Measuring social impact –
The conceptual and empirical advancement of an emergent concept

Thesis submitted for the award of a D.Phil.

Written by

Gorgi Krlev

University of Oxford

Department for Continuing Education (OUDCE)

Kellogg College

October 2017, Trinity Term

Subject area: Business and Management

Supervisors: Professor Jonathan Michie, Professor Alex Nicholls

Examiners: Professor Martin Ruhs (internal), Professor Patrick Kenis (external)

Measuring social impact –
The conceptual and empirical advancement of an emergent concept

Abstract

This thesis develops a theoretical foundation for measuring social impact. It focuses on 'social purpose organisations', for example non-profit organisations. Its original contribution lies in the conceptual proposition of measuring social impact through assessing how an organisation enhances social productivity by creating capitals, for instance social, cultural or political capital. It proposes measuring capitals as an alternative to currently prevalent metrics surrounding: (1) costs and benefits; (2) life-satisfaction; (3) capabilities.

The research tests the 'capital based approach' in three empirical cases located in Germany. The studies consider multiple capitals, but focus on the main one for each intervention: (1) social capital in community oriented housing for elderly people; (2) cultural capital in a school-based intervention on pro-sociality; (3) political capital related to an online platform for citizens and politicians. The conceptual proposition combined with this 'multi-site multi capitals' setup generates knowledge on social impact in four different ways.

First, by tapping into established capitals measures across the social sciences the research mitigates the measurement problem—the challenge of translating social impact into measurable components. Second, by using capitals as a set of interconnected measures, which are furthermore at close proximity to organisational activities, it helps address the attribution problem—the difficulty of establishing causal links between an intervention and its effects. Third, by turning to capitals it highlights intermediate outcomes and processes in the creation of social impact and thereby deals with the aggregation problem—a loss of information when focussing only on final social outcomes. Fourth, by seizing on the position of capitals which are located in between individuals and the community, it responds to the foundation problem—the potential discrepancy between group level welfare and individual utilities or capacities.

The research tackles the most pressing challenges in social impact measurement and thus significantly advances its theory and practice.

Keywords: social impact; measurement; capitals.

Thesis submitted for the award of a D.Phil. (approx. 93,600 words, excl. bibliography)

Written by Gorgi Krlev

University of Oxford

Department for Continuing Education (OUDCE)

Kellogg College

October 2017, Trinity Term

Subject area: Business and Management

Supervisors: Professor Jonathan Michie, Professor Alex Nicholls

Acknowledgements

Any work done is always and steadily influenced by the people and circumstances surrounding the one performing it—sometimes to the worse, mostly to the better. The last couple of years have been a personal journey as well as an intellectual one. I want to thank those who have influenced me throughout this time, and in their own way have contributed to the crafting of this piece of work.

If one thing could be learnt from my dissertation, it is that it would have taken a much more sophisticated investigation than the time I have invested into putting this together to account for the exact impact each individual had on what readers now see before them, and on me personally. I will instead try to highlight how the persons I mention made the biggest impression on me. To keep it short, yet personal and comprehensive, I decided to trade the usual prose for a list and also to deviate in part from the traditional order of acknowledgements.

Thus, my heartfelt thanks go to:

Sabrina, Malin and Emmi for being just the way they are, their love, and for enriching my life in a myriad of ways. My mum and my grandmother for their unshakable and ceaseless support of whatever I do.

Jonathan Michie and Alex Nicholls for their guidance, their trust in me and for encouraging me to go forward with what many others (would) have warned me against. Jonathan also for his great hospitality. Martin Ruhs and Marc Ventresca for their critical, yet extremely appreciative and productive comments in the Transfer and Confirmation assessments, which have pushed my ways of thinking. Martin additionally and Patrick Kenis for a great viva voce experience that generated further improvements to my work. Gayle Avery, Robert Isaak and Helmut Anheier for their recommendation letters which certainly have done their part in getting my accepted for the DPhil.

All supporters and participants of my extensive empirical research. Foremost: Berthold Broll and Markus Nachbaur at Stiftung Liebenau and Alexander Künzel at Bremer Heimstiftung for agreeing to fund my study on community oriented housing. Ulrich Kuhn and Elke Munderloh for being my prime and highly engaged contacts for the operational realisation of the project, along with Andreas Schlegel at Samariterstiftung. Heather Cameron, Daniel Meehan, Julia Gedeon, Melanie Marahrens and Jana Gottschalk at CamP Group for enabling access to the participating schools. Several others, but in particular Roman Ebener at abgeordnetenwatch.de for his great communication efforts in promoting participation in my survey and his inexhaustible patience in waiting for the results. The inhabitants, the teachers, the pupils, the citizens and the politicians for participating in my research.

All my (former) colleagues at Heidelberg University's Centre for Social Investment (CSI) for being my intellectual basecamp and a group of friends at the same time. In particular: Georg Mildenberger for his understanding and advice on all matters of life, and for granting me great liberties in my doing. Volker Then for doing just the same, for backing me in my endeavours and for openly sharing personal insights. Martin Hölz, Thomas Scheuerle and Björn Schmitz for the many discussions, the trips, and the fun. Ute Bongertz, Dominic Egger, Katharina Enz, Myriam Goebel, Julia Hildmann, Stephanie Hörnig, Hannah Jonitz, Alina Langenhorst, Josefa Loebell, Romy Naumann, Tabea Pottiez, Peter Schubert and

Jakob Schultz for doing a tremendous job in interviewing the older people for one of my empirical studies. Ursula Fischer for managing the funds of that project. In addition to several of the above: Lea Heyer, Gudrun Schimpf and Eva Kesternich for their quick and careful commenting on my chapters. Max Vetter for having a psychologist's eye on my scales. Stephan Stahlschmidt for his most competent statistics advice and for always being available for a long chat on short notice. Also my other colleagues at Heidelberg and abroad, in particular Nico Seifert, Jan Eckhard and Markus Lang as well as René Bekkers and Georgios Melios for debating the quantitative part of my work with me, and Anker Brink Lund as well as Annette Bauer for inspiration in thinking about social impact.

All members of Oxford University's Kellogg College and its Department for Continuing Education for having become my intellectual family abroad, with many enjoyable and always inspiring encounters. Especially: Emma Potts for the numerous lunch breaks, for her open ear and for being the best College Advisor I could have imagined. Adrian Stokes for making everything possible, from methods training to ethics reviews, and in such a swift and pleasant manner. Claire O'Mahoney for her poetry, her sympathy and the tea. Sarah O'Brien for enduring my incessant inquiries and providing all administrative and personal support conceivable, reaching from College grants as far as to recommending the best possible childminder. All evidence-based healthcare researchers that have provided such effective and useful methods training. Specifically: Clare Bankhead, Anne-Marie Boylan, Jenny Hislop, Emily McFadden, Sharon Mickan, Jason Oke, and Braden O'Neill. Anne-Marie and Clare additionally for their extra time in commenting on my projects, and Clare in particular for her warm words. Liz Sanders and Órlaith Burke for organising a memorable Rewley House Research Seminar and Kellogg Colloquium on Impact, respectively.

