finding the characterising educational system where PISA is situated and/or new respondents reflexive modes with PISA. Altogether, the conceptual framework from Archer contributes to filling empirical, methodological and theoretical gaps in the PISA effects field, which provides a new contribution to the PISA effects field. Since the empirical data on reflexive modes is limited, more studies need to be done to capture the dominant reflexive modes with PISA beyond these cases presented here. The question is if generalisation will ever be possible with the reflexive modes, however, random samples might provide stronger indications.

7.6 Ideas for further research

Since I have applied a new conceptual framework to the PISA effects field; this has consequences for new undertakings with PISA. I can either discuss previous research (reanalysis) or embark on new empirical collections. Since there are limited publications from the micro-level and with open access to the data, I will concentrate on ideas that are grounded

in discussing macro-level PISA research and the collection of new data that are systematic and cumulative and based on this conceptual framework.

An interesting application of Archer's conceptualisations of educational systems is to examine the structures and processes going on in other countries' educational systems before and after 'major PISA events' to examine the possible impacts of such events. It would also be interesting to add updated voices on PISA (longitudinal studies), for example, expanding my previous work on reflexivity. Reflexivity and reflexive modes could also be applied with several of other stakeholders with PISA. For instance, at the school level (also with ethnographic studies in schools, with participant observation of events, such as when the PISA test is administered, or when the PISA results are published. This would require meetings that are scheduled prior to the publication of the results (and prior to the interviews). Furthermore, interviews with the meso-level (municipality officials/school owner) and macro-level (Directorate of Education and the Ministry of Education). Interviews could also be conducted with Union representatives in schools, municipalities, counties and on the national level, voicing their commitment to PISA. Also, interviews with researchers at the University of Oslo, who organise the PISA test at the national level, analyse PISA data, and organise PISA seminars/courses for staff members in lower secondary schools, could be of interest. In some cases, focus group interviews could be a strategy for validating and discussing (other) PISA effects. From a theoretical perspective, Archer's cultural situational logics connected to complementarities and contradictions would be interesting to apply in further research on PISA. Updated and new surveys amongst school personnel can be of interest for comparisons.

7.7 Summary

In this chapter I have argued that the reconceptualisation of 'PISA effects' has had consequences for two levels: the macro and micro-level. The macro-level effects are decision-making that have consequences for the structure of the educational system. The micro-level effects emerge from macro-level decisions. Hence, causes and effects on the macro-level of decision-making on the system are different from causes and effects at the micro-level of school personnel.

Empirical contributions from this research are *theory generation on the characterising Norwegian educational system* before and after the introduction PISA, and theorising and *illuminating voices from the micro-level* and establishing that there are *deep PISA effects on the macro- and micro-level*. Methodological contributions to the field of PISA effects from

the Norwegian context is the use of Archer's *morphogenetic cycle*, Archer's conceptualisations for *reanalysing* interview data with school leaders on PISA, and a mathematics teacher interview in combination with Archer's *ICONI*. This research has also applied *counterfactual questions* that critical realism support for reclaiming reality. Lastly, the theoretical contribution for the PISA effects field is the use of *critical realism as the meta-theory* for starting cumulative research within a paradigm, which is further supported by *Archer's conceptual framework*.

The strength of this research is the use of analytical dualism: identifying generative mechanisms in the educational system and in incumbents where deep PISA effects are situated. However, the analysis of the charactering educational system is limited to a historical period. Further analysis should be prompted. There is no possibility for statistical generalisation of reflexive modes. Only hypothesis. The interview guide can also be further optimised for grasping the PISA phenomenon better. Further research, based on Archer's conceptual framework, can add analysis of PISA's impact on other education systems. Further research can also continue interviews with teachers and school leaders, and include a broader variety of respondents, and observations and surveys, for different comparisons.

Works cited

- Adams, Tracey, and Peter H. Sawchuk. 2020. "Professional-Organizational Contradictions and Hybridization of Knowledge: Insights from the Study of Engineering and Nursing in Canada." *Vocations and Learning* 14 (1):75–93. doi: 10.1007/s12186-020-09253-1.
- Adamson, Bob, Katherine Forestier, Paul Morris, and Christine Han. 2017. "PISA, Policymaking and Political Pantomime: Education Policy Referencing between England and Hong Kong." *Comparative Education* 53 (2):192–208. doi: 10.1080/03050068.2017.1294666.
- Addey, Camilla. 2020. "The Appeal of PISA for Development in Ecuador and Paraguay: Theorising and Applying the Global Ritual of Belonging." *Compare: A Journal of Comparative and International Education* 50 (8):1159–1174. doi: 10.1080/03057925.2019.1623653.
- Aini, Indrie Noor, Zulkardi, Ratu Ilma Indra Putri, and Poppy Yaniawati. 2023. "Developing PISA-Like Math Problems in the Content of Space and Shape Through the Context of Historical Buildings." *Journal on Mathematics Education* 13 (4):723–738. doi: 10.22342/jme.v13i4.pp723-738.
- Al-Amoudi, Ismael. 2007. "Redrawing Foucault's Social Ontology." *Organization* 14 (4):543–563. doi: 10.1177/1350508407078052.
- Al-Amoudi, Ismael, and Hugh Willmott. 2011. "Where Constructionism and Critical Realism Converge: Interrogating the Domain of Epistemological Relativism." *Organization Studies* 32 (1):27–46. doi: 10.1177/0170840610394293.
- Andersson, Christian, and Sofia Sandgren Massih. 2023. "PISA 2018: Did Sweden Exclude Students According to the Rules?" *Assessment in education: Principles, Policy & Practice* 30 (1):33–52. doi: 10.1080/0969594X.2023.2189566.
- Andrews, Therese. 2021. "Bourdieu's Theory of Practice and the OECD PISA Global Competence Framework." *Journal of Research in International Education* 20 (2):154–170. doi: 10.1177/14752409211032525.
- Archer, Margaret. 1981. "On Predicting the Behaviour of the Educational System." *British Journal of Sociology of Education* 2 (2):211–219. doi: 10.1080/0142569810020206.
- Archer, Margaret S. 1983. "Process without System." *European Journal of Sociology* 24 (1):196–221.
- Archer, Margaret S. 1984. *The University Edition of Social Origins of Educational Systems*. London: Sage Publications.
- Archer, Margaret S. 1995. *Realist Social Theory: The Morphogenetic Approach*. Cambridge: Cambridge University Press.
- Archer, Margaret S. 2000. *Being Human: The Problem of Agency*. Cambridge: Cambridge University Press.
- Archer, Margaret S. 2003. *Structure, Agency and the Internal Conversation*. Cambridge: Cambridge University Press.
- Archer, Margaret S. 2007. *Making Our Way through the World: Human Reflexivity and Social Mobility*. Cambridge: Cambridge University Press.
- Archer, Margaret S. 2012. *The Reflexive Imperative in Late Modernity*. Cambridge: Cambridge University Press.
- Archer, Margaret S. [1979] 2013. *Social Origins of Educational Systems*. London: Sage.
- Archer, Margaret S. [1988] 1996. *Culture and Agency: The Place of Culture in Social Theory*. Cambridge: Cambridge University Press.
- Aursand, Leah Rose. 2018. "What [Some] Students Know and Can Do: A Case Study of Norway, PISA, and Exclusion." Master's Thesis, Department of Education, University of Oslo, Oslo, Norway.
- Aursand, Leah Rose, and David Rutkowski. 2021. "Exemption or Exclusion? A study of Student Exclusion in PISA in Norway." *Nordic Journal of Studies in Educational Policy* 7 (1):16–29. doi: 10.1080/20020317.2020.1856314.
- Baird, Jo-Anne, Sandra Johnson, Therese N. Hopfenbeck, Talia Isaacs, Terra Sprague, Gordon Stobart, and Guoxing Yu. 2016. "On the Supranational Spell of PISA in Policy." *Educational Research* 58 (2):121–138. doi: 10.1080/00131881.2016.1165410.
- Ball, Stephen J. 2003. "The Teacher's Soul and the Terrors of Performativity." *Journal of Education Policy* 18 (2):215–228. doi: 10.1080/0268093022000043065.

- Ball, Stephen J. 2012. *Global Education Inc : New Policy Networks and the Neo-Liberal Imaginary*. Abingdon, Oxon and New York: Routledge.
- Ball, Stephen J. 2016. "Subjectivity as a Site of Struggle: Refusing Neoliberalism?" *British Journal of Sociology of Education* 37 (8):1129–1146. doi: 10.1080/01425692.2015.1044072.
- Bergesen, Helge Ole. 2006. *Kampen om kunnskapsskolen [The Battle for the Knowledge-School]*. Oslo: Universitetsforlaget.
- Bhaskar, Roy. 2011. *Reclaiming Reality: A Critical Introduction to Contemporary Philosophy*. London: Routledge.
- Bhaskar, Roy. 2016. *Enlightened Common Sense: The Philosophy of Critical Realism*. London: Routledge.
- Bhaskar, Roy. [1975] 2008. *A Realist Theory of Science*. London: Routledge.
- Bhaskar, Roy. [1979] 1998. *The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences*. London: Routledge.
- Bhaskar, Roy, and Berth Danermark. 2006. "Metatheory, Interdisciplinarity and Disability Research: A Critical Realist Perspective." *Scandinavian Journal of Disability Research* 8 (4):278-297. doi: 10.1080/15017410600914329.
- Bieber, Tonia, and Kerstin Martens. 2011. "The OECD PISA Study as a Soft Power in Education? Lessons from Switzerland and the US." *European Journal of Education* 46 (1):101-116.
- Biesta, G. 2009. "Good Education in an Age of Measurement: On the Need to Reconnect with the Question of Purpose in Education." *Educational Assessment, Evaluation and Accountability* 21 (1):33-46.
- Biesta, Gert J. J. 2010. *Good Education in an Age of Measurement: Ethics, Politics, Democracy, Interventions: Education, Philosophy & Culture*. London: Routledge.
- Biesta, Gert, Mark Priestley, and Sarah Robinson. 2015. "The Role of Beliefs in Teacher Agency." *Teachers and Teaching, Theory and Practice* 21 (6):624–640. doi: 10.1080/13540602.2015.1044325.
- Billett, Stephen, and Margarita Pavlova. 2005. "Learning Through Working Life: Self and Individuals' Agentic Action." *International Journal of Lifelong Education* 24 (3):195–211. doi: 10.1080/02601370500134891.
- Björnsson, Julius Kristjan, and Rolf Vegar Olsen. 2018a. "Tjue år med internasjonale skoleundersøkelser i Norge: Bakgrunn, læringspunkter og veien videre [Twenty Years of International School Surveys in Norway: Background, Learning Points and the Way Forward]." In *Tjue år med TIMSS og PISA i Norge : trender og nye analyser [Twenty Years with TIMSS and PISA in Norway: Trends and New Analyses]*, edited by Julius Kristjan Björnsson and Rolf Vegar Olsen, 12–34. Oslo: Universitetsforlaget.
- Björnsson, Julius Kristjan, and Rolf Vegar Olsen. 2018b. *Tjue år med TIMSS og PISA i Norge : trender og nye analyser [Twenty Years with TIMSS and PISA in Norway: Trends and New Analyses]*. Oslo: Universitetsforlaget.
- Blom, Björn, and Stefan Morén. 2011. "Analysis of Generative Mechanisms." *Journal of Critical Realism* 10 (1):60–79. doi: 10.1558/jcr.v10i1.60.
- Breakspear, Simon. 2012. The Policy Impact of PISA: An Exploration of the Normative Effects of International Benchmarking in School System Performance. *OECD Education Working Papers* (71): 31 pages.
- Bringeland, Terje André. 2015. "PISA: Et styringsinstrument for og i skoler i relasjon til matematikkopplæringen? [PISA: A Management Instrument for and in Schools in Relation to Mathematics Education?]." Master's Thesis, Bergen University College.
- Bringeland, Terje André. 2022a. *The Impact of PISA on Education in Norway: A Morphogenetic Perspective on Structural Elaboration in an Education System*. Edited by Margaret S. Archer, Unn-Doris K. Bæck and Tone Skinningsrud, *The Morphogenesis of the Norwegian Educational System: Emergence and Development from a Critical Realist Perspective*. Milton: Taylor & Francis Group.
- Bringeland, Terje André. 2022b. "The Impact of PISA on Education in Norway: A Morphogenetic Perspective on Structural Elaboration in an Education System." In *The Morphogenesis of the Norwegian Educational System: Emergence and Development from a Critical Realist Perspective*, ed Margaret S. Archer, Unn-Doris K. Bæck and Tone Skinningsrud. Milton:

- Taylor & Francis Group.
<https://www.taylorfrancis.com/books/edit/10.4324/9781003163527/morphogenesis-norwegian-educational-system-margaret-archer-unn-doris-b%C3%A6ck-tone-skinningrud>.
- Bringeland, Terje André. 2022c. "School Leaders' Reflexive Mode in their Internal Conversations on PISA." In *The Morphogenesis of the Norwegian Educational System: Emergence and Development from a Critical Realist Perspective*, edited by Margaret S. Archer, Unn-Doris K. Bæck and Tone Skinningsrud, 181–210. Milton: Taylor & Francis Group.
- Bringeland, Terje André, and Tone Skinningsrud. 2023. PISA and Teachers' Reflexivities. A Mixed Methods Case Study. Unpublished manuscript, last modified November 27, 2023. Portable Document Format.
- Buch-Hansen, Hubert, and Peter Nielsen. 2005. *Kritisk realisme [Critical Realism], Problemer, positioner og paradigmer [Problems, Positions, and Paradigms]*. Frederiksberg: Roskilde Universitetsforlaget.
- Bürgi, Regula. 2019. "Learning Productivity: The European Productivity Agency—An Educational Enterprise." In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 17–37. Cham: Palgrave Macmillan.
- Camphuijsen, Marjolein K., Jorunn Møller, and Guri Skedsmo. 2021. "Test-based Accountability in the Norwegian Context: Exploring Drivers, Expectations and Strategies." *Journal of Education Policy* 36 (5):624–642. doi: 10.1080/02680939.2020.1739337.
- Centeno, Vera G. 2019. "The Birth of the OECD's Education Policy Area." In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 63–82. Cham: Palgrave Macmillan.
- Cruikshank, Justin. 2003. *Realism and Sociology: Anti-Foundationalism, Ontology and Social Research*. Vol. 5. London & New York: Routledge.
- Dalland, Cecilie Pedersen. 2011. "Utfordringer ved gjenbruk av andres kvalitative data [Challenges in Reusing Other People's Qualitative Data]." *Norsk pedagogisk tidsskrift* 95 (6):449–459.
- Danermark, Berth, Mats Ekström, Liselotte Jakobsen, and Jan Ch. Karlsson. 2002. *Explaining Society: Critical Realism in the Social Sciences*. 1st ed. London: Routledge.
- Danermark, Berth, Mats Ekström, Liselotte Jakobsen, and Jan Ch. Karlsson. 2003. *Att förklara samhället [Explaining Society]*. Lund: Studentlitteratur.
- Dilekçi, Atilla. 2022. "Evaluation of Turkey's PISA Reading Literacy Scores." *International Journal of Education and Literacy Studies* 10 (1):138. doi: 10.7575/aiac.ijels.v.10n.1p.138.
- Djupedal, Elise Farstad. 2022. "PISA-politikk på timeplanen. Vil politikerne fjerne PISA-testen eller politikken som ble utformet i kjølvannet av den? [PISA policy on the timetable. Will the Politicians Remove the PISA test or the Policies that were Designed in its Wake?]." *Dagsavisen*, Last Modified 01.09.2022, accessed 30.08.2022.
<https://www.dagsavisen.no/debatt/2022/09/01/pisa-politikk-pa-timeplanen/>.
- Donati, Pierpaolo, and Margaret S Archer. 2015. *The Relational Subject*. Cambridge: Cambridge University Press.
- Eggen, Astrid Birgitte. 2010. "PISAs gyldighet blant skoleledere og lærere [PISA's Validity among School Leaders and Teachers]." In *PISA: sannheten om skolen? [PISA: The Truth About the School?]*, edited by Kirsten Sivesind and Eyvind Elstad, 281–297. Oslo: Universitetsforlaget.
- Eggen, Astrid Birgitte. 2011. "Agency as the Ability and Opportunity to Participate in Evaluation as Knowledge Construction." *European Educational Research Journal* 10 (4):533–544. doi: 10.2304/eeerj.2011.10.4.533.
- Elfert, Maren. 2019. "The OECD, American Power and the Rise of the "Economics of Education" in the 1960s." In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 39–61. Cham: Palgrave Macmillan.
- Elstad, Eyvind. 2010. "Internasjonale storskalaundersøkelser: en sammenlikning av PISA, TIMSS og PIRLS [International Large-Scale Surveys: A Comparison of PISA, TIMSS and PIRLS]." In *PISA: sannheten om skolen? [PISA: The Truth About School?]*, edited by Eyvind Elstad and Kirsten Sivesind, 54–66. Oslo: Universitetsforlaget.
- Elstad, Eyvind, and Kirsten Sivesind. 2010. "OECD setter dagsorden [The OECD Sets the Agenda]." In *PISA: sannheten om skolen? [PISA: The Truth About School?]*, edited by Eyvind Elstad and Kirsten Sivesind, 20–41. Oslo: Universitetsforlaget.