Paul du Gay for nice and stimulating Oxford meetings. Helena Fox for being a gorgeous host during my visits to Oxford and intellectual companion along the journey, if on separate roads.

All those listed for being wonderful people. Getting to know you has not only helped advance my DPhil work, but been personally valuable and become dear to me.

Gorgi Krlev
Mannheim/Heidelberg
March 2018

Content

Content **7**

List of figures **11**

List of tables **12**

1 Introduction **15**

1.1 The evolution of interest in social impact 15
1.2 The meaning of social impact 21
1.3 Why ‘measure’ social impact, its functions and constituents 22
1.4 Current tools and trends 26
1.5 Research questions 30
1.6 The theoretical argument and empirical design of the thesis 32

2 Survey of the literature **35**

2.1 Existing conceptual approaches to social impact 35
2.1.1 Cost-Benefit Analysis 36
2.1.2 Life satisfaction Approaches 40
2.1.3 The Capability Approach 43
2.2 Summary – challenges in existing approaches 45

3 Conceptual advancement **49**

3.1 Measuring social impact – a new approach 49
3.2 Social impact as a contribution to social productivity 50
3.3 Challenges in assessing social impact within the system of social productivity 56
3.4 Properties of the capital based approach to social impact 58
3.5 The different forms of capital 69
3.5.1 Social capital 73
3.5.2 Cultural capital 75
3.5.3 Political capital 78
3.5.4 Relating the capitals 80
3.6 Advantages of the ‘Capital Based Approach’ 82

4 Methodological framework **85**

4.1 Measuring capitals – breadth and focus 85
4.2 Research design versus the conceptual approach 87
4.3 The studied interventions and points of reference for capital measures 90
4.4 Analytical procedure 95
4.4.1 Accounting for complexity and research interest 96
4.4.2 Defining scope and measures 99
4.4.3 Choosing the appropriate research design 101
4.4.4 Composing instruments 104
4.4.5 Refining measures 106

4.4.6 Pre-testing and executing the study ... 108
4.4.7 Analysing data and presenting results ... 109
4.5 Properties of the methodological framework ... 112

5 Social impact of community oriented housing ... 115

5.1 Background ... 116
5.2 Specifics of community oriented housing ... 118
5.3 Methods ... 120
5.3.1 Setting ... 120
5.3.2 Participants ... 121
5.3.3 Analytic strategy ... 122
5.3.4 Accounting for group characteristics: matching and regression ... 123
5.3.4.1 Initial differences between groups ... 123
5.3.4.2 Matching procedure ... 126
5.3.4.3 Group characteristics after weighting ... 127
5.3.4.4 Regression analysis ... 130
5.3.5 Effect measures ... 132
5.4 Results ... 137
5.4.1 Cultural capital ... 142
5.4.2 Political capital ... 143
5.4.3 Social capital ... 144
5.4.4 Secondary variables ... 149
5.5 Discussion ... 150
5.5.1 Main effects produced ... 150
5.5.2 Complementary and follow-up issues ... 153
5.5.3 Un-pursued perspectives ... 155
5.5.4 Limitations ... 156
5.6 Conclusion ... 157

6 Social impact of a school-based intervention ... 159

6.1 Background ... 160
6.2 Specifics of the programme ... 163
6.3 Methods ... 165
6.3.1 Setting and participants ... 165
6.3.2 Analytic strategy ... 167
6.3.3 Effect measures ... 169
6.4 Results ... 172
6.4.1 Quantitative scoping ... 172
6.4.2 Qualitative analysis ... 175
6.4.2.1 Pupils' own accounts ... 175
6.4.2.1.1 General assessment of the programme ... 175
6.4.2.1.2 Cultural capital ... 176
6.4.2.1.3 Social capital ... 179
6.4.2.1.4 Political capital ... 180

6.4.2.2 Teachers' assessments 180
6.4.3 Mechanisms and moderating context factors 181
6.4.3.1 Confounding factors 182
6.4.3.2 Effect mechanisms of the programme 183
6.4.3.2.1 Boxing 183
6.4.3.2.2 Video work 185
6.4.3.2.3 Trust exercises and discussions 185
6.5 Discussion 186
6.5.1 Main effects produced and missed 187
6.5.2 Limitations 191
6.6 Conclusion 192

7 Social impact of a political web-platform 193

7.1 Background 194
7.2 Specifics of the web-platform 196
7.3 Methods 197
7.3.1 Participants 197
7.3.2 Analytic strategy 197
7.3.3 Socio-demographics of the groups of respondents 198
7.3.4 Further characteristics: capitals baselines and satisfaction 202
7.3.5 Effect measures 209
7.4 Results 213
7.4.1 Immediate effects on capitals 214
7.4.2 Mediated effects on capitals 217
7.4.2.1 Comparison of capital levels 218
7.4.2.2 Citizens' and politicians' assessment of mediated effects 222
7.4.2.3 Probing active engagement 224
7.4.3 Effects of organisational activities on political capital 226
7.5 Discussion 229
7.5.1 Main effects produced from citizens' perspective 229
7.5.1.1 Immediate effects 229
7.5.1.2 Mediated effects 231
7.5.1.3 Effects of organisational activities 232
7.5.2 Main effects produced from politicians' perspective 233
7.5.3 Sensitivity of results 235
7.5.4 Limitations 237
7.6 Conclusion 238

8 Discussion 241

8.1 The 'multi-sites multi-capitals' setup and its uses 241
8.1.1 Social capital in the community context 244
8.1.2 Pro-social behaviour as cultural capital 247
8.1.3 Political capital and democratic functioning 250
8.1.4 Capitals combined 251

8.2 The ‘Capital Based Approach’ and research designs 256
8.3 The ‘Capital Based Approach’ and organisational learning 262
8.4 Shortcomings, challenges and open questions 267

9 Conclusions 275

9.1 Contributions 275
9.2 Relations between the different conceptual approaches 280
9.3 Lessons learnt for measuring social impact 285
9.4 Limitations 288
9.5 Implications for management theory, policy and practice 290
9.6 Prompts for future research 292
9.7 Outlook 295

10 Bibliography 297

11 Appendix 325

Appendix 1 Scales community oriented housing 325
Appendix 2 Regression results community oriented housing 327
Appendix 3 Scales and interview guide school-based intervention 331
Appendix 4 Additional analysis political web-platform 334
Appendix 5 Scales political web-platform 342

List of figures

Figure 1 From industrial production to social productivity 52
Figure 2 Radar chart: exemplary display of capitals measured 111
Figure 3 Size of capital formation CO 141
Figure 4 Network strength to neighbours 145
Figure 5 Capital formation by RA 189
Figure 6 Watching news and political information on TV 206
Figure 7 Reduction of income inequalities 207
Figure 8 Trust in people 207
Figure 9 Satisfaction with how democracy works 208
Figure 10 Immediate effects: differences between citizens and politicians 216
Figure 11 Mediated effects: differences CS and ESS 221
Figure 12 Influence of AW on capital formation: differences CS and PS 222
Figure 13 Parties' accessibility AW (frequent responses) 341
Figure 14 Parties' accessibility AW (rarer responses) 341

Abbreviations in lists of figures and tables:

CO (community oriented housing); RA (RespAct); AW (abgeordnetenwatch); CS (citizen survey); PS (politician survey); ESS (European Social Survey); SROI (Social Return on Investment).

List of tables

Table 1 Evolvement of social impact in economic thinking 18
Table 2 Challenges in existing conceptual approaches 46
Table 3 Indicators used in SROI studies 71
Table 4 Differences in the complexity of interventions 97
Table 5 Matching variables and differences 124
Table 6 Control variables after weighting (part 1) 128
Table 7 Control variables after weighting (part 2) 129
Table 8 Capital forms and secondary variables 133
Table 9 Overview of effects CO 139
Table 10 Socio-demographics of participating pupils 167
Table 11 Applied scales 170
Table 12 Quantitative effects 173
Table 13 Socio-demographics CS, PS and ESS 200
Table 14 Federal states CS, PS and ESS 201
Table 15 Voting and political orientation CS, PS and ESS 203
Table 16 Political capital, cultural capital and satisfaction CS, PS and ESS 204
Table 17 Effect measures 210
Table 18 Capital formation: immediate effects 214
Table 19 Mediated effects: levels of capitals CS and ESS 219
Table 20 Influence of AW on capital formation: levels CS and PS 223
Table 21 Active political engagement: differences CS and ESS 224
Table 22 Knowledge of AW's activities CS and PS 227
Table 23 Effects of AW's activities 228
Table 24 Multi-capitals framework 252
Table 25 Secondary effect variables and control variables 255
Table 26 Specifics of the empirical research designs 256
Table 27 Scales effects CO (part 1) 325
Table 28 Scales effects CO (part 2) 326
Table 29 Scales matching variables CO-AL 326
Table 30 Regression results CO (part 1) 328
Table 31 Regression results CO (part 2) 329
Table 32 Regression results CO (part 3) 330
Table 33 Scales quantitative scoping RA (part 1) 331
Table 34 Scales quantitative scoping RA (part 2) 332

Table 35 Semi-structured interview guide for pupils RA 333
Table 36 Frequency of use AW 335
Table 37 Intensity of use AW 335
Table 38 Main purpose of use AW 336
Table 39 Alternative modes of communication AW 336
Table 40 Main policy levels AW 339
Table 41 Contact with parties AW 340
Table 42 ESS questions AW 342

1 Introduction

Be not simply good; be good for something.
Henry David Thoreau
(Familiar Letters, 1895)

"What are organisations' actions good for in society, and how can we assess these actions?" This is what measuring social impact is about. The introductory chapter sheds light on five questions related to social impact: (1) Where does the interest in it come from?; (2) What is it exactly that people want to measure?; (3) What and whom does the measurement serve?; (4) How is it currently done?; and (5) Why is this thesis needed? It then provides a short outline of the thesis' theoretical argument and empirical work, which seek to advance our understanding of social impact measurement.

1.1 The evolution of interest in social impact

For a long time, scholars have performed research in economics and management in a largely isolated fashion, focussing on the technical aspects of supply and demand for market goods without giving much attention to the context in which economic activity takes place. Coase (1992, p. 73) refers to this circumstance in an ironic remark quoting poet Roy Campbell: "I see the bridle and the bit all right. But where's the bloody horse?" Coase uses another, similar figure of speech when he laments that economists used to study the circulation of the blood without a body. It is claimed that this started to change through the rise of 'new institutional economics' (a term coined by Williamson (1975)), which has embedded economic activity into the wider social context. This context is embodied by informal and formal institutions, ranging from culture to the polity as the manifested structure of political governance (Williamson, 2000, p. 597).

Two of the milestones in the scholarly tradition of taking context and social effects seriously were 'The Nature of the Firm' (Coase, 1937) and 'The Problem of Social Cost' (Coase, 1960). Both deal with costs that arise through economic transactions. Many of these costs do not affect the firm however, but its surrounding social and natural environment. At around this point, scholars began to think systematically about the negative spill-overs and unintended effects economic activity can have, such as environmental pollution. Instead of maximizing firm profit ceteris paribus, the damage caused by economic activity had to be taken into account and weighed against the benefits produced by it. This is not a purely rationalist process though and always contains a normative dimension, more specifically a judgement of what is dear and valuable. Coase (1960, p. 21) hints at this by saying that decisions on economic activity "[…] must ultimately dissolve into a study of aesthetics and morals." In the wake of this development, attention in economics and management shifted slowly but steadily from a focus on profits to the creation or preservation of other kinds of value. This urge persisted despite some opposition, such as the one found in Friedman's (1970) notorious article in which he argued that the sole responsibility of business was to increase its profits.

It took years though before variables other than gross national income started penetrating the systems of national accounts. Some scholars nowadays promote happiness, life satisfaction or 'well-being' as increasingly important factors in grasping the value created in societies (Frey, 2010; Frey & Stutzer, 2002b). Others go even further and suggest that traditional economic categories such as unemployment, income or inflation should be replaced by happiness and related measures (Fioramonti, 2013). This proposition is embedded in the wider discussion of 'new wealth indicators' (Diefenbacher & Zieschank, 2010), which aim at integrating aspects such as the value of civic engagement or the prevalence of crime rates into national accounts, in order to reflect the wealth of a nation, region or city

more accurately than is possible by economic metrics alone. The status of health and health care provision, education, "political voice and governance" or "social connections and relationships" form part of a broader set of indicators that are gaining more attention (Stiglitz, Sen, & Fitoussi, 2009, p. 14).