- Engelsen, Britt Ulstrup. 2008. *Kunnskapsløftet : sentrale styringssignaler og lokale strategidokumenter [The Knowledge Promotion: Key Management Signals and Local Strategy Documents]*. Oslo: Pedagogisk Forskningsinstitutt.
- Eteläpelto, Anneli, Katja Vähäsantanen, Päivi Hökkä, and Susanna Paloniemi. 2013. "What is Agency? Conceptualizing Professional Agency at Work." *Educational Research Review* 10:45–65. doi: 10.1016/j.edurev.2013.05.001.
- Evetts, Julia. 2011. "A New Professionalism? Challenges and Opportunities." *Current Sociology* 59 (4):406–422. doi: 10.1177/0011392111402585.
- Frønes, Tove Stjern, and Fredrik Jensen. 2020. *Like muligheter til god leseforståelse? 20 år med lesing i PISA [Equal Opportunities for Good Reading Comprehension? 20 Years of Reading in PISA]*. Oslo: Universitetsforlaget.
- Giberti, Chiara, and Andrea Maffia. 2020. "Mathematics Educators are Speaking about PISA, aren't they?" *Teaching Mathematics and its Applications* 39 (4):266-280. doi: 10.1093/teamat/hrz018.
- Gjone, Gunnar. 2010. "Kompetanser for matematikkfaget i PISA og i norske læreplaner [Competences for the Mathematics Subject in PISA and in Norwegian Curricula]." In *PISA: sannheten om skolen? [PISA: The Truth About School?]*, edited by Eyvind Elstad and Kirsten Sivesind, 157–175. Oslo: Universitetsforlaget.
- Gorur, Radhika. 2011. "ANT on the PISA Trail: Following the Statistical Pursuit of Certainty." *Educational Philosophy and Theory* 43 (December 2011):76-93. doi: 10.1111/j.1469-5812.2009.00612.x.
- Grek, Sotiria. 2009. "Governing by Numbers: The PISA 'Effect' in Europe." *Journal of Education Policy* 24 (1):23-37.
- Grey, Sue, and Paul Morris. 2018. "PISA: Multiple 'Truths' and Mediatized Global Governance." *Comparative Education* 54 (2):109-131. doi: 10.1080/03050068.2018.1425243.
- Groff, Ruth. 2004. *Critical Realism, Post-positivism and the Possibility of Knowledge*. Vol. 11, *Routledge Studies in Critical Realism*. Abingdon, Oxon: Routledge.
- Groff, Ruth. 2013. *Ontology Revisited: Metaphysics in Social and Political Philosophy*. Abingdon, Oxon: Routledge.
- Gunnulfsen, Ann Elisabeth, and Jorunn Møller. 2017. "National Testing: Gains or Strains? School Leaders' Responses to Policy Demands." *Leadership and Policy in Schools* 16 (3):455-474.
- Heaton, Janet. 2008. "Secondary Analysis of Qualitative Data: An Overview." *Historical Social Research* 33 (3):33–45.
- Hopfenbeck, Therese N., and Marit Kjærnsli. 2016. "Students' Test Motivation in PISA: The Case of Norway." *The Curriculum Journal* 27 (3):406–422. doi: 10.1080/09585176.2016.1156004.
- Hopfenbeck, Therese N., Jenny Lenkeit, Yasmine El Masri, Kate Cantrell, Jeanne Ryan, and Jo-Anne Baird. 2018. "Lessons Learned from PISA: A Systematic Review of Peer-Reviewed Articles on the Programme for International Student Assessment." *Scandinavian Journal of Educational Research* 62 (3):333–353. doi: 10.1080/00313831.2016.1258726.
- Hopfenbeck, Therese Nerheim. 2010. "Elevstemmer fra PISA 2006 [Student Voices from PISA 2006]." In *PISA: sannheten om skolen? [PISA: the Truth about School?]*, edited by Eyvind Elstad and Kirsten Sivesind, 268–280. Oslo: Universitetsforlaget.
- Hopfenbeck, Therese, Astrid Tolo, María Flórez Petour, and Yasmine El Masri. 2013. *Balancing Trust and Accountability? The Assessment for Learning Programme in Norway. A Governing Complex Education Systems Case Study*. Paris: OECD.
- Hopmann, Stefan Thomas. 2008. "No child, No School, No State Left Behind: Schooling in the Age of Accountability." *Journal of Curriculum Studies* 40 (4):417-456. doi: 10.1080/00220270801989818.
- Hopmann, Stefan Thomas. 2007. "Introduction: PISA According to PISA – Does PISA Keep What It Promises?" In *PISA zufolge PISA - PISA According to PISA*, edited by Stefan Thomas Hopmann, Gertrude Brinek and Martin Retzl, 9–19. Berlin: Lit Verlag.
- Hornskov, Søren, Helle Bjerg, and Laura Høvsgaard. 2016. *Review: Brug af data i skoleledelse [Review: Use of Data in School Management]*. København: UCC.

- Hossain, Mobarak. 2023. "Perceptions of Key Education Actors towards PISA: The Case of Scotland." *Oxford Review of Education* ahead-of-print (ahead-of-print):1-18. doi: 10.1080/03054985.2023.2207817.
- ILS. n.d-a. "Om PISA [About PISA]." Institutt for lærerutdanning og skoleforskning (ILS) [Institute for Teacher Education and School Research], accessed 24.08.22. <https://www.uv.uio.no/ils/forskning/prosjekter/pisa/om-pisa/>.
- ILS. n.d-b. "PISA." accessed 18.09.2023. <https://www.uv.uio.no/ils/forskning/prosjekter/pisa/>.
- Imsen, Gunn, Ulf Blossing, and Leif Moos. 2017. "Reshaping the Nordic Education Model in an Era of Efficiency. Changes in the Comprehensive School Project in Denmark, Norway, and Sweden since the Millennium." *Scandinavian Journal of Educational Research* 61 (5):568-583. doi: 10.1080/00313831.2016.1172502.
- Jensen, Fredrik, Sonja Mork, and Marit Kjærnsli. 2018. "En sammenligning av naturfagkompetanser i PISA-rammeverket 2015 og den norske læreplanen [A Comparison of Science Competences in the PISA Framework 2015 and the Norwegian Curriculum]." In *Tjue år med TIMSS og PISA i Norge : trender og nye analyser [Twenty Years with TIMSS and PISA in Norway: Trends and New Analyses]*, edited by Rolf Vegar Olsen and Julius Kristjan Björnsson, 94–127. Oslo: Universitetsforlaget.
- Jiang, Tao, Ji-gen Chen, and Wei Fang. 2021. "The Generative Mechanism of Secondary School Students' Occupational Expectations in the Baltic Countries: Influence of Family, School, and Individual Science Learning Achievement." *Journal of Baltic Science Education* 20 (5):759-774. doi: 10.33225/jbse/21.20.759.
- Kanes, Clive, Candia Morgan, and Anna Tsatsaroni. 2014. "The PISA Mathematics Regime: Knowledge Structures and Practices of the Self." *Educational Studies in Mathematics* 87 (2):145-165. doi: 10.1007/s10649-014-9542-6.
- Kelly, Peter, and Hans-Georg Kotthoff. 2017. "PISA, National and Regional Education Policies and their Effect on Mathematics Teaching in England and Germany." *Research in Comparative and International Education* 12 (3):318-332. doi: 10.1177/1745499917724286.
- Kjærnsli, Marit, and Fredrik Jensen. 2016. *Stø kurs: norske elevers kompetanse i naturfag, matematikk og lesing i PISA 2015 [Steady Course: Norwegian Students' Competence in Science, Mathematics and Reading in PISA 2015]*. Oslo: Universitetsforlaget.
- Kjærnsli, Marit, Svein Lie, Rolf Vegar Olsen, and Astrid Roe. 2007. *Tid for tunge løft : norske elevers kompetanse i naturfag, lesing og matematikk i PISA 2006 [Time for Heavy Lifts: Norwegian Students' Competence in Science, Reading and Mathematics in PISA 2006]*. Oslo: Universitetsforlaget.
- Kjærnsli, Marit, Svein Lie, Rolf Vegar Olsen, Astrid Roe, and Are Turmo. 2004. *Rett spor eller ville veier? Norske elevers prestasjoner i matematikk, naturfag og lesing i PISA 2003 [Straight Tracks or Wild Roads? Norwegian Students' Achievements in Mathematics, Science and Reading in PISA 2003]*. Oslo: Universitetsforlaget.
- Kjærnsli, Marit, and Rolf Vegar Olsen, eds. 2013. *Fortsatt en vei å gå: norske elevers kompetanse i matematikk, naturfag og lesing i PISA 2012 [Still a Way to Go: Norwegian Students' Competence in Mathematics, Science and Reading in PISA 2012]*. Oslo: Universitetsforlaget.
- Kjærnsli, Marit, and Astrid Roe. 2010. *På rett spor: norske elevers kompetanse i lesing, matematikk og naturfag i PISA 2009 [On the right Track: Norwegian Students' Competence in Reading, Mathematics and Science in PISA 2009]*. Oslo: Universitetsforlaget.
- Komatsu, Hikaru, and Jeremy Rappleye. 2021. "Rearticulating PISA." *Globalisation, Societies and Education* 19 (2):245-258. doi: 10.1080/14767724.2021.1878014.
- Krejsler, John Benedicto. 2019. "How a European 'Fear of Falling Behind' Discourse Co-produces Global Standards: Exploring the Inbound and Outbound Performativity of the Transnational Turn in European Education Policy." In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 245–267. Cham: Palgrave Macmillan.
- Lewis, Steven, and Bob Lingard. 2023. "Platforms, Profits and PISA for Schools: New Actors, Bypasses and Topological Spaces in Global Educational Governance." *Comparative Education* 59 (1):99-117. doi: 10.1080/03050068.2022.2145006.

- Lewis, Steven, Sam Sellar, and Bob Lingard. 2016. "PISA for Schools: Topological Rationality and New Spaces of the OECD's Global Educational Governance." *Comparative Education Review* 60 (1):27-57. doi: 10.1086/684458.
- Lie, Svein, Marit Kjærnsli, Astrid Roe, and Are Turmo. 2001. *PISA 2000 med få ord: en kortversjon av den nasjonale rapporten: "Godt rustet for framtida? Norske 15-åringers kompetanse i lesing og realfag i et internasjonalt perspektiv" [PISA 2000 in a Few Words: A Short Version of the National Report: "Well Equipped for the Future? Norwegian 15-year-olds' Competence in Reading and Science in an International Perspective"]*. Oslo: Universitetet.
- Lingard, Bob. 2020. "Foreword." In *PISA, Policy and the OECD: Respatialising Global Educational Governance Through PISA for Schools*, edited by Steven Lewis, v–viii. Singapore: Springer.
- Lundgren, Ulf. 2010. "PISA." In *PISA: sannheten om skolen? [PISA: The Truth About School?]*, edited by Eyvind Elstad and Kirsten Sivesind, 42–51. Oslo: Universitetsforlaget.
- Mausethagen, Sølvi. 2013. "Governance Through Concepts: The OECD and the Construction of "Competence" in Norwegian Education Policy." *Berkeley Review of Education* 4 (1):161–181.
- MBN. n.d. "What is the OECD? Definition and meaning." accessed 05.09.22. <https://marketbusinessnews.com/financial-glossary/oecd-definition-meaning/>
- Merriam, Sharan B. 1988. *Case Study Research in Education: A Qualitative Approach, Jossey-Bass education series*. San Francisco: Jossey-Bass.
- Meyer, Heinz-Dieter, and Katie Zahedi. 2014. "An Open Letter: To Andreas Schleicher, OECD, Paris." *Policy Futures in Education* 12 (7):872-877.
- Morgan, Clara. 2018. "The Spectacle of Global Tests in the Arabian Gulf: A Comparison of Qatar and the United Arab Emirates." *Comparative Education* 54 (3):285-308. doi: 10.1080/03050068.2017.1348018.
- Mumford, Stephen, and Rani Lill Anjum. 2010. "A Powerful Theory of Causation." In *The Metaphysics of Powers: Their Grounding and their Manifestations*, edited by Anna Marmodoro, 143–159. New York: Routledge.
- Mumford, Stephen, and Rani Lill Anjum. 2013. "With Great Power Comes Great Responsibility." In *Critical Essays on Causation and Responsibility: Critical Essays*, edited by Benedikt Kahmen and Markus Stepanians, 219–238. Berlin/Boston: De Gruyter, Inc.
- Møller, Jorunn, and Eli Ottesen. 2011. "Styring, ledelse og kunnskapsutvikling i skolen [Management, Leadership and Knowledge Development in the School]." In *Rektor som leder og sjef: om styring, ledelse og kunnskapsutvikling i skolen [Principal As Leader and Manager: On Management, Leadership and Knowledge Development in the School]*, edited by Jorunn Møller and Eli Ottesen, 15-26. Oslo: Universitetsforlaget.
- Nash, Roy. 2005. "Explanation and Quantification in Educational Research: The Arguments of Critical and Scientific Realism." *British Educational Research Journal* 31 (2):185-204. doi: 10.1080/0141192052000340206.
- NESH. 2021. *Guidelines for Research Ethics in the Social Sciences and the Humanities*. 5th ed. Oslo: The Norwegian National Research Ethics Committees.
- Nilsen, Trude, and Sigrid Blömeke. 2018. "Lærerkvalitet, undervisningskvalitet, -kvantitet og presentasjon [Teacher Quality, Teaching Quality, -Quantity and Performance]." In *Tjue år med TIMSS og PISA i Norge : trender og nye analyser [Twenty Years with TIMSS and PISA in Norway: Trends and New Analyses]*, edited by Julius Kristjan Björnsson and Rolf Vegar Olsen, 57–72. Oslo: Universitetsforlaget.
- Nordkvelle, Yngve, and Lene Nyhus. 2017. "Management by Objectives as an Administrative Strategy in Norwegian Schools: Interpretations and Judgements in Contrived Liberation." In *Administrative Strategies of Our Time*, edited by Noralv Veggeland, 220–260. New York: Nova Science Publishers, Inc.
- Nortvedt, Guri A. 2018. "Policy Impact of PISA on Mathematics Education: The Case of Norway." *European Journal of Psychology of Education* 33 (3):427–444. doi: 10.1007/s10212-018-0378-9.
- Nortvedt, Guri A., Andreas Pettersen, Astrid Petterson, and Samuel Sollerman. 2016. "Is PISA relevant to Maths Education in Norway and Sweden? ." In *Northern Lights on PISA and*

- TALIS, edited by The Nordic Evaluation Network group, 27–56. Copenhagen: Nordic Council of Ministers.
- NOU [Green Paper] 2023:1. "Kvalitetsvurdering og kvalitetsutvikling i skolen. Et kunnskapsgrunnlag [Quality Assessment and Quality Development in Schools. A Knowledge Base]." <https://www.regjeringen.no/contentassets/8758c1a973ae4f94ab67e3830c96b9c9/no/pdfs/nou202320230001000dddpdfs.pdf>.
- Nusantara, Duano Sapta, Zulkardi Zulkardi, and Ratu Ilma Indra Putri. 2021. "Designing PISA-Like Mathematics Task Using a COVID-19 Context (PISACOMAT)." *Journal on Mathematics Education* 12 (2):349-364. doi: 10.22342/jme.12.2.13181.349-364.
- Nusche, Deborah, Lorna Earl, William Maxwell, and Claire Shewbridge. 2011. *OECD Reviews of Evaluation and Assessment in Education: Norway 2011, OECD Reviews of Evaluation and Assessment in Education*. Paris: OECD Publishing.
- Nyhus, Lene, Eirin Annamo, and Trond Gansmo Jakobsen. 2019. "Sustainability, Interdisciplinarity and Transformative Change: A Critical Realist Response to the Crisis System." *Journal of Critical Realism* 18 (4):347–351. doi: 10.1080/14767430.2019.1667731.
- O'Shea, John, and Aisling M. Leavy. 2013. "Teaching Mathematical Problem-Solving from an Emergent Constructivist Perspective: The Experiences of Irish Primary Teachers." *Journal of Mathematics Teacher Education* 16 (4):293-318. doi: 10.1007/s10857-013-9235-6.
- OECD. 2019. *PISA 2018 Results (Volume I): What Students Know and Can Do*. Paris: OECD publishing.
- OECD. n.d-a. "List of OECD Member countries - Ratification of the Convention on the OECD." accessed 13.12.22. <https://www.oecd.org/about/document/ratification-oecd-convention.htm>.
- OECD. n.d-b. "PISA 2022 participants." accessed 08.10.22. <https://www.oecd.org/pisa/aboutpisa/pisa-2022-participants.htm>.
- OECD. n.d-c. "PISA for Schools - FAQs." accessed 01.09.2023. <https://www.oecd.org/pisa/aboutpisa/pisa-based-test-for-schools-faq.htm>.
- OECD. n.d-d. "PISA: FAQ: Overview." accessed 08.10.22. <http://www.oecd.org/pisa/pisafaq/>
- OECD. n.d-e. "PISA: FAQ: Process." accessed 12.09.23. <http://www.oecd.org/pisa/pisafaq/>
- OECD. n.d-f. "PISA: FAQ: Results." accessed 12.09.23. <http://www.oecd.org/pisa/pisafaq/>
- OECD. n.d-g. "Programme for International Student Assessment (PISA) 2000 - List of Participating Countries/Economies." accessed 30.10.22. <https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/programmeforinternationalstudentassessmentpisa2000-listofparticipatingcountriесеconomies.htm>.
- Ozga, Jenny, and Martin Lawn. 1988. "Schoolwork: Interpreting the Labour Process of Teaching." *British Journal of Sociology of Education* 9 (3):323–336. doi: 10.1080/0142569880090305.
- Pettersen, Hilde M , and Kjell Arne Røvik. 2014. "Masterideér [Master ideas]." In *Reformideer i norsk skole : spredning, oversettelse og implementering [Reform ideas in Norwegian schools: dissemination, translation and implementation]*, edited by Kjell Arne Røvik, Tor Vidar Eilertsen and Eli Moksnes Furu, 53–79. Oslo: Cappelen Damm akademisk.
- Pham, Lien. 2019. "Capital and Capabilities in Education: Re-examining Australia's 2015 PISA Performance and Context Assessment Framework." *Policy Futures in Education* 17 (5):599-617. doi: 10.1177/1478210318808615.
- Pons, Xavier. 2017. "Fifteen Years of Research on PISA Effects on Education Governance: A Critical Review." *European Journal of Education* 52 (2):131–144. doi: 10.1111/ejed.12213.
- Price, Leigh, and Lee Martin. 2018. "Introduction to the Special Issue: Applied Critical Realism in the Social Sciences." *Journal of Critical Realism* 17 (2):89-96.
- Radišić, Jelena, and Aleksandar Baucal. 2018. "Teachers' Reflection on PISA Items and why they are so hard for Students in Serbia." *European Journal of Psychology of Education* 33 (3):445-466. doi: 10.1007/s10212-018-0366-0.
- Rautalin, Marjaana, Pertti Alasuutari, and Eetu Vento. 2019. "Globalisation of Education Policies: Does PISA have an Effect?" *Journal of Education Policy* 34 (4):500-522. doi: 10.1080/02680939.2018.1462890.
- Reid, Alan. 2003. "Understanding Teachers' Work: Is There Still A Place for Labour Process Theory?" *British Journal of Sociology of Education* 24 (5):559–573. doi: 10.1080/0142569032000127134.

- Rienecker, Lotte, Peter Stray Jørgensen, Signe Skov, and Wenche Landaas. 2013. *Den gode oppgaven: håndbok i oppgaveskriving på universitet og høyskole [The Good Assignment: Handbook in Assignment Writing at University and College]*. 2nd ed. Bergen: Fagbokforlaget.
- Ringarp, Johanna, and Martin Rothland. 2010. "Is the Grass Always Greener? The Effect of the PISA Results on Education Debates in Sweden and Germany." *European Educational Research Journal* 9 (3):422-430.
- Robson, Colin, and Kieran McCartan. 2016. *Real World Research: A Resource for Users of Social Research Methods in Applied Settings*. 4th ed. Chichester: Wiley.
- Rutkowski, Leslie, and David Rutkowski. 2016. "A Call for a More Measured Approach to Reporting and Interpreting PISA Results." *Educational Researcher* 45 (4):252-257. doi: 10.3102/0013189X16649961.
- Rutkowski, Leslie, David Rutkowski, and Yan Zhou. 2016. "Item Calibration Samples and the Stability of Achievement Estimates and System Rankings: Another Look at the PISA Model." *International Journal of Testing* 16 (1):1-20. doi: 10.1080/15305058.2015.1036163.
- Sáenz, César. 2009. "The Role of Contextual, Conceptual and Procedural Knowledge in Activating Mathematical Competencies (PISA)." *Educational Studies in Mathematics* 71 (2):123-143. doi: 10.1007/s10649-008-9167-8.
- Sahlberg, Pasi. 2011. "The Fourth Way of Finland." *Journal of Educational Change* 12 (2):173-185. doi: 10.1007/s10833-011-9157-y.
- Santos, Íris, Luís Miguel Carvalho, and Benedita Portugal e Melo. 2022. "The Media's Role in Shaping the Public Opinion on Education: A Thematic and Frame Analysis of Externalisation to World Situations in the Portuguese Media." *Research in Comparative and International Education* 17 (1):29-50. doi: 10.1177/17454999211057753.
- Sayer, Andrew. 2000a. *Realism and Social Science*. London: Sage.
- Sayer, Andrew. 2000b. "System, Lifeworld and Gender: Associational Versus Counterfactual Thinking." *Sociology* 34 (4):707-725.
- Scott, David. 2014. "Ontology, Epistemology, Strategy and Method in Educational Research. A Critical Realist Approach." *Magis. Revista Internacional de Investigación en Educación* 7 (14):29-38. doi: 10.11144/Javeriana.M7-14.OESM.
- Sellar, Sam, and Bob Lingard. 2013. "Looking East: Shanghai, PISA 2009 and the Reconstitution of Reference Societies in the Global Education Policy Field." *Comparative Education* 49 (4):464-485. doi: 10.1080/03050068.2013.770943.
- Sellar, Sam, and Bob Lingard. 2018. "International Large-Scale Assessments, Affective Worlds and Policy Impacts in Education." *International Journal of Qualitative Studies in Education* 31 (5):367-381. doi: 10.1080/09518398.2018.1449982.
- Sennett, Richard. 1998. *The Corrosion of Character : The Personal Consequences of Work in the New Capitalism*. New York: W.W. Norton & Company.
- Serder, Margareta, and Malin Ideland. 2016. "PISA Truth Effects: The Construction of Low Performance." *Discourse* 37 (3):341-357. doi: 10.1080/01596306.2015.1025039.
- Sivesind, Kirsten. 2019. "Nordic Reference Societies in School Reforms in Norway: An Examination of Finland and the Use of International Large-Scale Assessments." In *Understanding PISA's Attractiveness : Critical Analyses in Comparative Policy Studies*, edited by Florian Waldow and Gita Steiner-Khamsi, 89-108. London: London: Bloomsbury Publishing Plc.
- Sjøberg, Svein. 2014a. "Hva PISA-testen måler [What the PISA Test Measures]." *Bedre Skole* (04):70-75.
- Sjøberg, Svein. 2014b. "PISA-syndromet – Hvordan norsk skolepolitikk blir styrt av OECD [The PISA Syndrome – How Norwegian School Policy is Directed by the OECD]." *Nytt Norsk Tidsskrift* 31 (01):30-43.
- Sjøberg, Svein. 2014c. "PISAfisering av norsk skole: en suksesshistorie fra OECD [PISAfication of Norwegian School: A Success Story From the OECD]." In *Reformideer i norsk skole: spredning, oversettelse og implementering [Reform Ideas in Norwegian Schools: Dissemination, Translation and Implementation]*, edited by Kjell Arne Røvik, Tor Vidar Eilertsen and Eli Moksnes Furu, 195-225. Oslo: Cappelen Damm akademisk.
- Sjøberg, Svein. 2017. "PISA Testing: A global Educational Race?" *Europhysics News* 48 (4):17-20. doi: 10.1051/eprn/2017402.

- Sjøberg, Svein. 2019. "The PISA-syndrome – How the OECD has Hijacked the Way We Perceive Pupils, Schools and Education." *Confero: Essays on Education, Philosophy and Politics* 7 (1):12–65. doi: 10.3384/confero.2001-4562.190125.
- Sjøberg, Svein. 2022. "Da skolen gikk ad PISA [When the School went to PISA]." In *Skolens mening : femti år etter Nils Christies. Hvis skolen ikke fantes [The School's Meaning: Fifty Years after Nils Christies. If the School did not Exist]*, edited by Espen Schaanning and Willy Aagre, 131–153. Oslo: Universitetsforlaget.
- Sjøberg, Svein. 2023. "Nasjonalt kvalitetsvurderingssystem for skolen [National Quality Assessment System for Schools]." Last Modified 22.05.2023, accessed 10.10.2023. https://snl.no/Nasjonalt_kvalitetsvurderingssystem_for_skolen.
- Skedsmo, Guri, and Sølvi Mausestaden. 2017. "Nye styringsformer i utdanningssektoren - spenninger mellom resultatstyring og faglig-profesjonelt ansvar [New Forms of Management in the Education Sector—Tensions between Performance Management and Professional Responsibility]." *Norsk pedagogisk tidsskrift* 101 (2):169-179.
- Skedsmo, Guri, and Jorunn Møller. 2016. "Governing by New Performance Expectations in Norwegian Schools." In *New Public Management and the Reform of Education: European Lessons for Policy and Practice*, edited by Helen M Gunter, Emiliano Grimaldi, David Hall and Roberto Serpieri, 53-65. London: Routledge.
- Skinningrud, Tone. 2006. "Komparativt perspektiv på framveksten av statlige utdanningssystemer [Comparative Perspective on the Emergence of State Education Systems]." In *Å greie seg i utdanningssystemet i nord og sør: innføring i flerkulturell og komparativ pedagogikk, utdanning og utvikling [Dealing with the Education System in the North and South: Introduction to Multicultural and Comparative Pedagogy, Education and Development]*, edited by Birgit Brock-Utne and Liv Bøyese, 135–148. Bergen: Fagbokforlaget.
- Skinningrud, Tone. 2012. "Fra reformasjonen til mellomkrigstiden: Framveksten av det norske utdanningssystemet [From the Reformation to the Interwar Period: The Emergence of the Norwegian Education System]." Doctoral thesis, The Faculty of Humanities, Social Sciences and Teacher Education, UiT - The Arctic University of Norway, Tromsø, Norway.
- Skinningrud, Tone. 2019. "Vindicating Archer's Concepts of Educational Systems – Centralized and Decentralized – As Exemplars of Critical Realist Theorizing." *Journal of Critical Realism* 18 (4):453–473. doi: 10.1080/14767430.2019.1656924.
- Solberg, Mariann. 2021. "Hvilken rolle spiller menneskesyn i pedagogisk tenkning? [What Role Does Views of Humanity Play in Educational Thinking?]." In *Grunnspørsmål i pedagogikken [Fundamental Questions in Pedagogy]*, edited by Line Torbjørnsen Hilt and Lars Petter Storm Torjussen, 151–168. Bergen: Fagbokforlaget.
- Stake, Robert E. 1978. "The Case Study Method in Social Inquiry." *Educational Researcher* 7 (2):5-8. doi: 10.3102/0013189X007002005.
- Stake, Robert E. 1995. *The Art of Case Study Research*. Thousand Oaks, California: Sage.
- Steiner-Khamsi, Gita. 2016. "Standards are Good (for) Business: Standardised Comparison and the Private Sector in Education." *Globalisation, Societies and Education* 14 (2):161-182. doi: 10.1080/14767724.2015.1014883.
- Stray, Janicke Heldal, and Bronwyn Wood. 2020. "Global-local Education Policy Dynamics: A Case Study of New Zealand and Norway." *Scandinavian Journal of Educational Research* 64 (2):256–269. doi: 10.1080/00313831.2018.1541818.
- Sun, Jingping, Robert Przybylski, and Bob J Johnson. 2016. "A Review of Research on Teachers' Use of Student Data: From the Perspective of School Leadership." *Educational Assessment, Evaluation and Accountability* 28 (1):5-33.
- Svensson, Lennart G. 2006. "New Professionalism, Trust and Competence: Some Conceptual Remarks and Empirical Data." *Current Sociology* 54 (4):579–593. doi: 10.1177/0011392106065089.
- Telhaug, Alfred Oftedal, and Odd Asbjørn Mediås. 2003. *Grunnskolen som nasjonsbygger: Fra statspietisme til nyliberalisme [Primary School As A Nation-Building: From State Pietism to Neoliberalism]*. Oslo: Abstrakt Forlag.
- Thagaard, Tove. 2018. *Systematikk og innlevelse: en innføring i kvalitative metoder [Systematics and Immersion: An Introduction to Qualitative Methods]*. 5th ed. Bergen: Fagbokforlaget.

- The Guardian. 2014. "OECD and PISA tests are Damaging Education Worldwide - Academics." Last Modified Tue 6 May 2014 07.00 accessed 08.10.2022. <https://www.theguardian.com/education/2014/may/06/oecd-pisa-tests-damaging-education-academics>.
- Thomas, Paul. 2021. "The Discourse of International Standard-Setting: PISA tests and Norway, A Critical Discourse Analysis." *Education In The North* 28 (2):6–23. doi: 10.26203/nns1-bt94.
- Thuen, Harald. 2010. "Skolen – Et liberalistisk prosjekt? 1860–2010 [The school—A liberal project? 1860–2010]." *Norsk pedagogisk tidsskrift* 94 (4):273–287.
- Troman, Geoff. 1996. "The Rise of the New Professionals? The Restructuring of Primary Teachers' Work and Professionalism." *British Journal of Sociology of Education* 17 (4):473–487. doi: 10.1080/0142569960170404.
- Tveit, Sverre. 2014. "Educational Assessment in Norway." *Assessment in Education: Principles, Policy & Practice* 21 (2):221-237. doi: 10.1080/0969594X.2013.830079.
- UiO. n.d. "Om senteret [About the Centre]." accessed 21.09.2022. <https://www.uv.uio.no/cemo/om/>.
- UiT. n.d. "PhD at UiT - The Arctic University of Norway: Faculty of Humanities, Social Science and Education: Monograph or article collection?". University of Tromsø - The Arctic University of Norway, accessed 16.09.2022. <https://uit.no/phd-en#v-pills-669539>.
- Utdanningsdirektoratet. n.d. "Internasjonale studier [International Studies]." accessed 26.09.2023. <https://www.udir.no/tall-og-forskning/internasjonale-studier/>.
- Utdanningsforbundet. 2008. Spørreundersøkelsen om PISA blant Utdanningsforbundets medlemmer [The Survey on PISA among the Members of the Union of Education Norway]. Oslo: The Union of Education Norway.
- Verger, Antoni, Clara Fontdevila, and Lluís Parcerisa. 2019. "Constructing School Autonomy with Accountability as a Global Policy Model: A Focus on OECD's Governance Mechanisms." In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 219–243. Cham: Palgrave Macmillan.
- Yang, Wenjie, and Guorui Fan. 2019. "Post-PISA Education Reforms in China: Policy Response Beyond the Digital Governance of PISA." *ECNU Review of Education* 2 (3):297-310. doi: 10.1177/2096531119878557.
- Ydesen, Christian. 2019a. "The Formation and Workings of a Global Education Governing Complex." In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 291–303. Cham: Palgrave Macmillan.
- Ydesen, Christian. 2019b. "Introduction: What Can We Learn About Global Education from Historical and Global Policy Studies of the OECD?" In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 1–14. Cham: Palgrave Macmillan.
- Yin, Robert K. 1994. *Case Study Research: Design and Methods*. 2nd ed. Vol. 5, *Applied Social Research Methods Series*. Thousand Oaks: SAGE Publications.
- Zhao, Yong. 2020. "Two Decades of Havoc: A Synthesis of Criticism Against PISA." *Journal of Educational Change* 21 (2):245-266. doi: 10.1007/s10833-019-09367-x.
- Zheng, Jia-qi, Kwok-cheung Cheung, and Pou-seong Sit. 2022. "Insights from Two Decades of PISA-related Studies in the New Century: A Systematic Review." *Scandinavian Journal of Educational Research* ahead-of-print (ahead-of-print):1-18. doi: 10.1080/00313831.2022.2148273.
- Ørskov, Frederik Forrai. 2019. "Australian Education Joins the OECD: Federalism, Regionalization, and the Role of Education in a Time of Transition." In *The OECD's Historical Rise in Education: The Formation of a Global Governing Complex*, edited by Christian Ydesen, 85–107. Cham: Palgrave Macmillan.
- Østerud, Svein. 2016. "Hva kan norsk skole lære av PISA-vinneren Finland? [What can Norwegian Schools Learn from the PISA Winner Finland?]." *Nordisk Tidsskrift for Pedagogikk og Kritikk* 2 (2):14-35.
- Aas, Marit, and Christian Brandmo. 2018. "Assessment Results for Transforming Practice: School Leaders' Role." *Nordic Studies in Education* 38 (2):174-194. doi: 10.18261/issn.1891-5949-2018-02-06.

Appendixes

Appendix A: The Norwegian Centre for Research Data (NSD) approval

Vurdering

Referansenummer

██████████

Type

Med vilkår

Dato

07.04.2022

Prosjektittel

Effekter av PISA i det norske utdanningssystemet (overordnet norsk tittel)

Behandlingsansvarlig institusjon

UiT Norges Arktiske Universitet / Fakultet for humaniora, samfunnsvitenskap og lærerutdanning / Institutt for lærerutdanning og pedagogikk

Prosjektansvarlig

Terje André Bringeland

Prosjektperiode

14.10.2019 - 31.12.2023

Kategorier personopplysninger

Alminnelige

Rettslig grunnlag

Samtykke (art. 6 nr. 1 bokstav a)

Behandlingen av personopplysningene kan starte så fremt den gjennomføres som oppgitt i meldeskjemaet. Det rettslige grunnlaget gjelder til 31.12.2023.

[Meldeskjema](#) 

Kommentar

Vi viser til endring registrert 13.03.2022. Vi kan ikke se at det er gjort noen oppdateringer i meldeskjemaet eller vedlegg som har innvirkning på vår vurdering av hvordan personopplysninger behandles i prosjektet.

OPPFØLGING AV PROSJEKTET

Personverntjenester vil følge opp ved planlagt avslutning for å avklare om behandlingen av personopplysningene er avsluttet.

Lykke til videre med prosjektet!