The increasing appreciation of social and environmental dimensions at the national level has had its counterpart on the meso level of markets and the micro level of organisations. The 'triple-bottom-line' is one of several new principles guiding organisational activity. Elkington (1998), who coined the term, stresses the ecological and social dimension in business activity and thereby strengthens the position of the latter two aspects against the goal of profit maximization. In addition to fostering cooperation between firms and non-governmental organisations, the concept introduces the assessment of environmental impacts of projects or products as a critical variable in strategic decision making. Elkington argues that the concept is a response to growing stakeholder influence, which brings us back to the very foundation of new institutional economics.

The rationale of taking environmental and social impacts into account has been pushed further by the discussion of 'shared value', "[…] which involves creating economic value in a way that *also* creates value for society by addressing its needs and challenges. […] It is not on the margin of what companies do but at the center." (Porter & Kramer, 2011, XX, italics in original). While we cannot expect the striving for social value creation to trespass the one for financial profits in business corporations, the social element has made a race up to the priority list of economic activity. The dissolving separation between social and financial "realities" has not only driven organisations to act on 'blended value' motives but also led to the evolvement of an "integrated capital market" interested in both, financial and social returns (Emerson, 2003).

Table 1 summarises the shifts in economic and management thinking that have occurred, from treating social and environmental issues as a non-causa (Stage 1), to acknowledging potential trade-offs between economic activity and the latter (Stage 2), to embracing social and environmental issues as central to business activity and thus social impact as a relevant measure of performance (Stage 3). Along the three stages social aspects have moved up in their significance, from mattering only in direct relation to the production process in stage 1, to being outcomes of a process of production which are of value equal to that of economic contributions in Stage 3. The dominant principle in economic and business action can be categorised as: (1) 'industrial organisation'; (2) 'externalities'; and (3) 'blended value' or 'multiple bottom lines'.

Table 1 Evolvement of social impact in economic thinking

	Stage 1	**Stage 2**	**Stage 3**
Regard for social effects	• Negligence of the social effects of economic activity	• Trade-off between economic activity and its social effects	• Economic activity and the creation of social impact as distinct priorities
Hierarchical positioning of social aspects	• Social aspects only factors in economic production process	• Social aspects are granted an inherent value, but one that is subordinate to economic profits	• Social and economic 'production' as equally valued contributions
Dominant principle	*Industrial organisation – profit maximization*	*Externalities*	*Blended value, multiple bottom-lines*

There are several established strands of discussion that intertwine business activity and social mission and thereby bridge the divide between the two that once was. The 'moralization of markets' (Stehr, Henning, & Weiler, 2010) for instance is concerned with the increasing influence of 'political consumers' on industries or that of the virtues of fairness, solidarity and sustainability on business ethics. Descriptive terms such as 'Caring Capitalism' (Barman, 2016) comprise 'corporate social responsibility' (CSR) initiatives (Garriga & Melé, 2004), 'responsible investment' (Renneboog, Ter Horst, & Zhang, 2008), or

'base-of-the-pyramid' business (Prahalad & Hart, 2002)—all of which embrace social mission as central or at least highly complementary to business activity.

These developments in the for-profit sector have been met by three concurrent developments in the public and non-profit sectors. First, there has been a move towards 'evidence-based practice' in public service provision. It was established as a defined set of action principles in medicine and health care in the early 1990s (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996) and is now also promoted in many other fields, including development assistance (Savedoff, Levine, & Birdsall, 2006), public policy (Pawson, 2002), or social reform (Schorr, 2016). Precedents of evidence-based thinking go back a long way. They can be found in the context of more effective decision-making in government, demanded in view of social problems and the availability of alternative interventions (Rivlin, 1971). Behind all such approaches is the desire to find out what works best in serving target groups, be they patients or citizens.

Second, social mission and economic action have been merged with the hope of a mutually reinforcing relationship that would lever the magnitude of contributions to the 'common good'. The desire to combine both worlds is embodied in the very terminology used: 'Venture philanthropy' (Grossman, Appleby, & Reimers, 2013) is only one example. It builds on the assumption that increasing strategic action and treating donations as 'investments' into good causes would increase the effectiveness of such actions, for the benefit of all involved (Mair & Hehenberger, 2014). These rationales are also underlying the emergence of a 'social or impact investment market' (Nicholls, 2010b; Social Impact Investment Taskforce, 2014) that in many ways works like a traditional investment market, with the exception that social returns are given priority. 'Social entrepreneurship' (Dees, 2001) in

turn marks the tapping of enterprising behaviour and innovative action for solving pressing social problems.

Third, public and non-profit organisations have generally become more professionalized (Hwang & Powell, 2009), more "business-like" some say (Dart, 2004). In terms of accountability this has led to a shift in focus from 'responsibility' (following the rules) and 'controllability' (fulfilling a principal's expectations) to increased 'transparency' through revealing the facts of organisational performance and 'liability' for this performance (Koppell, 2005). Non-profit performance metrics are no longer only concerned with basic financial information or how resources have been deployed, but relate for instance to the number of services performed and their quality (Epstein & Yuthas, 2014; Poister, Hall, & Aristigueta, 2015). Advocacy, participation and innovation have come to complement efficiency, effectiveness and equity in the assessment of how well voluntary organisations do (Kendall & Knapp, 2000). Management tools originally applied in firms have been adapted for use in non-profits (Kaplan, 2001; Liket, Rey-Garcia, & Maas, 2014).

There is a debate about whether the above developments embody a neo-liberal agenda that is promoting the hollowing out of the welfare state and an economization of pressing social issues (see for instance Nicholls & Teasdale, 2017 on social enterprise). Independent of one's judgement of whether these developments are a solution or a new problem in disguise, most would agree that it was exactly this simultaneous move from all sides that has promoted interest in this question: What has been achieved for society in environmental and social terms? In other words: What 'social impact' has been produced?

1.2 The meaning of social impact

Social impact typically refers to the positive (and potentially negative) consequences of a social intervention on its target group and other key stakeholders. Common definitions of social impact mostly relate to Carol Weiss' (1972; 1998) logic model of a causal line of effects spanning from inputs to impacts. This is often also referred to as the 'theory of change' of an intervention (Social Impact Investment Taskforce, 2014). The logic model, a revitalised classic from the evaluation literature, describes how organisations move from their organisational mission, to performed activities to outputs and then to outcomes. Despite the amounts of criticism of the logic model for being too simplistic and deterministic (Wholey, Hatry, & Newcomer, 2010, p. 26), it is useful for defining what we actually mean by social impact.

There are two different versions of how we move on from outcomes to impact. There are some who argue that impact refers to "changes, or effects, on society or the environment that follow from outcomes that have been achieved" (Social Impact Investment Taskforce, 2014, p. 6)—an example for outcomes being the reduction of recidivism in ex-offenders, while the reduction of crime at a broader scale would be regarded as impact. Others have drafted impact as "outcomes minus what would have happened anyway" (Clark, Rosenzweig, Long, & Olsen, 2004, p. 7) to discern the relative contribution of actors to solving a problem or improving the situation for those affected by an intervention.

I follow the latter though older conception for two reasons: Firstly, in the first concept impact is merely delineated from outcome by the element of scope—individual level outcomes lead to societal impacts. This only adds an additional stage to the causal chain without providing any implication of how to deal with the confounding invited by this move. As will be seen in the research a focus on final outcomes and aggregated measures, which

is spurred by interpreting social impact as changes within the wider system of society, contributes to problems in assessing social impact. Secondly, the use of terminology of the second concept fits more neatly the various conceptual approaches to social impact measurement discussed in my research. They commonly refer to outcomes rather than impacts. For these reasons I use 'outcomes' and 'impacts' interchangeably and essentially refer to 'attributable outcomes' or 'net outcomes' when I apply the term outcomes.

1.3 Why 'measure' social impact, its functions and constituents

My work explicitly focuses on 'measurement' rather than alternative modes of assessment for mainly two reasons, one of which is situated in social science theory and its empirical application, the other in organisational practice:

First, social impact as described in relation to the 'theory of change' comes with the idea of causality and traceability. It therefore requires a structured and systematic assessment, which needs to be adaptable to the specifics of the analysed intervention, but also replicable and transferrable to other interventions or fields of activity. For this to happen the approach applied needs to adhere to the standards of scientific research. As Peter (1979, p. 6) has pointed out "behavioral measures are seldom if ever totally reliable and valid, but the degree of their validity and reliability must be assessed if research is to be truly scientific". The ambition of measuring social impact, at least in the long run, should be to judge discrete interventions and compare them with others in terms of the effects they are creating. For doing so common standards need to evolve on how to assess, specify and display ef-

fects. The theory and practice of measurement in the social sciences entails these aspects (Miller & Salkind, 2002; Zeller & Carmines, 2009).[1]

Second, social impact has come to complement other dimensions of organisational performance which are grounded in a system of metrics. Figures or other standard ways of reporting the social effects created by an intervention are thus demanded by the organisational context in which the assessment of social impact is embedded. Performance measurement systems are of vital importance to organisational functioning and striving (Haas & Kleingeld, 1999; Neely, Gregory, & Platts, 1995), and the communicative functions of an impact analysis plays an important role in this regard. As will be seen below the illustration and documentation of the results of such an analysis are relevant to several constituents. Although a comprehensive social impact measurement may be a singular event, it provides information that can be integrated into an organisation's continuous reporting system, which serves to inform the different constituents. In this respect social impact and its measurement are intertwined with social reporting (Mahoney & Roberts, 2004), sustainability accounting (Lindgreen, Hirsch, Vallaster, & Yousafzai, 2018), or blended value accounting (Nicholls, 2009a).

Against this background measuring social impact has three different functions, which are primarily relevant for one distinct constituent each: (1) strategy formation; (2) investment allocation; (3) regulation and control.

[1] In deviation from or extension of established practice I posit later on that 'measurement' should not only rest on quantitative data. Qualitative data, if assessed by rigorous methods of the social sciences, also allows for systematic and standardised analysis, which is not least visible through the discussion of causality, validity and reliability in qualitative research Golafshani (2003); Gläser and Laudel (2013); Cho and Trent (2006).

Strategy formation is most central to the organisation, whose social impact is under study (micro level). In view of the developments sketched initially, an organisation needs to account for an extended set of responsibilities in order to "[…] maintain an effective alignment with its environment" (Miles, Snow, Meyer, & Coleman, 1978, p. 547). Forming an integrated view on economic and social impacts is a central thread in stakeholder-based approaches to strategic management (Freeman, 2010, p. 40). When forming strategies in particular non-profit and governmental managers need to take into account the "public value to be created" and "the operational capacity to deliver the value" (Moore, 2000, p. 183). But also for-profit managers need to think about "impacts and purposes that stretch far beyond firm performance", if they are interested in pursuing "big strategy" (Whittington, 2012, p. 23). However, organisations can only effectively formulate their mission and operational strategy, and revise it if necessary, if they have a clear account of whether or not they make a difference for their stakeholders and how meaningful the difference is. Measuring social impact can provide the necessary information.

The information gathered through impact measurement is also needed to allocate investment decisions within organisations across a portfolio of services, projects or interventions (Wood, Thornley, & Grace, 2013). The guiding question is: Which activities are most worth pursuing? The latter, however, is even more relevant to (philanthropic) investors or other funders of social causes (meso level). Not only organisations performing activities, but in particular their financiers are increasingly interested in promoting 'blends' of social, environmental and economic value (Bonini & Emerson, 2005, p. 7). Emerson and Bonini are not the only ones to establish an explicit link between social impact measurement and 'social enterprise' as a perceptibly more effective response to complex social challenges. The link between the two has fostered the creation of new financing intermediaries (Edery, 2006), who are interested in the development of rationales for investment decisions. There-

fore they have endorsed the idea of social impact as the guiding principle in developing new performance metrics (Global Impact Investing Network, 2010).

Third and finally, measuring social impact is relevant in view of new practices in social service commissioning by the state. It is required by policy makers and regulators with regard to fostering as well as monitoring and controlling new ways of service delivery and social problem solving. Social impact bonds for instance, are an illustrative case. A social impact bond is a financial instrument where the state guarantees the repayment of the invested capital plus a premium to an investor if predefined success criteria of the provided intervention have been met (see (Edmiston & Nicholls, 2017) for an overview and (Fox & Albertson, 2011) for an application of the instrument in the criminal justice sector). Impact bonds aim to complement public grant making by private investment into social matters. In addition to the transfer of risk, they are set up to lever efficiency as well as effectiveness of the provided services. The achieved social impact needs to be assessed in order to decide whether higher effectiveness has been achieved, and as criteria to pay out the premium or withhold it from the investor (Mulgan, Reeder, Aylott, & Bo'sher, 2010).

My focus on ‘measuring’ social impact is distinct from but also shares some commonalities with other approaches at the nexus of process and effect analysis, and organisational change potentially triggered by the analysis (see Woodman, 1989 for a portrait of these interconnections). One such model is ‘appreciative inquiry’, a social constructivist approach to assessing organisational performance and redesigning organisational processes and structures in response to it (Cooperrider & Srivastva, 1987). A shared trait between appreciative inquiry and my approach to assessing social impact, is the recognition that the assessment of organisational performance—here in terms of how organisations affect their surroundings—needs to be flexible enough to respond to the nature of the assessed inter-

ventions, and also needs to maintain a certain degree of informational complexity. As will be seen as my work evolves, in particular informational complexity, if neglected, may lead to a critical loss of decision criteria to make an informed judgement about the performed intervention. Another commonality can be found in that the assessment of what works and what doesn't needs to lead to an incremental and iterative process of calibration, and if necessary to organisational change.