Appendix B: Informed letter of consent—mathematics teacher

Vil du delta i forskningsprosjektet

«PISA: Introduksjon og rolle i utdanningssystemer»

Dette er en forespørsel til deg som matematikklærer om å delta i et forskningsprosjekt hvor formålet er å få bedre innsikt i hvilken rolle PISA testen har i relasjon til deg som matematikklærer. I dette skrivet gir jeg deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

Formål

PISA er en internasjonal undersøkelse som brukes for å kartlegge elevers kompetanse iblant annet matematikk. PISA har blitt et sentralt referansepunkt i samfunnet når det gjelder norske matematikkresultater og implementering av utdanningspolitikk. Men vi vet lite om mulige virkninger som følge av dette på matematikklærerens rolle og arbeid. Formålet med denne studien er å undersøke erfarte PISA-effekter hos matematikklærere i ungdomsskolen. Undersøkelsen inngår i min doktorgradsavhandling med foreløpig tittel «PISA: Introduksjon og rolle i utdanningssystemer». Mitt bidrag til forskningsfeltet vil være bruk av nyere samfunnsteorier for å forklare PISA som fenomen og mulige effekter hos matematikklærere. Intervjuene med matematikklærere har som mål senere å danne grunnlaget for en kvantitativ spørreundersøkelse, som kan undersøke problemstillingen hos matematikklærere nasjonalt for å få innsikt i hvordan en eventuell PISA-effekt gjør seg gjeldende i en skolehverdag preget av mål- og resultatstyring. Spørsmål knyttet til din utdanningsbakgrunn, hvordan det er å være matematikklærer på din skole, og profesjonsfaglige spørsmål er derfor aktuelle i tillegg til spørsmål om PISA. I tillegg ønsker jeg at du svarer på et kort papirbasert spørreskjema som skal belyse det vi har snakket om i intervjuet.

Hvem er ansvarlig for forskningsprosjektet?

Universitetet i Tromsø er ansvarlig for prosjektet.

Hvorfor får du spørsmål om å delta?

Du får forespørsel om å delta siden du kan bidra med bevisstgjøring og innsikt om PISA-effektens utbredelse og til å forstå samspillet mellom utdanningspolitikk og matematikklærere.

Du kan bidra til å gi studenter, forskere og lærere bedre forståelse for PISA i utdanningssystemet. Du kan bidra til å supplere opplæringen i lærerutdanningen og pedagogikk med informasjon som kan brukes i undervisning av kommende lærere. Du kan bidra til å få frem matematikklæreres stemmer og belyse deres handlingsmuligheter i praksis. Du kan bidra til å gi innsikt i hva det vil si å være en profesjonell matematikklærer i en tid med mål- og resultatstyring.

Utvalgskriteriet for denne undersøkelsen er at du må arbeide som matematikklærer på ungdomsskolen. Det er ikke et krav til at du må ha gjennomført PISA testen hos dine elever. Dette informasjonsbrevet ble sendt til din skoleleder som videre distribuerer informasjon om dette forskningsprosjektet til deg som matematikklærer.

Hva innebærer det for deg å delta?

Datainnsamlingen innebærer først et SKYPE intervju med deg. Skype-intervjuet er estimert til ca. 50 minutter. De opplysningene som innhentes er hovedsakelig dine erfaringer og tanker som matematikklærer. Etter intervjuet besvarer du et papirbasert spørreskjema med svaralternativer som sendes ut til deg på epost. Dette spørreskjemaet tar ca. 10 minutter å besvare. Spørreskjemaet returneres så til meg på mail.

Det er frivillig å delta

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykke tilbake uten å oppgi noen grunn. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg. Dersom du trekker deg, vil datamaterialet som er samlet inn fra deg bli slettet og ikke bli brukt i oppgaven.

Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger

Ditt og skolens navn vil ikke registreres og undersøkelsen vil bli anonymisert. Det vil bli brukt fiktive navn og deltakerne vil ikke kunne gjenkjennes i publikasjonene. Dataene registreres ved hjelp av opptaksfunksjonen på SKYPE og ved hjelp av notater. Opptaket lagres på en kryptert USB-minnepenn med passord. Opptaket og det transkriberte materialet vil i løpet av undersøkelsen bare være tilgjengelig for meg, som prosjektleder. Navnet og kontaktopplysningene dine vil jeg erstatte med en kode som lagres på egen navneliste adskilt fra øvrige data.

Jeg vil bare bruke opplysningene som du gir til formålene vi har fortalt om i dette skrivet. Jeg behandler opplysningene konfidensielt og i samsvar med personvernregelverket. Det er kun jeg som vil ha tilgang til dataene som samles inn for dette forskningsprosjektet.

Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?

Prosjektet skal avsluttes i løpet av 2023 med forbehold om forutsette hendelser som eksempelvis sykdom. Du vil bli bedt om å ta stilling til om datamaterialet kan lagres for videre bruk. Hvis du samtykker i dette vil det muligens bli benyttet av meg eller andre forskere for videre forskning innenfor feltet ved en senere anledning. Skype-opptak vil bli slettet etter prosjektslutt. Det er det besvarte papirbaserte spørreskjemaet, intervjuguide (spørsmålene) og det transkriberte datamaterialet som eventuelt lagres etter prosjektslutt.

Dine rettigheter

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg,
- å få rettet personopplysninger om deg,
- få slettet personopplysninger om deg,
- få utlevert en kopi av dine personopplysninger (dataportabilitet), og
- å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine personopplysninger.

Hva gir oss rett til å behandle personopplysninger om deg?

Jeg behandler opplysninger om deg basert på ditt samtykke. På oppdrag fra Universitetet i Tromsø har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med

Universitet i Tromsø ved:

Stipendiat Terje A. Bringeland (Prosjektleder): [95826258](tel:95826258) / terje.a.bringeland@uit.no

Professor Emerita Tone Skinningsrud (Hovedveileder): [97734158](tel:97734158) / tone.skinningsrud@uit.no

Personvernombud Joakim Bakkevold: 776 46 322 og 976 915 78 / personvernombud@uit.no

Eller

NSD – Norsk senter for forskningsdata AS: 55 58 21 17 / (personverntjenester@nsd.no)

Med vennlig hilsen

Prosjektleder
(Stipendiat)

X

Terje A. Bringeland

Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet «PISA: Introduksjon og rolle i utdanningssystemer», og har fått anledning til å stille spørsmål om hva deltakelse i prosjektet innebærer. Jeg samtykker:

- til å delta i intervju gjennomført på Skype
- til å delta på den papirbaserte spørreundersøkelsen
- til at datamaterialet (svar på spørreskjema, intervjuguide og transkribert intervju) kan lagres og brukes etter prosjektslutt til bruk av prosjektleder
- til at datamaterialet (svar på spørreskjema, intervjuguide og transkribert intervju) kan lagres og brukes av andre forskere

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet, ca. 31.12.2023

(Signert av prosjektdeltaker/informant, dato)

Appendix C: The interview guide with the mathematics teacher

Tema / Intervjusspørsmål**Bakgrunn**

1. Hvor gammel er du?

18-25	26-35	36-50	51+
-------	-------	-------	-----

2. Mann eller kvinne?

Mann	Kvinne
------	--------

3. Hvilken utdanningsbakgrunn har du?

Videregående	Lærerutdanning	PPU	Annen
--------------	----------------	-----	-------

4. Hvor mange studiepoeng har du i matematikk?

0-30	31-60	61-120	121-300
------	-------	--------	---------

5. Hvor mange år har du arbeidet som matematikklærer på ungdomsskolen?

0-5	6-11	12-17	18+
-----	------	-------	-----

6. Hva var din motivasjon for å bli matematikklærer?

7. Opplever du at læreryrket er som du trodde det ville være?

8. Opplever du at din bakgrunn påvirker din praksis? I så fall, hvordan?

Miljøet

9. Har skolen teamarbeid blant matematikklærere? I så fall, hvordan foregår dette?

10. Har skolen matematikk som et satsningsområde? I så fall, hvordan utspiller dette seg?

11. Blir det truffet beslutninger av skoleeier (kommunen) som får konsekvenser for din matematikkundervisning?

12. Har skoleledelsen synspunkter som får konsekvenser for din matematikkundervisning?

13. Opplever du at matematikkfaget har prestisje på din skole i forhold til andre fag? I så fall hvordan?

14. Er det faglig samarbeid mellom matematikklærerne på denne skolen for matematikkundervisningen? I så fall, hvordan foregår dette?

15. Hvordan oppfatter du relasjonen mellom deg som lærer og skoleleder?
16. Har skolen sanksjoner for dårlige resultater på normerte prøver i matte? I så fall, hvilke?
17. Har skolen belønningssystem av gode resultater på tester? I så fall, hvilke?
18. Opplever du at kollegaer jukser med tester på din skole? I så fall hvilke og hvordan?
19. Hvordan vil du beskrive elevrollen i matematikkundervisningen hos dere?
20. Har du brukt oppgaver fra nasjonale og internasjonale tester til å forberede elever før testen skal gjennomføres? I så fall hvilke tester og hvorfor?
21. Hvilke læringsteorier/perspektiver mener du sammenfaller med din undervisning?
22. Er det noen tanker i utdanningssystemet, enten på lokalt eller sentral nivå, som du opplever som dominerende for din praksis?
23. Er det noen personer i utdanningssystemet, enten på lokalt eller sentral nivå, som du opplever som dominerende for din praksis?

Profesjonalitet

24. Har du som matematikklærer mulighet til å gjennomføre undervisningen på en profesjonelt forsvarlig måte, eller slik du mener er best? Hvis ikke, hva hindrer deg?
25. Opplever du at undervisningen din har endret seg med arbeidserfaringen din? I så fall, hvordan?
26. Opplever du at din måte å gjøre undervisningen på påvirkes av hendelser og engasjement utenfor skolen? I så fall, hvordan?
27. Hvordan opplever du dine kollegaers muligheter for å gjennomføre undervisningen på en profesjonelt forsvarlig måte på din skole? Hvis ikke, hva hindrer de?
28. Opplever du at du må dokumentere resultater? I så fall hvordan?
29. Har du opplevd endringer i din lærerrolle over tid? I så fall hvilke?
30. Har du opplevd endringer i hvordan skoler blir styrt i din karriere? I så fall hvilke?
31. Hvordan opplever du ytringsfriheten på din skole?
32. Bruker du internasjonale og nasjonale undersøkelser knyttet til matematikkundervisning? I så fall, hvilke og hvorfor?
33. Hvordan opplever du din selvbestemmelse (autonomi) i henhold til bruk av tester?

PISA

34. Har du gjennomført PISA i matematikk?
35. Har skolen din gjennomført PISA i matematikk?
36. Hvordan forstår du intensjonen med PISA?
37. Hvordan stiller du deg til PISA testen?
38. Hvordan har du fått kjennskap til PISA?

39. Hvordan var reaksjonen til PISA på skolen din når testen ble introdusert i 2000?
40. Har du merket deg noe endringer i holdninger til PISA testen fra testen ble introdusert i 2000 til nå?
41. Mener du at det har en betydning at en kommersiell aktør som Pearson utarbeider rammeverket til PISA testen? I så fall, på hvilken måte?
42. Opplever du at organisasjonen for økonomisk samarbeid og utvikling (OECD) påvirker din praksis? I så fall, hvordan?
43. Mener du at lærerprofesjonen som gruppe har blitt påvirket av PISA testen? I så fall, på hvilken måte?
44. Forberedes det til PISA undersøkelsen i matematikk ved skolen? I så fall, hvordan?
45. Opplever du at PISA har påvirket matematikkundervisningen ved skolen? I så fall, hvordan?
46. Har du blitt tilbudt PISA-kurs? I så fall, er dette obligatoriske kurs?
47. Opplever du at PISA har påvirket dine matematikkollegaer? I så fall, hvordan?
48. Opplever du at PISA har påvirket elevens foreldre? I så fall, hvordan?
49. Opplever du at PISA har påvirket dine elever? I så fall, hvordan?
50. Har PISA påvirket din lærerrolle i matematikk?
51. Opplever du at skolen har gjennomført tiltak som følge av PISA-resultater i matematikkfaget? I så fall, hvilke?
52. Hvem bestemmer om PISA-resultatene skal brukes for videre arbeid i matematikkfaget ved din skole?
53. Hvilke muligheter og begrensninger opplever du PISA gir?
54. Hvordan opplever du kompetansen din til å tolke PISA resultater?
55. Hvordan opplever du dine kollegaers kompetanse til å tolke PISA resultater?
56. Hvor relevant mener du PISA i matematikk er i forhold til matematikkopplæringen på din skole?

Appended papers

Article I: The Impact of PISA on Education in Norway: A Morphogenetic Perspective on Structural Elaboration in an Education System

Bringeland, Terje André. 2022. "The impact of PISA on Education in Norway: A Morphogenetic Perspective on Structural Elaboration in an Education System." In *The Morphogenesis of the Norwegian Educational System: Emergence and Development from a Critical Realist Perspective*, edited by Margaret S. Archer, Unn-Doris K. Bæck and Tone Skinningsrud, 147–180. Milton: Taylor & Francis Group.

Article II: PISA and Teachers' Reflexivities. A Mixed Methods Case Study

Bringeland, Terje André, and Tone Skinningsrud. 2023. PISA and Teachers' Reflexivities. A Mixed Methods Case Study. Unpublished manuscript, last modified November 27, 2023. Portable Document Format.

PISA and Teachers' Reflexivities. A Mixed Methods Case Study

Abstract

Neoliberal educational reforms include extensive use of standardized tests. We examine the Program for International Student Assessment (PISA) initiated and developed by the Organization for Economic Cooperation and Development (OECD). Examining previous research on teachers' reactions to neoliberal reforms altering their work context, we have identified three theoretical frameworks in use: labour process theory, derived from structural Marxism; post-structuralism, relying primarily on Foucault's conceptualizations; and 'new professionalism', which has emerged from the theory of professions. A major weakness in these frameworks is their inability to account for different reactions to the same structural and cultural conditions. Therefore, we suggest utilizing Archer's theories of reflexivity and reflexive modes to understand these reactions. Presenting one Norwegian lower secondary school teacher's reflexive engagement with PISA in the Norwegian school context as an example, our mixed methods case study indicates autonomous reflexivity as the prevailing mode. Meta-reflexivity rates second in prevalence.

Keywords: teachers' work, labour process theory, post-structuralism, professionalism, Archer, reflexivity, and reflexive modes

Introduction

Research on the consequences of neoliberal educational reforms at the micro level of the school and the classroom started out in the 1980s and 1990s when standardized testing, the demand for performativity and, increasingly, accountability regimes were introduced in schools. Some have traced the neoliberal turn in educational policies to 'New Right' think tanks, which in the 1980s furnished the US, the British and the Chilean government with ideas for a new educational policy (Fuller 2019). The central theoretical paradigm in early investigations on how neoliberal educational reforms impacted teachers was the neo-Marxist labour process theory formulated by Harry Braverman (1974). He saw the new forms of work control as a general tendency in capitalist economies, resulting in de-skilling, intensification of work, and a separation of the conception of work from its execution. Applications of the labour process approach to teachers' work were adopted on both sides of the Atlantic. Jenny Ozga and Martin Lawn (1988) in Britain and Michael W. Apple (2013) in the United States were among the

prominent early representatives of this theoretical tradition. However, at the time, Ozga and Lawn (1988) pointed to a lack in labour process theory and argued for the incorporation of (collective) agency in explaining the historical development of the skilling and de-skilling of the work force. They claimed that the reorganization of work and de-skilling was not inevitable and uncontested, though they still argued for upholding the labour process approach. This early paradigm, however, was criticized for its structural determinism and its lack of a subject (Reid 2003, 563). The structuralist perspective overlooked teachers' subjectivities, their agency, and the admittedly varied teacher reactions elicited by the new types of controls introduced by neoliberal reforms (Reid 2003).

Following the ebb of labour process theory, post-structuralist theory, leaning on Foucault's conceptualizations, took over as the leading theoretical paradigm. Though appearing to study a neglected domain in labour process theory, namely 'subjectivity', this approach tended to perpetuate some inadequacies of structural determinism not by neglecting subjectivity and agency, but by assuming that social agents and their subjectivities are totally determined by external forces, resulting in a kind of 'neoliberal subjectivity'. Thus Steven J. Ball provocatively asserted that neoliberal educational reforms not only determine what teachers do, but 'who they are' (Ball 2003, 215). Ball later modified this statement by claiming that subjectivity is 'a site of struggle' (Ball 2016, 1129), which may include 'resistance' (Ball and Olmedo 2013), but did not venture into studying more closely a broader variety of reactions, except for 'resistance'.

Despite the theoretical bias of labour process theory and post-structural theory, which emphasize the uniformity of teacher reactions, both early and more recent empirical studies guided by neither of these two divergent theoretical approaches have reported teacher responses that are more varied and defied the expectation of uniformity (Troman 1996; Lewis and Hardy 2014). What is missing from many of the empirical studies, however, is a general and coherent conceptual framework that can explain the variety of teacher reactions to the new forms of control.

Addressing the general question of how structure influences agency and how agency itself is a cause contributing to structurally situated practices, Margaret Archer (2000, 2003, 2007, 2012) has revitalized the concept of reflexivity¹. Archer has suggested that reflexivity, or internal conversations, is the mediating process linking structure, culture, and agency. Various reflexive

¹ Reflexivity refers to real ongoing internal conversations in which all normal individuals engage when they discuss with themselves which course of action to pursue.

modes, i.e., communicative, autonomous, and meta- reflexivity entail that persons have different stances on or orientations to their structural and cultural environment. Some persons, however, have fractured reflexivity, which means they are unable to engage their personal reflexive powers, due to disabling internal or external circumstances. Applying the concepts of ‘reflexivity’ and ‘reflexive mode’ can account for why different agents may react differently to the same structural and cultural context. Agents’ modes of reflexivity, or their ‘way of being in the world’, co-determine reactions to given structural surroundings. In this way Archer’s theory can account for how social agents’ constellations of personal concerns, that is, what they deeply care about, mediate and modify structural influences on their individual courses of action.

We consider Archer’s theory of structure, culture, and agency, introducing reflexivity as the mediator of structural and cultural impacts, as a promising alternative to both structuralist and post-structuralist approaches. Observing the ontological distinctions between structure, culture, and agency, Archer’s approach avoids conflating them in concepts such as ‘neoliberal subjectivity’ and instead sees them as representing distinct causal powers. The conceptualization of reflexivity and reflexive modes as activated in individuals’ internal conversations about their personal concerns, deliberating on how these may be pursued in a given structural and cultural context can help to explain why teachers vary in their reactions to new types of control and ideas associated with tests such as the Program for International Student Assessment (PISA).

Archer’s concepts such as the internal conversation, reflexivity, and reflexive modes are based in a philosophical and social ontology of emergence (Archer 1995, 135ff), seeing reality as hierarchically stratified into levels with distinct causal mechanisms (Bhaskar 2016, 32). They are an integrated part of, and a further refinement of her morphogenetic approach, providing a more specified content to the ‘vague’ notion of structural and cultural ‘conditioning’ of action and interaction (Archer 2003, 2). This ensemble of concepts enable theoretically coherent explanations of the reported diversity of individual reactions to neoliberal educational reforms, which have been missing in previous empirical studies, for example Troman (1996, 474), Ball (2003, 215), and Ball, Maguire, and Braun (2012, 145).

In this article, we will illustrate how Archer’s theoretical approach enables the analysis of teachers’ responses to their structural and cultural context by reporting on one case study of a Norwegian secondary school teacher. In the following, we will first provide a short presentation of the Norwegian educational context, the implementation of New Public Management (NPM) reforms, which are a subspecies of neoliberal reforms, and Norway as a PISA-participant from

the very beginning. We will then discuss the three major theoretical approaches that have predominated in international studies of teacher reactions to neoliberal educational reforms: labour process theory; post-structuralist theory based on Foucault; and the conception of 'new professionalism'. Our presentation of these research traditions examines typical examples from the vast literature that is available and are not meant as comprehensive research reviews. We will then present Archer's theory of structure, culture and (reflexive) agency, explaining how reflexivity mediates, in different ways, the impact of structure and culture on individual action. By presenting a case study of one Norwegian secondary school teacher, we illustrate how Archer's concepts may be applied in practical research. The teacher interview shows how social structure, in this case the teacher's obligation to administer the PISA test as part of his job, and expectations to use PISA to improve his work (assessment for learning) activates various modes of reflexivity. It also shows how his compliance in administering the test, which he is very critical of, is a product of both his structural context and his predominant modes of reflexivity.

The uneven adoption of NPM reforms in Europe

In this article we subsume NPM reforms under the general label of neoliberal reforms, though NPM is a separate branch of neoliberal ideas that concerns the organization of the public sector in particular. Neoliberal and NPM principles refer to market organization, management techniques and accountability regimes that are adopted from the private sector, aiming to improve effectiveness and efficiency in the public sector including education (Møller and Skedsmo 2013; Gunter et al. 2016). Investigations studying the introduction and implementation of NPM in ten different European countries showed discrepant and uneven developments. The investigators explained this by differences in the national and local reform contexts (Gunter et al. 2016). Even among the Nordic countries, there are major differences in how neoliberal NPM-reforms have been implemented. While Sweden, since the 1990s, has introduced private schools on a previously unprecedented scale, Norway is described as a hesitant reformer, resisting educational competition and privatization (Møller and Skedsmo 2013). However, a common element in recent educational policies implemented in Norway, Sweden, Denmark, and Finland is their participation in the PISA test. Norway has participated in the test since its inception in the year 2000, and in 2004, the PISA test was included in the Norwegian National Quality Assessment System (NQAS) for education.

The introduction of NPM reforms in Norway

Describing the introduction of NPM-reforms in Norwegian education, Møller and Skedsmo (2013) claim that this happened in two separate waves. The first wave, starting at the end of the 1980s and the beginning of the 1990s, consisted in a restructuring of the governance of education, introducing management by objectives (MbO) and the restructuring of local governance. These changes reduced the influence of teacher professionals on the local governance of compulsory schools. The second wave started soon after the turn of the millennium, one milestone being Norway's participation in the PISA test for the first time. Other major structural innovations constituting the second wave were the introduction in 2004 of the National Quality Assessment System (NQAS), which included PISA among other international large-scale assessments such as Progress in International Reading Literacy Study (PIRLS) and Trends in International Mathematics and Science Study (TIMSS), and the introduction in 2006 of a national curriculum reform called the Knowledge Promotion Reform (K06), which reshaped the old national curriculum by introducing learning outcomes as its central feature. Being 'a hesitant reformer', Norway introduced these NPM reforms more than a decade after similar arrangements had been implemented in the UK. The introduction of PISA, however, is seen by Møller and Skedsmo (2013) as a turning point in Norwegian educational policy, releasing the second wave of NPM reforms. These changes to the Norwegian educational policy through decades have contributed to stronger unification processes in the Norwegian educational system, making it a centralized educational system (Skinningsrud 2019; Bringeland 2022a).

PISA in the Norwegian educational structure

The PISA test is administered every third year. It is incorporated into the National Quality Assessment System (NQAS) together with other international largescale assessments and National Tests, and other student tests and questionnaires. The national random sample participating in the PISA test is composed of single individuals from various schools and school classes. Test results are therefore not representative of specific schools or class units. The test is administered by Norwegian teachers who happen to be teaching students in the random sample sitting for the test. Administering the test is part of their duties as teachers. Moreover, as part of a feedback procedure, national PISA-results are disseminated to all Norwegian schools, where school leaders and teachers are expected to 'learn from them', that is, use them as feedback to improve school leadership and teaching practice.

The importance attached to PISA by Norwegian education authorities is underlined by the ‘PISA-courses’ to which leaders in Norwegian lower secondary schools are invited. These courses are offered on a regular basis by the PISA research team at the Department of Teacher Education and School Research at the University of Oslo (Aursand 2018; Aursand and Rutkowski 2021; Bringeland 2022b). As part of the National Quality Assessment System, PISA, along with other international and national standardized tests, constitutes a part of the Norwegian educational structure: it is mandatory for schools to take part in the test; teachers must administer the test if their school is in the randomly drawn national sample; and national results are disseminated in schools, where school leaders and teachers are expected to discuss, and possibly implement the results. A previous study on reflexivity and reflexive modes have identified the prevalent reflexive mode of three school leaders in their internal conversations about the PISA test as ‘autonomous’. Meta-reflexivity was the next prevalent mode. Firstly, the school leaders are mainly concerned with administrating the test and presenting and discussing the test results for possible further implementation. Secondly, they are to some extent critical of the test, especially the validity of the test results for their school due to random sampling (Bringeland 2022b).

Theoretical frameworks guiding previous research

As mentioned, three theoretical traditions stand out as distinct in their conceptual approach to the study of teacher responses to neoliberal educational reforms. Labour process theory focuses on the increased external control of teachers’ work, resulting in loss of autonomy, deskilling and the separation of conceptualizations and execution of work tasks. The focus of this tradition was structural change in the work context.

An alternative approach, gaining ascendancy after the heyday of labour process theory, focused on changes in ‘subjectivity’ resulting from structural change. In this post-structural approach, relying heavily on Michel Foucault’s conceptualizations, subjects were considered as ‘disciplined’ and ‘obedient’, almost totally malleable by their surroundings. The post-structural approach conceptualizes the worker/teacher as a ‘subject’ but lacks a conception of ‘agency’. An analysis of English schools operating within this research paradigm refers to Foucault’s concept ‘*dispositif*’, whose clarification is the stated overall aim of the research. ‘Dispositif’ is defined as ‘*a thoroughly heterogeneous ensemble consisting of discourses, institutions, architectural forms, regulatory decisions, laws, administrative measures, scientific statements, philosophical, moral and philanthropic propositions – in short, the said as much as the unsaid*’.

(Cited from Foucault 1977, in Gordon 1980, and in Ball, Maguire, and Braun 2012, 141). By its definition, this ensemble contains structures, ideas, and material artefacts without any theoretical clarifications of how they are to be distinguished and possibly related to each other. The post-structural approach, accordingly, avoided the structure-agency debate which was ignited by studies based on the labour process, and instead introduced the concept of ‘discourse’ as a key in analysing education and educational institutions. Arguably, this represented a ‘cultural turn’ in the theoretical approach to the consequences of neoliberal educational reforms.

A third theoretical framing in studies of neoliberal educational reforms and their consequences for teachers’ work practice is the theory of professions, epitomized in the concept ‘new professionalism’, which raises the issue of recent structural change and their consequences for agency. The structural issue concerns the extent to which structural change affects autonomy at work, and the agency issue concerns the possible emergence of a new type of professionalism among teachers, produced by neoliberal reforms.

Labour Process Theory: focus on structural change, de-skilling, or up-skilling

Labour process theory originated in studies of industrial work. Accounting for the theory and its trajectory in the study of teachers’ labour process, Reid (2003) claims that the theory lost its momentum in face of both theoretical and empirical critique. Only a few central concepts have survived in subsequent research, such as ‘de-skilling’ and ‘intensification’ of work. At the theoretical level the theory was criticized for being too universalizing and deterministic, universalizing in the sense of seeing ‘scientific management’² as the only form of work control and deterministic in the sense of seeing ‘de-skilling’ and ‘intensification’ of work as inevitable in a capitalist economy. The role of (collective) agency in social development was neglected. Also, critics pointed out that ‘scientific management’ was not the predominant form of work control in education, and labour process theory was therefore not applicable to teachers’ work. Reid (2003), defending the relevance of labour process theory, despite its shortcomings, argues that when applied in the field of education more attention should be paid to the particular circumstances and contexts of teachers’ work (Reid 2003, 560).

² ‘Scientific management’ refers to Frederick Winslow Taylor’s theory of management, developed at the beginning of the 20th century, aiming to improve productivity and efficiency in industrial work by analysing and synthesizing work operations in new ways, for example time-motion studies and assembly line operations. It was criticized by ‘human relations theory’ for focusing only on physical and behavioural aspects of work, while neglecting social relations in the workplace (see Elton Mayo’s Hawthorn studies).

Admitting that de-skilling was not inevitable, and that labour process theory lacked a conception of subjectivity and agency, Ozga and Lawn (1988) incorporated agency in their study of teachers' work by providing historical examples. Their approach was, however, criticized for merely providing descriptions and subjective experiences of teachers' work without formulating a theory which included both structure and agency and how they were related. Other studies of teachers' work claimed, on empirical grounds, that structurally determined de-skilling did not grasp the current situation at all, since teachers were developing more skills and experiencing greater autonomy in their work than ever before (Reid 2003, 563).

Neoliberal policy for the public sector, that is NPM, is based in 'public choice theory' developed by James Buchanan in the 1960s. This theory postulates that public sector workers, bureaucrats, and professionals, as well as everyone else, are pursuing their self-interest and thus perpetuating inefficiency under the guise of serving the common good (Hodge et al. 2018). This assumption fuelled caution among politicians who, according to the theory of public choice, would see teachers in a new light as the maximisers of self-interest rather than the public good. This idea promoted a new policy towards teachers (Connell 2009; Robertson 2012). The consequences of a policy informed by public choice theory were that public sector workers, bureaucrats, and professionals were targeted for maximizing their own interests (Hodge et al. 2018).

The assumption that teachers were a hindrance to raising educational standards, notably due to their pursuit of self-interest was, however, challenged when the OECD in the 2005 publication 'Teachers Matter: Attracting, Developing and Retaining Effective Teachers' (OECD 2005) argued that high quality teachers was an important precondition for high quality education, playing a potentially vital role in preparing students for the new knowledge economy. The OECD recommendation of upgrading teachers' qualifications and providing good work environments for them apparently defied the trend, which labour process theory had predicted, towards the de-qualification and de-skilling of teachers. There was a renewed interest in what constituted a 'good teacher', and in several countries, for example England and Australia, lists of competencies were worked out indicating the kinds of knowledge and skills effective teachers would have to develop (Connell 2009; Robertson 2012).

However, long before OECD called for an upskilling of teachers, many countries, for example Norway, Finland, and Sweden, had upgraded their teacher education in what has been described as an academization process (Bergmark and Hansson 2021). Before the turn of the millennium teacher education in these Nordic countries had become upgraded to university level programs.

In 2010 Sweden made a decisive move to make education scientific, passing an act which requires teaching practice to be based in ‘scientific knowledge’ and ‘proven experience’, the latter referring to examples of ‘best practice’ and ideas of ‘what works’. Though reportedly difficult to implement, this act confirms a policy determined to upgrade teachers’ competencies (Hansson and Erixon 2020; Bergmark and Hansson 2021).

Paradoxically, these upskilling efforts, and in particular their emphasis on teachers’ ability to understand and apply research results, can be a kind of de-skilling. Bergmark and Hansson (2021) point out that the Swedish 2010 act which requires education to be scientific by applying scientific knowledge, may in fact result in de-skilling, since teachers are called to apply scientific knowledge that has been developed by others. Thus, they may still be trapped in a work situation where the conceptualization of their work is separated from its execution.

Likewise, recent Canadian research on the professions of engineering and nursing show that these professions’ knowledge base is changing from being a clearly defined body of knowledge specialized for the professions to becoming a hybrid and more heterogenous field of knowledge which could, according to Adams and Sawchuk (2020), indicate a broader process of de-skilling in these professions. Thus, labour process theory, defended by Reid (2003) and others (Connell 2009), which thematizes the process of de-skilling, is not irrelevant to current occupational development. De-skilling may appear in new guises. Still, the theory’s one-sidedness in only thematizing structural change and not the varied responses of agency to these changes is a critique that points to an absence in the theory, which still remains.

Post-structuralism: focus on teachers’ subjectivities

Post-modernist and post-structuralist approaches, challenging ‘grand narratives’ and ‘totalizing’ accounts of social development, grew strong in the 1980s and 1990s. The French philosopher Jean-François Lyotard (1984) in his translated publication ‘The Post-Modern Condition’ claimed that knowledge was no longer legitimated by the aspiration to find truth. Its new source of legitimacy was ‘performativity’, that is, its contribution to the economy, and effective production in the technological sense of having a favourable input/output ratio. The collapse of the modern epoch’s ‘grand narratives’ of progress, or even emancipation, had lost their credulity, and this called for *‘petits récits’*, that is, localized narratives on a smaller scale. In the social sciences this critique coincided with increased attention given to uneven development, and contextual variations of developmental trends.

The post-structuralists' charge against 'grand narratives' of being too universalising and not accounting for the particularistic was detrimental to orthodox labour process theory. A leading voice in the Anglophone research literature on the consequences of neoliberal educational reforms for teacher subjectivities based in post-structuralist ideas is Ball (2003, 2016). His major source of intellectual inspiration is Michel Foucault. In Ball's widely read and much quoted article from 2003, 'The Teacher's Soul and the Terrors of Performativity', he considers performativity, the market and managerialism as 'policy technologies' included in 'the package of neoliberal educational reforms', attributing their promotion to international agencies like the World Bank and the OECD. Ball also wrote a book titled 'Global Education Inc.' in which he traces the foundations and think tanks that operate on a global scale with the explicit purpose of promoting neoliberal policies in all domains of society, including education (Ball 2012).

Ball sees performativity as 'a new mode of state regulation which makes it possible to govern in an "advanced liberal" way'. Essentially it requires self-disciplining, as individuals must organize themselves by responding 'to targets, indicators, and evaluation' (Ball 2003, 215). Emphasising the uniformity of effects on individuals, the new imperative according to Ball is to set aside personal beliefs and commitments and instead live an existence of calculations. He describes the new performative worker as 'a promiscuous self, an enterprising self, with a passion for excellence' (Ball 2003, 215). Not only does the new framing of teachers' work entail an influence on what teachers do, it also changes teachers' identities. The 'inner-life of the teachers' is profoundly influenced by the policy 'technologies of marketization, managerialism' and demands for performativity (Ball 2003, 226). However, admitting to variable individual responses to the situation, Ball makes a distinction between those who see neoliberal reforms as an opportunity to 'make a success of themselves', and others, to whom it may result in 'inner conflicts, inauthenticity and resistance.' (Ball 2003, 215).

Ball, Maguire, and Braun (2012) provide a concrete example of their theoretical approach and empirical findings in '*How schools do policy: policy enactments in secondary schools*'. Reporting on ethnographic case studies of four English secondary schools, Ball and his team tell how teacher audiences attending presentations of their research frequently asked whether they had encountered resistance against neoliberal educational policy in the schools they studied. Their reply was 'very little'. As they see it, in their research they had to choose between focusing on variances and difference versus studying 'the colonization of practice by performativity' (Ball, Maguire, and Braun 2012, 150). This choice resembled an optical figure/ground dilemma, where one may see either an urn or two faces in profile, depending on

what is seen as the figure and what is seen as the background. Despite opting for studying uniformity, or the total ‘colonization of practice’, the authors do report observations of differences among teachers’ reactions. They observed policy enthusiasts, critics, and receivers, as well as senior teachers for whom policy responsibility is a ‘career move’. ‘Thus, some of those who work in schools are “policy carriers” and some “policy careerists”.’ (Ball, Maguire, and Braun 2012, 145).