My approach differs from appreciative inquiry in terms of the localisation of the two modes of assessment in respect to the organisation. Appreciative inquiry starts from the vantage point of organisational change and designs it in a 'generative' process, taking into account what provides value to the organisation's stakeholders (Bushe, 2013). Measuring social impact instead in the first instance turns to the organisation's stakeholders and derives the need and implications for necessary organisational change from how those stakeholders are affected in positive or negative ways. Another difference between 'measuring' and performing appreciative inquiry is the former's focus on a systemised analysis and a predefined framework for reporting results. Both differentiating aspects place my approach more in positivist rather than constructivist theory in the philosophy of science (Bunge, 1996, 317, 335). Despite its positivist grounding my approach has a strong normative component, which will be laid out further below.

1.4 Current tools and trends

The growing fascination with social impact and the three primary functions it serves have resulted in a large variety of impact measurement tools applicable to all kinds of organisations (Clark et al., 2004; Maas, 2009; Olsen & Galimidi, 2008; Social Impact Investment Taskforce, 2014; Tuan, 2008).

Some of the tools available, although relating strongly to 'social impact' or to 'social value creation', are not really concerned with impact. Instead they provide sustainability indicators or perform organisational 'due diligence' processes to derive the studied organisations' *potential* for creating social impact (Mildenberger, Münscher, & Schmitz, 2012). The same applies to much of even the most recent literature on performance measurement in public and social purpose organisations (Berman, 2006; Epstein & Yuthas, 2014; Micheli & Kennerley, 2005; Moynihan, 2008; Poister et al., 2015). It deals with why and how performance is (or is not) measured, discusses specific tools and examines what the measurement imperative means for organisational management, but does not show how to actually measure social impact.

There is literature that attempts to do the latter, which helps fathom the breadth of indicators that can be assessed when studying the performance of social purpose organisations. However, the wide perspective taken often leads to a narrow set of assessment criteria for a single dimension. The indicators proposed by Arena et al. (2015) for studying the performance of a social enterprise in the energy sector for instance comprise: employment, contributions to local associations, the share of local suppliers bought from, the share of local families served, and the amount of CO_2 emissions reduced. Bagnoli and Megali's (2009) article on social impact in relation to a work integration social enterprise is a similar contribution. It stops at mentioning relevant areas of impact without spelling out what effective measures could be and how one would assess them.

This is consistent with observations on the practice of the 'Social Return on Investment' (SROI) as one of the most widely used impact measurement tools (Arvidson, Battye, & Salisbury, 2014; Arvidson & Lyon, 2013; Kehl, Krlev, Then, & Mildenberger, 2018; Maier, Schober, Simsa, & Millner, 2014; Millar & Hall, 2013). SROI as a method has done

a lot to promote interest in social impact and contributed to the evolvement of shared approaches and indicators. However, its practical application is also marked by significant shortcomings. The following insights relate to a comprehensive assessment of practice in SROI by means of a systematic review of more than 100 studies published between 2000 and early 2012.[2] Three main findings of the study stand out: the first one relates to the theoretical grounding of social impact measurement, the second to the rigorousness of designs and measurement, and the third to the use of measures.

First, in contrast to other metrics-related practices, such as 'social/blended value accounting' (Mook, 2013; Mook & Quarter, 2006; Nicholls, 2009a; Richmond, Mook, & Quarter, 2003), social impact measurement in the way we see it today has a shorter tradition and is much less theorised or even established in academic discourse. This is partly due to the fact that most tools for social impact measurement have been developed hands-on by philanthropic organisations such as Bill & Melinda Gates Foundation, The Acumen Fund, or The Robin Hood Foundation (Brest, Harvey, & Low, 2009). Other reasons are the low involvement of academic analysts in the performance of impact analyses as well as the dynamic development of the field with an exponential increase of the number of impact assessments in the recent past (see for both points in relation to SROI (Krlev, Münscher, & Mülbert, 2013)).

Second, as mentioned in relation to the issue of 'attribution', the questions "What would have happened anyway?" and "How much of the change can we ascribe to the activities of the organisation?" are central. There is a debate about whether these questions are best answered by experimental designs, non-experimental (observational) ones or qualitative evidence. Experimental methods are the gold standard in healthcare research (Evans,

[2] The research strategy and methodology are laid out in greater detail in Krlev et al. (2013).

Chalmers, & Thornton, 2006). But some argue that in assessing social interventions, due to varying context conditions, they are not necessarily more useful than non-experimental methods and generally argue in favour of combining quantitative and qualitative or experimental and observational evidence (Deaton & Cartwright, 2016).

No matter which design is applied it needs to be executed in a rigorous fashion and this is where impact measurement is currently lacking: Not even 3 percent of the SROI studies analysed had applied an observational control group design—not to speak of randomization—and less than 20 percent had performed any ex ante – ex post observation (Krlev et al., 2013, p. 44). That is most studies were cross-sectional, often without attempts of tracking influence of an intervention over time. About 80 percent of the analysed studies used qualitative evidence in a complementary or exclusive fashion. However, less than half of them were assessed as being of high quality. Those assessed as low quality did not perform a systematic analysis of qualitative evidence but instead provided 'stories' of a single person that participated in a programme or used other anecdotal evidence.

The low quality of many studies in these regards is not particular to SROI and might even be worse in fields with other traditions of evaluation. Development assistance is one such field, where reviewers find that "[o]f the hundreds of evaluation studies conducted in recent years, only a tiny handful were designed in a manner that makes it possible to identify program impact" (Savedoff et al., 2006, p. 17).

Third and finally, despite the ambition of assessing 'economic', 'socio-economic' and 'social' value, the SROI studies we analysed focussed on those aspects that were readily accessible and could easily be monetised, such as: reduced state benefits (for instance in case someone found employment), avoided costs (in case of prevented imprisonment), or in-

creased income and taxes (through economic development of a region). Genuine social effects in contrast, such as the strength of personal relations, the fortification of someone's personality, or political participation are largely disregarded. This happens despite the fact that such effects are central to the analysed organisations' activities, and usually more so than the effects that can be monetised.

This is again not particular to SROI. Other kinds of investigations occupied with the assessment of impact come to the same conclusion, be it a review of the measurement of social progress (Stiglitz et al., 2009) or an analysis of the impact of migration (Migration Advisory Committee, 2012): Disproportionate weight is given to measures that focus on financial returns, be they direct or indirect, in contrast to those that are less tangible and more fuzzy but central to the assessed intervention such as the promotion of social cohesion or cultural integration.