Ball and his team admit to not having raised the question of the relationship between power, agency and the space for alternatives (Ball, Maguire, and Braun 2012, 149). They claim that resistance or refusal of policies among teachers were only found in the form of ‘discontents, murmurings, indifference and disengagement’, which to some extent is ‘free-floating, rather than systematic’ (Ball, Maguire, and Braun 2012, 143–144). They conceptualize these responses as a kind of role distancing, i.e., ‘disdainful detachment of the performer from a role he is performing’ (Goffman 1961: 110, referred in Ball, Maguire, and Braun 2012, 150). Ball and his team, however, recommend that more work should be done on detailing the micro-politics of resistance and ‘resistance within accommodation’ (Ball, Maguire, and Braun 2012, 150).

The unwillingness, or perhaps inability, of Ball and his team to analyse and explain the variety of responses to neoliberal reforms might have something to do with the theory they use. The concept of social structure is absent from their analytical toolbox, as is a more elaborate theorizing of agency. Instead the authors aim to identify ‘a set of master discourses that define schooling’ (Ball, Maguire, and Braun 2012, 140). These discourses are identified as ‘learning’, ‘curriculum’ and ‘behaviour’, and are ‘what makes the school’. The challenge, which Ball and his research team saw in their study of ‘policy enactment’ in English secondary schools was to ‘join up politics and practices’ by ‘the concept of discourse’ (Ball, Maguire, and Braun 2012, 140)—an approach that overlooks social structures in the schools, which is produced by macro-level politics - and how these influence local practices.

Ball’s conceptual framing in studying the effects of neoliberal educational reforms has been widely influential in international educational research. Studies in countries as far apart as Australia, the US and Sweden have used Ball’s conceptions of ‘policy as discourse’ (Lewis and Hardy 2014; Holloway and Brass 2018).

Australian researchers studying teachers' subjective experiences within the post-structural framework give an account of how high stakes testing, combined with target setting and the schools' struggle for a positive reputation, influence the subjectivities of those who work and learn in schools. Lewis and Hardy (2014, 245) refer to Ball's assertion that national 'policies discursively constitute the teacher as a performative subject – not merely changing what teachers *do*, but also ultimately who teachers *are*' – and they seem alarmed by the idea that teachers are totally determined by the structures they inhabit. They conclude that 'teasing out alternative practices and dispositions is important work for thinking such practices differently.' (Lewis and Hardy 2014, 261).

Jessica Holloway and Jory Brass (2018), also working within Ball's post-structuralist paradigm positing 'neoliberal subjectivities', find that new generations of teachers are more accepting of neoliberal structuring of their work. They compared different generations of American teachers by studying two groups through and after their teacher training. One group of five teachers was followed for several years (2002–2005) through their one-year internship and first year of licenced teaching. Another group of seven middle school teachers was followed for two years about ten years later (2013–2014). The period of the first investigation coincided with the implementation of the No Child Left Behind (NCLB) Act during the Bush administration and the second with the implementation of 'Race to the Top' (RTTT) under the Obama administration.

Comparing the teachers' subjectivities/responses to the two reforms indicated a shift in the two groups concerning their 'subjectivities' and in their sense of professionalism. The teachers in the first group considered the steering mechanisms of the NCLB reform as an external intrusion on their autonomy, professionalism, and practice, while the second group accepted the RTTT's accountability mechanisms as 'the very modes by which they knew themselves and their quality' (Holloway and Brass 2018, 361). Holloway and Brass draw the conclusion that over time there is a 'collapse between the governed (i.e., the teachers) and the government (i.e., accountability mechanisms)' (Holloway and Brass 2018, 361). A normalization of the managed and marketized teacher has taken place. Holloway and Brass (2018, 361) take their findings to confirm Ball's claim that accountability reforms 'produce new kinds of teacher subjects', a finding which they claim has been further corroborated by Evetts (2011) and Anderson and Herr (2015). Evetts is a central proponent of the notion of 'new professionalism'.

Professionalism – ‘Old’ and ‘New’

The consequences of neoliberal educational reforms for teachers have also been studied from the perspective of changes in the professions and to professionalism. This theoretical framing overlaps with labour process theory in its focus on teachers’ loss of self-determination and autonomy in their work, resulting in de-skilling and de-professionalization. Studies that focus on teachers’ professionalism have developed conceptions of various ‘new professionalisms’, emerging as a consequence of structural change resulting from neoliberal reforms.

In social theory, the professions have received special attention as a group of occupations having more autonomy in their work than other groups. Professionals’ work has been delineated from other types of work by its reliance on specialized fields of knowledge, exclusive access to job markets through education and licencing and, in Talcott Parsons’ classical account, the professions are seen to adhere to generally accepted social values that serve the public interest. The historically established professions are in law, medicine and theology, but new professions have emerged such as teaching, nursing, and social work, which are called semi-professions by some authors since work autonomy in these fields has been less pronounced (Etzioni 1969).

Studies have indicated that the distinctiveness of the professions compared to other occupations has diminished, and the emergence of new kinds of professionalism is suggested (Svensson 2006; Evetts 2011). Neoliberal reforms are seen to threaten the traditionally distinctive characteristics of the professions such as relations of trust among the professionals themselves, between practitioners and clients, and between practitioners and employers. Likewise, egalitarian relations seem to have given way to legal-rational bureaucratic rationality and market-based competition, commercialism, and the commodification of services.

Evetts (2011) distinguishes between ‘old’ and ‘new’ professionalism, alternatively called ‘occupational’ and ‘organizational’ professionalism respectively, with special reference to public services. New/organizational professionalism entails that the practitioner identifies with and is loyal to the organization and organizational interests over normative commitment to professional values. Adams and Sawchuk (2020, 91) suggest that a hybrid type of professionalism is emerging. Hybridization, in their opinion, is a result of increased control of professionals’ work and might in essence be an aspect of de-professionalization.

The various suggestions of emerging new types of professionalism continue the emphasis on uniformity in individual responses to structural change. At most they point to the division of

professionals into two categories, 'old' and 'new' professionals, the new professionals being more accepting of and adapting to current structural change.

Already Troman (1996, 474) early study of English primary school teachers' reactions to the introduction of neoliberal reforms distinguished between 'old' and 'new' professionals depending on the extent to which they accepted and adapted to their new work context. Referring to Pollard et al. (1994) and Hargreaves (1994), Troman mentioned adaptive teacher responses indicating the emergence of a new type of professionals who consider the reforms as 'necessary measures to remedy deficiencies in the system' (Troman 1996, 474). This positive view of an emerging new professionalism fits well with the 'new professionalism' mentioned by Hargreaves (1994) which breaks teacher isolationism and promotes collaboration both internally in the school and with external groups such as parents and the local community. Another adaptive teacher response mentioned by Troman is the 'new entrepreneurs', who fully accept the new changes made in schools (Troman 1996, 474).

Troman also foreshadows Archer's approach when studying agents' reactions to their structural context by drawing attention to other studies of English schools, which show that individual agency is active in producing a variety of responses and reactions to changes in the work environment and to new definitions of teachers' work. He points out that 'teachers filter the policy of reforms and change through their existing professional ideologies, perspectives, and identities' (Troman 1996, 474). He lists a range of different reactions that have been reported, from compliance and accommodation to resistance and rejection. The missing element in these empirical accounts of divergent teacher reactions to current reforms, however, is a theory which can explain the variety of reactions. Is it by chance that some teachers become 'new professionals' or 'new entrepreneurs' and others take early retirement? When Troman (1996) suggests that variances in professional identities, ideologies and perspectives cause different adaptations to new structures, he does not specify which types of identity, ideology, and perspectives predispose a teacher to become a 'new entrepreneur' rather than seeking early retirement.

Research on new professionalism has also been carried out in the Scandinavian countries. One study, comparing Norwegian and Swedish teachers, indicates that despite variance within each country, teachers in these two countries diverge as groups (Helgøy and Homme 2007). Referring to Svensson (2006), Helgøy and Homme (2007) define old professionalism as 'professional practice relying on formal educational credentials, the monopolizing of certain occupations based on licensing', and new professionalism as 'competencies which are more

personal, implicit, individual and related to context, tasks and actual performance'. The capability of each professional may increase his or her autonomy and responsibility. In other words, old professionalism is oriented to the collective of professionals, while new professionalism is more individualistic. Their different orientations may correspond to different accountability regimes; holding the profession accountable promotes old professionalism, while holding the individual accountable encourages new professionalism (Helgøy and Homme 2007, 234).

Norwegian teachers on the whole practiced old professionalism in the sense of experiencing less individual autonomy in their teaching practice than Swedish teachers. Swedish teachers had to a greater extent adopted new professionalism, that is, loyalty to their employer over loyalty to their profession (Helgøy and Homme 2007, 232).

Similar results from comparative studies of Norwegian and Swedish teachers are reported by Ingrid Carlgren and Kirsti Klette (2008), who found that national policy documents from the two countries in the 1990s signalled different expectations regarding teachers' work. In Sweden teachers were portrayed as 'professionally empowered curriculum makers', while Norwegian teachers were seen as 'curriculum deliverers' (Carlgren and Klette 2008, 129). In the Norwegian National Curriculum from 1997, the state prescribed – in detail – the content of schooling. However, in agreement with the findings of Skedsmo and Mausethagen (2017), Norwegian teachers did not consider the Norwegian national curriculum as a constraint on their professionalism. Instead, they experienced it as an enablement. They could spend more time on teaching the curriculum, while Swedish teachers had to discuss and select curriculum content, choose teaching methods, formulate learning goals, and develop criteria for marking. With an expanding private education sector, Swedish teachers also had to attend to competition with other schools and negotiate their individual salaries.

Summing up

Studies on the effects of new educational structures on teachers' reactions and agency under the current 'neoliberal order' (Gerstle 2018) reveal disagreement among researchers. There is disagreement on the theoretical framing of such studies, and findings vary between countries regarding teachers' adaptation or resistance to their new structural environment. The divergent findings between countries could reflect that 'the neoliberal order' and 'neoliberal reforms' are not the same everywhere. The early labour process theorists tried to explain why the de-skilling and proletarianization of teachers' work did not stimulate collective protest. More recent post-

structuralist approaches have emphasized the wholesale changes in teachers' subjectivities and the production of 'neoliberal subjects', who are compliant, disciplined and employ 'technologies of the self' to cope with demands. The theory of professions points to another type of adaptation, 'new professionalism', which entails a shift in loyalty from the professional group to the organization where they are employed.

Studies informed by all these theoretical frameworks mention variations in individual responses, but without being able to refer to theory that can accommodate and explain why, for example, external control in terms of a fixed national curriculum and standardized tests as part of accountability schemes is seen by some as constraining their work while others see the same controls as facilitating. Reported reactions among teachers to the restructuring of their work include both taking early retirement and becoming 'careerists'. What seems to be particularly lacking in theories guiding research in this area is the ability to account for differences in individual teachers' reactions to the same given structures and circulating ideas.

Archer's theoretical framework: structure, culture, and (reflexive) agency

Archer's concept of reflexivity seeks to resolve the structure/culture/agency issue in social theory. This issue has been prominent in theoretical debates, resonating in areas of applied research, in the sociology of education, and in studies of policy implementation, i.e., how educational reforms, creating structural and cultural change at the macro level, impact practices at the micro level of schools and classrooms. Introducing the notion of reflexivity as mediating structural and cultural effects on agency entails explaining the course of action taken by agents as caused both by their structural and cultural situation and internal deliberations about how to reconcile their structural and cultural context with their personal concerns and values. Archer, in her first full volume exploring the concept of reflexivity titled 'Structure, Agency and the Internal Conversation' (2003), starts out by posing the fundamental question: 'How does structure influence agency?' Noting agreement among theorists that structure does not fully determine human agency, she suggests that there must be something else 'involved in the process' and she suggests that this 'something else' might be the 'properties and powers of agents themselves' (Archer 2003, 1).

Archer notes that in the critical realist research tradition the concept 'conditioning' has been used about the structure/agency relationship. That is, structure is said to 'condition social action and interaction'. Archer also postulates that the same logic applies for culture (Archer [1988] 1996; 1995, 193; 2003, 3). However, the concept 'conditioning' tends to give primary emphasis

to structure and culture, without accounting for how agency contributes to the action outcome. The question is how agents respond to structural and cultural conditioning, and what kind of processes are involved in producing subsequent action. Roy Bhaskar’s fundamental claim that ‘the causal power of social forms is mediated through human agency’ (Bhaskar [1979] 1998, 28), says nothing about the mediatory process and does not conceptualize the mediation. Archer’s contribution to clarifying the process of mediation is to introduce the concepts of ‘reflexivity’ and ‘reflexive modes’ (Archer 2003, 342).

Figure 1: Internal conversation and pursuit of the good life (Source: Archer (2007, 89)).

Defining and dovetailing one’s concerns →	Developing concrete courses of action →	Establishing satisfying sustainable practices
(Internal goods)	(Micro-politics)	(<i>Modus vivendi</i>)

Two major implications of incorporating reflexivity in the structure-agency and culture-agency relations are:

- i) a given structural [or cultural] context is not a *constraint* or an *enablement* per se; whether it is a constraint, or an enablement depends on the agent’s projects, i.e., what the agent wants to promote or obtain, that is, what the agent’s fundamental concerns are (Archer 2003, 5–6).

Neoliberal reforms in the education sector, such as the introduction of accountability mechanisms based on measured performance (test-based accountability, TBA) are generally seen to increase constraints on teachers’ work. Target setting and the use of readymade tests have, moreover, been seen to result in the de-skilling of teachers, or at least to their decreased autonomy and freedom to exercise discretionary judgement. However, this is not necessarily the case. External controls may be experienced as supporting the learning process when they are seen to promote activities that increase students’ motivation and learning.³

- ii) agents’ response to the same structurally [or culturally] defined situation are not uniform because they may activate different modes of reflexivity; since agents differ in what is their *predominant reflexive mode*, which is founded in their ultimate concerns, what they care about most. Differences between

³ Interestingly, Norwegian research on the reactions among teachers and principals to National Tests and a fixed National Curriculum indicates that increased state control is not necessarily experienced as a constraint by teachers (Carlgren and Klette 2008; Skedsmo and Mausestagen 2017, 178).

agents' ultimate or fundamental concerns contribute to agency responding in different ways to a given structural [or cultural] context. (Archer 2003, 41, 343–344).

Each mode of reflexivity entails a different stance, or orientation, towards society and its constraints and enablements. A stance (reflexive mode) is a mechanism at the personal level, which tendentially regulates relations between persons and their society. It is an overall response pattern to the totality of structural powers encountered by a person, and thus, it is a particular way of 'being in the world' (Archer 2003, 342).

By activating a particular mode of reflexivity, stance, or orientation, subjects attempt to regulate their personal-societal relationship (Archer 2003, 355). Depending on which stance is activated, the internal conversation leads the person-society relationship in different directions – 'thus, articulating the precise form of the micro-macro-link' (Archer 2003, 349). '[T]he "stance" is ventured as a generative mechanism, at the personal level, with the tendential capacity to regulate relations between the person and her society.' (Archer 2003, 343). Agents' stances towards society, their mode of reflexivity, has both internal effects on the agent and external effects on the agent's environment. Internal effects on the agent consist in agents' (i) prioritising personal concerns and developing personal projects, and their (ii) establishing and consolidating orientations (stances) in encountering constraints and enablements (Archer 2003, 349).

Figure 2 below is our condensed presentation of Archer's model of the relationship between agents' various modes of reflexivity, their ultimate concerns, stances towards structural and cultural constraints and enablements, and external effects upon the agent's immediate environment from activating these different modes.

Figure 2: Modes of reflexivity - their basis in ultimate concerns, expressed as stances towards structural enablements and constraints, and their external effects (Source: Archer (2003))

Mode of reflexivity	Ultimate concerns	Stance towards structural and cultural enablements and constraints	External effects
Communicative	Interpersonal relations (family and friends)	Evasive	Social reproduction; social integration (morphostasis)
Autonomous	Performative achievements (work)	Strategic	Goal achievement ('regional' social

			change; ‘regional’ morphogenesis)
Meta	Transcendent ideals	Subversive	No obvious effect (Negligible contextual effects but ‘reactivation of <i>Wertrationalität</i> in the cultural system) ⁴
Fractured ⁵	Vacillating concerns	No consistent stance	No effect due to lack of agency

Communicative reflexives, whose ultimate concerns are ‘interpersonal relations’ (family and friends), tend to adopt a stance of ‘*evasion*’ in relation to structural and cultural enablements and constraints (Archer 2003, 192, 342). The external effect of this stance, on a micro scale, is social reproduction. By not undertaking ambitious projects, communicative reflexives manage to avoid constraints. They also tend to avoid enablements facilitating their social advancement. ‘Communicative reflexives’ are ‘collectivists’ towards the social in the sense that concerns and context are inseparable. They contribute to social reproduction by their strengthening of social integration, i.e., agreement among members of their network.

Autonomous reflexives, whose ultimate concern is performative achievements (work), meaning task performance ‘at a level which satisfies external standards of assessment’ (Archer 2003, 265–266), tend to adopt a ‘*strategic*’ stance towards their structural and cultural context; its enablements and constraints. They are ‘accommodative’ towards the social, since to them, context is a means towards the realization of their concern, which is task performance. The external effect of this stance is increased goal achievement in whichever sector they are located, which means that they contribute to social change (morphogenesis) in the section of society where they work.