1.5 Research questions

Two main rationales underlie the disregards diagnosed above: (1) sophisticated designs are hard to implement, complicated to carry out and costly; and (2) some variables, although they might be most profound to the assessed activities, are still neglected because they are particularly hard to measure.

However, if the aim of the whole discussion about impact is to *accurately* capture, illustrate and judge the *social* effects created by organisations, we are facing a critical problem: In lack or omission of effective accounts of these specific effects we are using flawed or at least insufficient criteria for assessing the overall social impact created. The lack of theoretical grounding identified may be both, a reason for and a consequence of these two rationales. It is a reinforcer of the present deficiencies at any rate.

The research presented in this thesis therefore aims to address the prevalent gaps by focussing on two inter-related research questions:

(1) How can we define social impact and by means of which theories can we explain it so as to understand and meet the challenges it produces?

(2) Based on the theorising: How can we operationalise social impact (specifically the social effects) and measure it empirically?

Due to a vast knowledge base on capturing facets of impact in economic theory (e. g., in the fields of welfare, health or environmental economics) and neighbouring disciplines of the social sciences and humanities, it is unnecessary to 'invent something new'. Instead we need to identify how the current limitations of social impact measurement can be overcome by tapping existing points of reference.

In summary, this research establishes a theoretical foundation for social impact measurement, which informs its practical application and thereby enhances the exchange between praxis and academic research. This exchange is crucial for maintaining methodological rigor while ensuring practical applicability.

From the above it should have become clear that the question of social impact is relevant for basically any organisation, since all organisations have effects on their surrounding environment. The social impact discourse is however particularly relevant for 'social purpose organisations' (Nicholls, 2009a), an umbrella term typically applied to organisations that primarily pursue a social mission: non-profit and non-governmental organisations, charities, foundations or social enterprises. While the strategic setup and operational realization of their mission may differ remarkably, all of the listed organisational types aim at creating positive social impact or "social wealth" (Zahra, Gedajlovic, Neubaum, & Shul-

man, 2009) in that they help maintaining social stability, fighting exclusion or fostering political and civic engagement (Anheier, 2014). In consequence, social impact measurement as a 'proof of concept' and a justification for their existence is of central importance. For these reasons the theory chapters and the empirical work presented here will focus on this population of organisations. I will provide implications for the study of a wider set of organisations in the conclusions.

1.6 The theoretical argument and empirical design of the thesis

The great variety of available impact measurement tools notwithstanding, there are essentially three generic conceptual approaches to social impact that can be identified in the literature: (1) cost-benefit analysis (CBA) (Drèze & Stern, 1985), (2) life-satisfaction approaches (Kapteyn, Smith, & van Soest, 2010), and (3) the capability approach (Sen, 1985). All concepts, despite their merits, are subject to one or several of the following challenges, which are all discussed in detail in the *literature review* performed in Chapter 2.

First, the measurement problem, that is the challenge of translating social impact into measurable components (Nicholls, 2009a, p. 758; Paton, 2003, 1ff.; Smith, 2010, 135ff.). Second, the attribution problem, that is the difficulty of establishing causal links between an organisational activity and its effects, which results from confounding within complex social realities (Byford & Sefton, 2003; Milton et al., 2012). Third, the aggregation problem, manifested in a focus on final outcomes (benefits, happiness, or capabilities) and a relative disregard for intermediate outcomes and processes. Fourth, the foundation problem, that is the inhibition of the assessment of impact when individual utilities or capabilities do not add up to collective social welfare.

Against this background I promote a *conceptual advancement* in Chapter 3. I develop an approach to social impact measurement that is new but compatible with the existing approaches. Instead of focussing on the individual level in defining social impact, I access the subject from the societal perspective. I draw links from the production process of the economy to develop an image of society as a productive system, which differs from that of the economy in mainly two regards: (1) it is more complex and intertwined; (2) the capital factors employed are of a greater variety. In addition to economic capital, the latter range from natural capital to social capital. Based on this model I discuss how social impact can be understood by accounting for (targeted) shifts in these capitals. Especially social purpose organisations can be seen as capital producers. Assessing the creation of these capitals thus serves as an alternative unit of measurement for their social impact. I refer to this as the 'capital based approach'. Through conceptual reasoning and empirical examples I illustrate how the capital based approach moderates the weaknesses of the existing approaches. I then go on to explain why it is important to focus on social, cultural and political capital in particular in order to advance social impact measurement. I conceptualise each capital form in turn.

The *methodology* in Chapter 4 presents how the theoretical model developed in the previous chapters can be translated into empirical research. It also provides a general methodological commentary for the empirical cases I have analysed. This commentary only outlines the fundamental principles of the research and leaves the exact details for the individual empirical chapters.

Chapters 5-7 present the *empirical results* from the analysis of three cases located in Germany. Each study performed has taken into account multiple capitals, but focussed on one of them: (1) community oriented housing models were analysed as to their effects on the

social relations of elderly people (social capital); (2) a school-based intervention on pro-social behaviour of pupils has been tested as to how it helped pupils develop positive attitudes towards themselves and others (cultural capital); (3) a political online platform for use by citizens and politicians has been assessed as to whether it levered political dialogue, participation and engagement (political capital).

The *discussion* in Chapter 8 outlines whether and how the capitals approach is useful in informing social impact analyses. It provides a synthesis of the results and insights from all three independent, yet conceptually connected empirical cases. First, it does so as regards the learnings about the social impacts of the investigated interventions that have been made in the analyses. Second, it reviews the usefulness and limitations of applying a capital based concept of social impact to the selected organisations.

The *conclusions* in Chapter 9 provide a critical appraisal of the capital based approach. It also shows in which ways the approach has to be refined further to increase its value for the theory and practice of social impact measurement. This includes implications of the capital based approach for empirical research beyond the present cases, that is in different fields of activity or in relation to the actions of organisations that do not primarily pursue a social purpose. Furthermore, it provides links to several other theoretical strands of research. One link is established to the debates concerning the usefulness of the term 'capital' for elements other than money or physical assets. Another connection is made by outlining how the capital based approach may inform the research on (social) innovation.

2 Survey of the literature

This chapter provides the background needed to subsequently develop the thesis' main argument. The literature review abstracts from the tools referred to in the introduction and derives three generic conceptual approaches, which have previously been used or recently proposed for assessing social impact. It outlines their merits but also their shortcomings. The latter serve as a point of reference for the conceptual proposition developed thereafter.