The ideals and concerns of the ‘*meta-reflexives*’ transcend present social reality. To the meta-reflexives, the context is always inadequate in meeting their ideals and concerns (Archer 2003, 353). However, meta-reflexives tend to experience their job as an activity in which they can express their commitment to fundamental personal values (Archer 2003, 258). Nevertheless, meta-reflexives ‘pursue cultural ideals that cannot be accommodated by the current social structure and the array of contexts it defines’ (Archer 2003, 361). Because of their fundamental

⁴ (see Archer 2003, 360–361)

⁵ Fractured reflexivity is included in this table, though it refers to the lack of reflexivity, i.e., reflexivity is fractured and inoperative.

critique, they become *subversives* of the structural status quo and resist any ‘deal’ with it. At the micro level their actions may have negligible direct impact. However, one important effect is at the macro-level of the cultural system. By personifying their utopian ideals of truth and goodness, meta-reflexives are upholding cultural ideals, which may otherwise ‘sleep on in the Universal Library of Humankind’ (see Archer [1988] 1996, 104). ‘The meta-reflexives awaken these ideals and re-present them to society’ (Archer 2003, 361). Archer connects meta-reflexivity to Max Weber’s notion of *Wertrationalität*, that is, actions that have value in themselves and are not a means to something else (Archer 2003, 355–361).

The most pronounced feature characterizing ‘*fractured reflexives*’ is that they have not developed or are unable to consistently adopt any determinate personal stance towards their social surroundings (Archer 2003, 343). This lack may be caused by various circumstances, emotional distress etc.

This research is illustrative and exploratory and primarily aimed at demonstrating that Archer’s conceptions of reflexivity and reflexive modes are usable in analyzing a teacher’s internal conversation on PISA-related issues. If we can showcase the applicability of Archer’s concepts, i.e., if we are able to identify a teacher’s different reflexive modes and pinpoint the predominant mode, we will have illustrated that teachers’ subjectivity is not just an internalization of neoliberal principles and thoughts (the neoliberal subjectivity) or a facile adaptation to the aims of the organization or the school (new professionalism), but that teachers’ subjectivities are more complex, and personal ideals and concerns play an essential role in determining the course of action followed in encountering obstacles and demands from the school context. Analyzing just one case of course precludes any attempt at generalizations, for example about what kinds of reflexive modes are prevalent among Norwegian teachers. We have, however, at the end of our analysis offered a hypothesis about why the special constellation of reflexive modes discovered in this one case might be prevalent in the Norwegian context, primarily due to the characteristics of the Norwegian educational system. For these reasons our research questions are:

Which modes of reflexivity are activated in a mathematics teacher by the PISA test? Which mode(s) of reflexivity predominate(s)?

Methodology: selection of case and types of data

The mathematics teacher works in a lower secondary school. He is a regular member of the school staff and is not involved in the school leadership. He is, however, a member of two staff teams, one specific to the grade level he teaches and one specific to his subject specialism, mathematics.

Max was selected for the interview because he had first-hand experience in administering the PISA- test. He could therefore be expected to talk not just ‘from the top of his head’ about PISA, or to express some stereotype opinions about the test. Having administered it, he would know something about its content. Also, compared to teachers who had not been personally involved in administering the test, he was more likely to have reflected on its use and utility in the Norwegian educational setting.

Two types of data were collected⁶ in April 2020: i) quantitative data generated by the Internal Conversation Indicator (ICONI), which is a short questionnaire consisting of thirteen items of the Likert scale type, measuring degree of engagement in different modes of reflexivity to establish the dominant general nonreferential reflexive mode, and ii) qualitative data collected in a semi-structured Skype-interview focused on his opinions on PISA and its function in the policy and practice of Norwegian education to establish his reflexive modes activated by PISA.

Results

Quantitative data: Internal Conversation Indicator (ICONI)

The instrument ICONI, which provides a quantitative measurement of persons’ engagement of the various reflexive modes, is developed by Archer and her associates. The purpose of the index is primarily to serve as a screening device for selecting persons who practice a distinct dominant mode of reflexivity for further interview. In developing the questionnaire two requirements were that it should be quick to administer and items should be readily understandable. Besides, items should not contain any form of referential specificity which would preclude its use in other countries. When the questionnaire is administered, it is introduced to participants as ‘an investigation of the processes of decision-making in everyday life’ (Archer 2007, 331).

Based on results from previous interview studies on reflexivity (Archer 2003), subjects were expected to obtain scores on more than one reflexive mode, but to various degrees. Thus, the

⁶ The project and data collection has been approved by the Norwegian Centre for Research Data (NSD).

mode with the highest mean score would qualify as the dominant mode (Archer 2007, 331). In developing the questionnaire various versions were subjected to a series of trials, and the final version consists of thirteen items of the Likert scale type where the respondent may choose a score from 1 to 7 on each item. Consequently, each person tested with the ICONI may be described by a ‘reflexivity profile’ consisting of a mean score for each mode of reflexivity.⁷

The ICONI was administered to our interviewed teacher, Max. Figure 3 shows that his highest mean score was on the ‘autonomous reflexive mode’, which indicated that his major personal concern was task performance. Added to the thirteen Likert scale items was an open question, which asked him to list the three most important areas of his life, ‘those that you care about deeply’ in a chronological order from 1-3, where 1 is of highest importance. Here, Max mentioned his relationship to his family (1), his relationship to his friends (2), and genuine interests (based on inner motivation) (3) as the most important areas of his life.

The ICONI scores indicated that Max engaged in ‘meta-reflexivity’ as his second strongest mode of reflexivity, and to some extent in ‘communicative reflexivity’. He also obtained a score on ‘fractured reflexivity’. Thus, Max’ ICONI profile included scores on all modes of reflexivity. At first sight, there was a puzzling discrepancy between his scores on the Likert-scale questions and his reply to the open question, where he did not include ‘work’ as one of ‘the three most important areas of his life’. In our ‘Discussion’, we will argue that Max was referring to his work when he mentioned ‘genuine interests’ as one of the most important areas of his life.

Figure 3: Reflexivity profile of the interview subject on the ICONI (Internal Conversation Indicator)

Reflexive modes	ICONI mean scores
Autonomous reflexive	4,6
Meta-reflexive	3,6
Communicative reflexive	3
Fractured reflexivity ⁸	2

⁷ For further details about how ICONI was constructed, see Archer (2007, 326ff) (Methodological appendix).

⁸ See previous footnote 5.

Qualitative interview data and analysis⁹

The interview started with a question on how Max got to know about the PISA test. To this he replied that it was through his work as a teacher.

Reference 2

PI-I12: Do you remember how you got to know about PISA?

PI-R12: Oh my God, it was probably through work. I don't remember if I knew so much about the PISA survey before I was exposed to it, to put it that way.

Max further clarified his involvement with the PISA test as a random happening as it was only every third year that he was teaching 10th grade students, who are the target group for the PISA test.

Reference 1

E-R3: Well, I have to say that personally, I had to think about when I was engaged in PISA. It was a bit random, as a secondary school teacher at our school you follow a cycle, every three years you have 15-year-olds, the 10th grade. You are “lucky” to administer the test if it falls on that year you have a 10th grade class.

By making these initial remarks Max conveyed that it was not his choice to administer the PISA test. This was a result of circumstances beyond his control. He distanced himself from the test.

Autonomous reflexivity

The ultimate concern of ‘*autonomous reflexives*’ is performative achievements (work), and they tend to adopt a ‘*strategic*’ stance towards their structural and cultural context; its enablements and constraints. Because context is a means towards realizing their concerns, they accommodate to their context. The external effect of this mode (stance) is increased goal achievement. This entails that autonomous reflexives may contribute to social change (morphogenesis) in the section of society where they are located.

Tackling the situation in his school with an increasing number of standardized tests, Max explained that he had initiated the practice of ‘simulated test situations’. He used previous test items from national tests and exams to prepare students for upcoming tests. He justified this practice by saying that it would enable students’ achievement and remove their insecurity when taking tests. He claimed that choosing to practice what some would call ‘teaching to the test’

⁹ Reference numbers below refers to the number of selection coded/cited data to a node (category).

he succeeded in familiarizing the students with the test situation and made them feel more at ease when they were tested. He underlined that he had promoted this practice to the school administration and the team level, which resulted in it being adopted as a normal practice in his school.

Reference 2

E-I25: Have you used tasks from national and/ or international tests to prepare students for upcoming tests?

E-R25: Yes.

E-I26: What tests?

E-R26: National tests and exams.

E-I27: Why do you think it is right to practice on previous test items before a test is to be conducted?

E-R27: [...]. Eh, mathematical it is to, okay then, you can say it's to enable the students, or you can say, what should I call it, warm them up to have tests, get used to the type of assignment, eh, to remove insecurity before the tests are to be conducted.

E-I28: Is this a choice, a decision made at the team level, or administrative level?

E-R28: No, it is a choice I make. But there is also... I have probably been promoting it so that it is at least said that it should be done from the administration - and also at the team level.

These excerpts from the interview indicate that Max adopts the stance of an autonomous reflexive in administering tests in general. He is goal oriented in the sense of wanting his students to perform optimally on tests, giving them the opportunity to practice test-situations, and thus removing their sense of insecurity. He adapts to the situation in his school where students are regularly subjected to standardized tests, which he and the other teachers are obliged to administer. Practicing autonomous reflexivity, he uses his work context to exercise his performative skills as a teacher, obtaining optimal results from his students. His involvement in goal attainment on behalf of his school, i.e., increasing test scores, is further underlined by his initiative to make 'simulated test situations' a regular part of the school's program. However, PISA was not a test used in his simulations.

Meta-reflexivity

Meta-reflexivity means being reflexive about one's own acts of reflexivity (Archer 2003, 255). For example, reflecting 'why did I think this is Friday, when in fact it is Thursday?' (Archer 2003, 255). In internal conversations meta-reflexivity is not about whether propositions are true or false, right, or wrong, but why they have been uttered (Archer 2003, 255). One must distinguish between meta-reflexivity, which includes all acts of self-monitoring (Archer 2003, 256), and the shared characteristics of meta-reflexives (whose dominant mode is meta-reflexivity) some of which are that they are social critics and critical of themselves as persons and the lives they lead. Their criticism of society and themselves is caused by their ideals and that no existing social arrangements nor their own lives approximate to these ideals (Archer 2003, 258). Meta-reflexivity is adding an extra loop in one's internal conversations, i.e., reflecting on one's own reflexivity.

'Meta-reflexives' are often motivated for their work by a sense of 'vocation'. Commitment to a 'vocation' entails 'a subjective investment of the self', a personification of ideal qualities associated with a specific type of work (Archer 2003, 266–267). Having a 'vocation' is different from just occupying a role, and not all teachers experience teaching as their 'vocation'. They may have chosen to become teachers for quite different reasons such as family traditions, expediency, or practical considerations. Meta-reflexives are idealists constantly seeking a better fit between who they try to become and their social environment which, to various degrees, permits their expression of it (Archer 2003, 258). Meta-reflexives tend to adopt a *subversive* stance towards experienced enablements and constraints in the sense that they are willing to 'pay the price' for pursuing actions that will not be rewarded by society, thus subverting the causal powers of society (Archer 2003, 289). Their ideals make them into social critics since nothing around them measures up to their ideals. They represent 'the conscience of society' (Archer 2003, 274).

Some sections of the interview indicated that, to a considerable degree, Max also engaged in meta-reflexivity. This stance was expressed when he engaged in a general assessment of the PISA test. He understood the test to be a ranking instrument, the purpose of which he considered meaningless (compared to providing feedback for students' learning). In the beginning and midway through the interview he expressed a negative opinion of the test. However, towards the end, in a typically meta-reflexive way, he reflected on his own reflections and tried to understand and explain his own negative reactions to the PISA test.

Reference 1

E-R30: [...] all schools, all the principals that I know of, use the results [on standardized tests] to show the rankings and how they relate to other schools and municipalities, and in relation to the country as a whole. Even the PISA surveys are used in relation to our rating in the global context. Which is totally, even more meaningless.

Reference 4

PI-I7: How do you understand the intention of PISA?

PI-R7: A tool for rankings.

PI-I8: A tool for rankings.

PI-R8: Yes. Quite simply. I think I recall that there is a place that, well, that this is a survey being taken all over the world, and there are a lot of different political governance we have around and mindsets, that is, hello, it is really just nonsense, the whole PISA survey.

Reference 6

PI-I41: How do you experience your competence in interpreting PISA results?

PI-R41: Eh, to interpret the results.

PI-I42: Mhm.

PI-R42: Do you think in a statistical way?

PI-I43: Yes, for example.

PI-R43: No, that would work out fine. I should be able to do so.

At the end of the interview, Max reflected on his own reflection saying that at the start of the interview he might have been too critical about the PISA test, as he was colored by what happened when the results from the first test were announced in 2001. At that time, the Education Minister Kristin Clemet fronted a shift in Norwegian educational policy, which resulted in a stronger centralization of the system (Bringeland 2022a). Despite being critical of the test, Max confirmed that he was able to interpret PISA results.

Reference 5

PI-R12: [...]. I'm probably colored and negatively influenced by the ravages Kristin Clemet did in her time. This is something that is still stuck in my memory and that I cannot get over. She started, in my opinion, to destroy the Norwegian school when she came to power. So, ehm, that will probably be the contributing cause to my attitude to this test, I think. [...].

Reference 2

E-R31: After all, it was PISA that started her and K06 [the new National Curriculum of 2006].

Reference 6

PI-R53: I may sound very strict, or too oppositional, or what should I call it. There are tasks in the PISA survey that may be individually relevant and exciting for students' development of competence, but there is more, there is more, what should I say. It is the method that is a problem to me. That this should be a test, that is, a measuring instrument.

In Max' opinion, the PISA test is resisted by the teacher professions, and its continued use in the Norwegian education system is a result of the exertion of power from the top, serving the interests of politicians rather than the concerns of the teacher profession.

Reference 8

PI-I56: Would you say that the PISA survey is better suited at the political level than at the level of teaching practice, in formative assessments, or evaluation?

PI-R56: I think if you had asked, if you let the teachers in Norway decide, it would have ended long ago. To be honest.

PI-I57: So, you think it is due to the exertion of power that the PISA survey has not been terminated?

PI-R57: Clearly.

Reference 8

PI-R55: [...] the politicians need arguments for making changes, and they probably don't trust the method, the methodology, that what we practice today is good enough. They also think such a study will confirm that they were right about this. So, the educational policy situation we have in the country is miserable now. It shows, after all, that they

use the PISA survey to bang on their chest, so that politicians no matter where they stand, eh, would welcome such an argument.

This section of the interview exemplifies that Max also engages his meta-reflexivity in his internal conversation about the test. He carefully adds an extra loop in his reflections on the test, reflecting on his own attitude - whether it was influenced by his negative view of the educational policy which it is part of, and the way in which it is used by politicians. He is critical of not only the PISA test, which he considers to be an irrelevant ranking instrument, but also of the national education policy in general.

Communicative reflexivity

The ultimate concerns of '*communicative reflexives*' are 'interpersonal relations' (family and friends). They tend to adopt a stance of '*evasion*' in relation to structural and cultural enablements and constraints. By not undertaking ambitious projects, communicative reflexives manage to avoid constraints, and they also tend to avoid enablements that might facilitate their social advancement. 'Communicative reflexives' are 'collectivist' towards the social in the sense that concerns and context are inseparable. The external effect of this stance, on a minor scale, is social reproduction in the sense of strengthening social integration, i.e., agreement among members of their network. In the context of this investigation, we regard the concern of 'interpersonal relations' as referring to collegial relationships in the school where our informant works.

Max described his relationship with colleagues as amicable and fruitful and the work environment as good. He mentions talking about the PISA test with his colleagues when the results were published and announced and presented at his school.

Reference 2

E-I43: Is there anything more you want to say about the environment before moving on to the next topic, which could be of relevance?

E-R43: No, no other than that we have an environment that tries to follow what we agree on. It can at times be a quite fruitful and useful collaborative climate. There are few dominant figures that knock through their needs. It is a pretty good and practical environment within the collegium.

Max stated that all school leaders (principals) that he knew of presented test-results to the staff - even PISA results. The school leaders present the results by PowerPoint or projector, followed by discussions in smaller groups.

Reference 1

E-I31: Where is this presented, how is this happening at your school, who is taking the initiative to present these results, and how is it being presented?

E-R31: The principal presents them in plenary for the entire collegium. Happily, with a PowerPoint, or a projector. And then you are asked to discuss in groups, preferably in subject sections, and reflect on why things are as they are, why we are where we are, etc.

Reference 6

PI-I47: How do you perceive your colleagues' competence to possibly interpret PISA results?

PI-R47: My leader is good, I think, and probably competent, and the same with the mathematics teachers. After all, there will always be someone who doesn't care about those things [PISA].

This section of the interview indicates that communicative reflexivity was activated by Max, in relation to his colleagues, who seemed to function as a collective that was able to reach agreements among themselves on school matters. They discussed PISA and other test results among themselves when these were presented to the staff by the principal. Max also considered colleagues having competencies for interpreting PISA results, but not all staff members care about PISA.

Discussion

The reflexivity profile resulting from ICONI indicated that generally Max was an *autonomous reflexive*, someone who prioritized work before leisure activities and interpersonal relations, and whose ultimate concern was performative achievement in school, adopting a strategic stance to his environment. However, answering the open question included in ICONI, Max did not mention 'work' as one of the three most important areas of his life, but he mentioned 'genuine interests' in addition to 'family' and 'friends'. 'Genuine interests' could have referred to his ambition to be a good teacher, thus, to his work. This interpretation of the reply 'genuine

interests' as Max's interest in his work is underscored by the qualitative interview. For example, Max maintaining that student motivation in mathematics was a reason for him creating 'simulated test situations'. This concern seemed to be one of his top priorities in school due to his focus on goal achievements, performing well, and maintaining motivated students. The interview data concerning his internal conversations on PISA and other tests confirmed his engagement in autonomous reflexivity. Thus he adapted to the situation and made the best of it, exemplified by him introducing 'simulated exams and tests' for his students, and his promotion of this practice in the whole school. He wanted to enable his students to achieve optimally on tests, which both satisfied his own concern to perform as a teacher at a high level in accordance with external standards, and enabled goal attainment for his school by raising students' achievements. Max mentioned that he used items from National tests and exams in his 'simulated test situations'. Norwegian lower secondary school teachers are measured on their achievements by the results on National Tests in 8th and 9th grade, results which are accessible for the public, which might explain why Max is focused on national tests items in his test-simulations. The reason Max is focused on exams items in his test-simulations might be their relevance for his students' final exam(s) in 10th grade.

In a theory of professions perspective Max might be seen to represent 'new professionalism', as his desire to perform well as a teacher coincided with the interest of his school, perhaps at the cost of spending more time on promoting knowledge among his students. The latter would have indicated a stronger leaning towards 'Old Professionalism' and the norms of the professional community to which he belonged. The depth that Archer's concept of autonomous reflexivity brings to this analysis is that Max's 'project' and actions were derived from his concern about performing well as a teacher.

Although his ICONI scores confirmed Max as a predominantly autonomous reflexive, the interview data indicated that the meta-reflexive mode was almost equally prevalent. The major difference between his engagement in the two modes of reflexivity was that his reflection in the autonomous mode resulted in action. In his autonomous mode he completed the sequence concerns → project → practices (Archer 2007), exemplified by his introduction of trial exams and 'simulated tests'. His meta-reflexive mode, on the other hand, did not issue in projects and specific practices. Max engaged the meta-reflexive mode when he reflected on why he was so critical of PISA. He hinted that his ideals as a teacher were not compatible with using the PISA test, when he emphasized that he had been assigned to administer the test rather than chosen to apply it; when he expressed his disapproval of using the test as an instrument for rating student

performances; and when he situated the test as part of ‘the miserable state’ of Norwegian education and Norwegian educational policy. Despite being critical, Max reassured that he was able to interpret PISA results. Max adopted a ‘subversive stance’ towards the larger picture of Norwegian education, but this stance had no definite project and thereby no practices, which suggests negligible contextual effects. However, Max was perhaps a person who experienced teaching as his ‘vocation’ in the sense of investing his ‘self’ in his work as a teacher and having certain moral standards. No definite conclusion can be drawn regarding this from the present interview though it confirms that he embraced certain ideals about teaching, which were contrary to ranking students’ results.

To some extent Max also engaged in communicative reflexivity in his collaborative work with the rest of the staff, describing his work environment as amicable, characterized by agreements among colleagues and his willingness to discuss PISA-results with other colleagues. Besides discussing PISA-results with his colleagues now and then, he had no clear projects and practices himself concerning PISA that involved his colleagues. Max confirmed that his teacher colleagues and the principal were able to interpret PISA results, but not all staff members cares (deeply) about PISA. Thereby Max contributed to social integration amongst the staff. However, how his autonomous, meta-, and communicative reflexivity contributed to his ‘modus vivendi’, balancing his different concerns, is an open question that would require further investigations.

Perhaps the combination of autonomous and meta- reflexivity which combines criticism with strategic adaptation to the status quo is specific to countries that have centralized educational systems like Norway (Skinningsrud 2019; Bringeland 2022a). In centralized systems, educational structures are generally determined in central political arenas and consequently difficult to challenge at the level of the individual school and by the individual teacher (Archer [1979] 2013, 1984; Skinningsrud 2019; Bringeland 2022b). Further research on modes of reflexivity among teachers encountering structural and cultural constraints in educational systems with various degrees of centralization might explore whether teachers in countries that have undertaken various types of neoliberal educational reforms activate different modes of reflexivity in dealing with their structural and cultural settings. For example, the possible identification of a predominance of ‘autonomous reflexives’ in centralized educational systems could suggest that teachers easily accommodate to central policy and are flexible towards change as they are concerned with ‘goal achievement’ and successful performance. Thereby

they could possibly contribute to social morphogenesis in their situated context, which is in line with central policy and agendas.

Conclusion

This article outlines the various theoretical approaches that have been applied in the study of teachers' reactions to the introduction of neoliberal reforms in education. Firstly, labour process theory, which emphasises de-skilling, intensification of work, and the separation of conceptions and execution of work tasks; secondly, the post-structural approach, which applies Lyotard's conception of performativity and Foucault's notions of power, discipline, and discourse; and thirdly, conceptions of 'new professionalism'. All three theories have been criticised. Labour process theory emphasises structural change and generally lacks a conception of agency. It has also been criticized for not being sufficiently specific when describing the new educational control regimes. Post-structural theory introduces the notion of 'subjectivity' but not agency since a uniform 'neoliberal subjectivity' is seen to result from neoliberal structures. Thus, like labour process theory, post-structural theory emphasizes the determining force of external influences on subjectivity. Likewise, theories of the professions, distinguishing between old and new professionalism consider the impact of new controlling structures to result in uniform reactions, shaping discourses and identities. At the same time, they conflate structure, culture, and agency under the umbrella of 'new professionalism'. All these theories, although two of them incorporate subjectivity, lack a precise conception of agency in the sense of persons possessing causally effective personal powers that codetermine their action.

As an alternative to these approaches, we suggest Margaret Archer's theory of structure, culture, and reflexive agency with her conceptions of various modes of reflexivity. Archer founds her theory of reflexivity in a basic tenet put forward in the philosophy of critical realism, that 'the causal power of social forms is mediated through human agency' (Bhaskar [1979] 1998, 28). Exploring the process through which agency mediates structure she has empirically identified three major modes of reflexivity, the *communicative*, the *autonomous* and the *meta reflexive* mode. In addition, *fractured reflexivity* occurs when reflexivity is thwarted and disoriented due to external or internal disabling conditions.

We have presented a case study of a Norwegian secondary school teacher, who answered the questionnaire ICONI measuring his reflexivity profile, i.e., his mean score on the various modes of reflexivity in a general nonreferential manner, and, in addition, data from a qualitative semi-structured interview about his reflections on the PISA test related to his work context. The

major merits of Archer's theoretical approach are maintaining a clear ontological distinction between structure, culture, and agency, thus postulating their independent causal powers, and distinguishing between various modes of reflexivity based on agents' ultimate concerns. The various modes of reflexivity have the potential to explain why reactions to the same environments (structure and culture) differ between individual teachers, depending on their different personal concerns and what they care deeply about.

Our case study with a secondary school teacher shows how a specific combination of modes of reflexivity, the autonomous and the meta-reflexive modes, under the given circumstances, operates through a *strategic* and *critical* stance towards existing state of affairs. This might be a feasible *modus vivendi* for teachers with PISA. Interestingly, a similar prevalence of the autonomous and meta-reflexive modes regarding PISA was discovered among Norwegian School leaders (Bringeland 2022b). Based on these joint findings our hypothesis is that this specific combination of the autonomous and meta-reflexive modes might have been promoted by the type of educational system in which both school leaders and teachers are located, namely the centralized Norwegian educational system. A prominent feature of such systems is that major decisions regarding educational structures are made at the central level of political decision making. In such systems individual teachers, despite their criticism of current arrangements, are unable to directly influence structural change. A substantiation of our hypothesis would require further studies of larger samples of teachers and comparative studies between countries whose educational systems differ regarding degrees of centralization and their specific configuration of neoliberal reforms.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Adams, Tracey, and Peter H. Sawchuk. 2020. "Professional-Organizational Contradictions and Hybridization of Knowledge: Insights from the Study of Engineering and Nursing in Canada." *Vocations and Learning* 14 (1):75–93. doi: 10.1007/s12186-020-09253-1.
- Anderson, Gary, and Kathryn Herr. 2015. "New Public Management and the New Professionalism in Education: Framing the Issue." *Education Policy Analysis Archives* 23 (84):1–9. doi: 10.14507/epaa.v23.2222.
- Apple, Michael W. 2013. "Controlling the Work of Teachers." In *Knowledge, Power and Education: The Selected Works of Michael W. Apple*, edited by Michael W Apple, 126–141. New York and Abingdon, Oxon: Routledge.
- Archer, Margaret S. 1984. *The University Edition of Social Origins of Educational Systems*. London: Sage Publications.

- Archer, Margaret S. 1995. *Realist Social Theory: The Morphogenetic Approach*. Cambridge: Cambridge University Press.
- Archer, Margaret S. 2000. *Being Human: The Problem of Agency*. Cambridge: Cambridge University Press.
- Archer, Margaret S. 2003. *Structure, Agency and the Internal Conversation*. Cambridge: Cambridge University Press.
- Archer, Margaret S. 2007. *Making Our Way through the World: Human Reflexivity and Social Mobility*. Cambridge: Cambridge University Press.
- Archer, Margaret S. 2012. *The Reflexive Imperative in Late Modernity*. Cambridge: Cambridge University Press.
- Archer, Margaret S. [1979] 2013. *Social Origins of Educational Systems*. London: Sage.
- Archer, Margaret S. [1988] 1996. *Culture and Agency: The Place of Culture in Social Theory*. Cambridge: Cambridge University Press.
- Aursand, Leah Rose. 2018. "What [Some] Students Know and Can Do: A Case Study of Norway, PISA, and Exclusion." Master's Thesis, Department of Education, University of Oslo, Oslo, Norway.
- Aursand, Leah Rose, and David Rutkowski. 2021. "Exemption or Exclusion? A study of Student Exclusion in PISA in Norway." *Nordic Journal of Studies in Educational Policy* 7 (1):16–29. doi: 10.1080/20020317.2020.1856314.
- Ball, Stephen J. 2003. "The Teacher's Soul and the Terrors of Performativity." *Journal of Education Policy* 18 (2):215–228. doi: 10.1080/0268093022000043065.
- Ball, Stephen J. 2012. *Global Education Inc : New Policy Networks and the Neo-Liberal Imaginary*. London: Routledge.
- Ball, Stephen J. 2016. "Subjectivity as a Site of Struggle: Refusing Neoliberalism?" *British Journal of Sociology of Education* 37 (8):1129–1146. doi: 10.1080/01425692.2015.1044072.
- Ball, Stephen J., Meg Maguire, and Annette Braun. 2012. *How Schools Do Policy: Policy Enactments in Secondary Schools*. Abingdon, Oxon: Routledge.
- Ball, Stephen J., and Antonio Olmedo. 2013. "Care of the Self, Resistance and Subjectivity Under Neoliberal Governmentalities." *Critical Studies in Education* 54 (1):85-96. doi: 10.1080/17508487.2013.740678.
- Bergmark, Ulrika, and Kristina Hansson. 2021. "How Teachers and Principals Enact the Policy of Building Education in Sweden on a Scientific Foundation and Proven Experience: Challenges and Opportunities." *Scandinavian Journal of Educational Research* 65 (3):448-467. doi: 10.1080/00313831.2020.1713883.
- Bhaskar, Roy. 2016. *Enlightened Common Sense: The Philosophy of Critical Realism*. London: Routledge.
- Bhaskar, Roy. [1979] 1998. *The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences*. London: Routledge.
- Braverman, Harry. 1974. *Labor and Monopoly Capital : The Degradation of Work in the Twentieth Century*. New York: Monthly Review Press.
- Bringeland, Terje André. 2022a. *The Impact of PISA on Education in Norway: A Morphogenetic Perspective on Structural Elaboration in an Education System*. Edited by Margaret S. Archer, Unn-Doris K. Bæck and Tone Skinningsrud, *The Morphogenesis of the Norwegian Educational System: Emergence and Development from a Critical Realist Perspective*. Milton: Taylor & Francis Group.
- Bringeland, Terje André. 2022b. "School Leaders' Reflexive Mode in their Internal Conversations on PISA." In *The Morphogenesis of the Norwegian Educational System: Emergence and Development from a Critical Realist Perspective*, edited by Margaret S.

- Archer, Unn-Doris K. Bæck and Tone Skinningsrud, 181–210. Milton: Taylor & Francis Group.
- Carlgren, Ingrid, and Kirsti Klette. 2008. "Reconstructions of Nordic Teachers: Reform Policies and Teachers' Work during the 1990s." *Scandinavian Journal of Educational Research* 52 (2):117-133. doi: 10.1080/00313830801915754.
- Connell, Raewyn. 2009. "Good Teachers on Dangerous Ground: Towards a New View of Teacher Quality and Professionalism." *Critical Studies in Education* 50 (3):213-229. doi: 10.1080/17508480902998421.
- Etzioni, Amitai. 1969. *Semi-Professions and Their Organization; Teachers, Nurses, Social Workers*. Edited by Amitai Etzioni. New York: The Free Press.
- Evetts, Julia. 2011. "A New Professionalism? Challenges and Opportunities." *Current Sociology* 59 (4):406–422. doi: 10.1177/0011392111402585.
- Fuller, Kay. 2019. ""That Would Be My Red Line": An Analysis of Headteachers' Resistance of Neoliberal Education Reforms." *Educational Review (Birmingham)* 71 (1):31-50. doi: 10.1080/00131911.2019.1522042.
- Gerstle, Gary. 2018. "The Rise and Fall (?) of Americas's Neoliberal Order." *Transactions of the Royal Historical Society* 28:241-264. doi: 10.1017/S0080440118000129.
- Gunter, Helen M, Emiliano Grimaldi, David Hall, and Roberto Serpieri. 2016. *New Public Management and the Reform of Education: European Lessons for Policy and Practice*. Edited by Helen M Gunter, Emiliano Grimaldi, David Hall and Roberto Serpieri. Abingdon, Oxon and New York: Routledge.
- Hansson, Kristina, and Per-Olof Erixon. 2020. "Academisation and Teachers' Dilemmas." *European Educational Research Journal* 19 (4):289-309. doi: 10.1177/1474904119872935.
- Helgøy, Ingrid, and Anne Homme. 2007. "Towards a New Professionalism in School? A Comparative Study of Teacher Autonomy in Norway and Sweden." *European Educational Research Journal* 6 (3):232-249. doi: 10.2304/eej.2007.6.3.232.
- Hodge, Steven, John Holford, Marcella Milana, Richard Waller, and Sue Webb. 2018. "Economic Theory, Neoliberalism and the Interests of Educators." *International Journal of Lifelong Education* 37 (3):279-282. doi: 10.1080/02601370.2018.1484009.
- Holloway, Jessica, and Jory Brass. 2018. "Making Accountable Teachers: The Terrors and Pleasures of Performativity." *Journal of Education Policy* 33 (3):361-382. doi: 10.1080/02680939.2017.1372636.
- Lewis, Steven, and Ian Hardy. 2014. "Funding, Reputation and Targets: The Discursive Logics of High-Stakes Testing." *Cambridge Journal of Education* 45 (2):245-264. doi: 10.1080/0305764X.2014.936826.
- Lyotard, Jean-François. 1984. *The Postmodern Condition : A Report on Knowledge*. Vol. 10, *La Condition Postmoderne*. Minneapolis: University of Minnesota Press.
- Møller, Jorunn, and Guri Skedsmo. 2013. "Modernising Education: New Public Management Reform in the Norwegian Education system." *Journal of Educational Administration and History* 45 (4):336–353.
- OECD. 2005. *Teachers Matter : Attracting, Developing and Retaining Effective Teachers, Education and training policy*. Paris: OECD.
- Ozga, Jenny, and Martin Lawn. 1988. "Schoolwork: Interpreting the Labour Process of Teaching." *British Journal of Sociology of Education* 9 (3):323–336. doi: 10.1080/0142569880090305.
- Reid, Alan. 2003. "Understanding Teachers' Work: Is There Still A Place for Labour Process Theory?" *British Journal of Sociology of Education* 24 (5):559–573. doi: 10.1080/0142569032000127134.

- Robertson, Susan L. 2012. "Placing Teachers in Global Governance Agendas." *Comparative Education Review* 56 (4):584-607. doi: 10.1086/667414.
- Skedsmo, Guri, and Sølvi Mausestaden. 2017. "Nye styringsformer i utdanningssektoren - spenninger mellom resultatstyring og faglig-profesjonelt ansvar [New Forms of Management in the Education Sector—Tensions between Performance Management and Professional Responsibility]." *Norsk pedagogisk tidsskrift* 101 (2):169-179.
- Skinningsrud, Tone. 2019. "Vindicating Archer's Concepts of Educational Systems – Centralized and Decentralized – As Exemplars of Critical Realist Theorizing." *Journal of Critical Realism* 18 (4):453–473. doi: 10.1080/14767430.2019.1656924.
- Svensson, Lennart G. 2006. "New Professionalism, Trust and Competence: Some Conceptual Remarks and Empirical Data." *Current Sociology* 54 (4):579–593. doi: 10.1177/0011392106065089.
- Troman, Geoff. 1996. "The Rise of the New Professionals? The Restructuring of Primary Teachers' Work and Professionalism." *British Journal of Sociology of Education* 17 (4):473–487. doi: 10.1080/0142569960170404.

Article III: School Leaders' Reflexive Mode in their Internal Conversations on PISA

Bringeland, Terje André. 2022. "School Leaders' Reflexive Mode in their Internal Conversations on PISA." In *The Morphogenesis of the Norwegian Educational System: Emergence and Development from a Critical Realist Perspective*, edited by Margaret S. Archer, Unn-Doris K. Bæck and Tone Skinningsrud, 181–210. Milton: Taylor & Francis Group.