2.1 Existing conceptual approaches to social impact

The multitude of individual tools for measuring social impact has been mapped and discussed repeatedly (Arvidson & Lyon, 2013; Clifford, 2014; Maas, 2009; Tuan, 2008) and there is little use in repeating this exercise. What I will do instead in the literature review is distinguish three main approaches to social impact measurement that originated in different areas within the scope of public choice, welfare economics and programme evaluation.

The first area is that of public sector planning where since long ago the effects of programmes, spanning from tax policies to construction projects, have been analysed by accounting for shifts caused in the 'social welfare function'. This has usually been done by applying cost-benefit analysis or variations thereof (Drèze & Stern, 1985). The second area is that of research on well-being, both at the level of the individual and at the level of the nation state. The concepts of life satisfaction or happiness are used to assess the general level of well-being as well as shifts in the latter through specific interventions (Frey, 2010; Kapteyn et al., 2010). The third area is that of the evaluation of development assistance. Evaluations in this area are employed to verify whether target groups of developmental interventions have effectively benefitted from them. The capability approach is one of the

main frameworks guiding such assessments (Alkire & Foster, 2011a; Sen, 1985; Sen, 1993b).

Each of the mentioned conceptual approaches has advanced our understanding of social impact in many regards as will be discussed below. However, the specific ways in which the approaches relate to and draft impact also provoke challenges. The following sub-chapters discuss each approach in turn. The analysis is not meant to be comprehensive but epistemic in that it highlights the strong and weak points the approaches have. The weak points serve as links for developing an alternative approach.

2.1.1 Cost-Benefit Analysis

Cost-benefit analysis (CBA) originally emerged as a tool of project evaluation in public sector planning (Layard & Glaister, 1994; Rossi, Lipsey, & Freeman, 2009). The decision of whether to perform an action, for instance for the public sector to provide a specific good or pass a certain regulation, is guided by a projection of the likely costs and benefits involved. The benefits of the action are assessed by its supposed effects on the 'social welfare function', which represents the aggregate individual preferences or utilities resulting from the consumption of the good (see Arrow, 1970; Pollak, 1979; Sen, 1970 for variations of the social welfare function).

CBA has made great contributions to the rationalisation of decision making. It has provided a standard evaluative procedure to compare two or more alternatives or to assess the situation with and without the project (Drèze & Stern, 1985). Only those projects are realised that have a positive effect on overall social welfare. CBA has always had a strong foundation in environmental protection and policy (Nas, 2016; Pearce, 1998). It has since been used to target even more complex interventions, for example in health and social care

(Byford & Sefton, 2003), and thereby come to affect an ever wider variety of fields. It has also extended its reach in terms of organisational types addressed. It has moved from its initial focus on public policy to include the activities performed by firms (see (McWilliams & Siegel, 2001) for a cost-benefit assessment of CSR) as well as those by social purpose organisations. In fact CBA is the archetype for the most widely recognized social impact measurement tools applied to such organisations (Brest et al., 2009): The 'Social Return on Investment' (SROI) or the 'Best Alternative Charitable Option'. However, in particular when CBA is applied in this context, challenges occur.

The first challenge is that there is a discrepancy between the mission of social purpose organisations and "the measuring rod of money" used in CBA (Bliss 1993, p. 417 citing Pigou, 1952, p. 11). The raison d'être of most of these organisations is other than producing financial or socio-economic benefits and they have emerged due to market (Hansmann, 1980; Titmuss, Oakley, & Ashton, 1970) and state failures (Weisbrod, 1975). Thus, if their purpose and mission is clearly other than financial, why do we assume that monetary benefits of all performance criteria are best to assess their performance? This clash of principles leads to a more specific challenge. Due to CBA's focus on monetisation, the dimensions that are most easily priced are the ones most pronounced in the analysis. This is likely one of the reasons for the observed priority given to 'tangible' indicators by CBA-based approaches such as SROI (Krlev et al., 2013). This shortcoming has been pointed out by others too (Marée & Mertens, 2012), who for this reason have criticised the dominant influence of monetisation within CBA. In consequence, CBA is particularly useful where financial aspects dominate. This is the case in work integration efforts, where social impact rests mainly on the unemployed person entering the job market and the income this creates. In many other cases applying CBA alone might offer an incomplete image of the key impacts produced.

The circumstance just diagnosed leads to a further issue: Over and again, assessors of impact invest time and effort into pinning down outcomes comprehensively, that is into covering the entire impact chain. But if some key outcomes are disregarded in cost-benefit analysis due to their intangible nature and despite their centrality to the activity under study, efforts allocated on tracing follow-up effects that depend on these key outcomes are of limited use. Violence prevention can serve as an example: When we have no exact way of assessing the formation of pro-social behaviour in adolescents, that is the variable by which follow-up effects are varied, calculating the savings produced by reducing juvenile crime is in vain.

The second challenge is that there are some tensions between the mission of social purpose organisations and the principles underlying CBA, specifically (1) that individual preferences when aggregated equal collective benefits and (2) that these preferences are typically regarded as 'fixed' or 'stable' over time. Bliss (1993, p. 425) says CBA's assumption that "[…] preferences are fixed and given is probably the least appealing idea economists have ever come up with." This is particularly relevant when analysing social impact, which in some fields is fundamentally grounded in the alteration of personal preferences and attitudes. The main social impact of a violence prevention programme for example lies in the transformation of individual preferences from anti-social actions to conformity or pro-social behaviour. This change is hard to assess when the underlying assumption of the instrument applied is that no such change happens.

Other inconsistencies arise in relation to the logic that aggregated individual preferences or utilities equal collective well-being, and that pareto efficient solutions are to be sought, where simply put no one can gain further from a reallocation of resources without putting someone else off (see Drèze & Stern, 1985 for these criteria and some of the challenges

they produce). This may collide with the 'public goods' or 'merit goods' character (Musgrave, 1951) of many social outcomes targeted by social purpose organisations. Take for instance an intervention that aims to strengthen community ties in a deprived village. These ties may suppress conflict and stimulate productivity across that community, independent of the particular involvement of individuals or their economic power. While in principle benefitting the community as a whole and in equal amounts—since everybody can potentially benefit from mutual help and a constructive climate of cooperation—some individuals will benefit relatively more from a fortified network of ties than others. In particular individuals with lower status at baseline are likely to gain against the better-off, since those more resourceful to begin with might have established a position of power within a weaker community network. Thus, disadvantaged individuals would prefer the community ties to be advanced, while the better-off might refuse it.

Any intervention now performed, must make the conscious decision of acting on the 'deserving' character of the disadvantaged individuals, while producing an intended interference with the individual preferences of others. An example of such action can be found in the empowerment and community building activities of BRAC in Bangladesh for instance (see Chowdhury & Bhuiya, 2004). There are some altering forces at play in social purpose organisations pushing them towards pursuing the 'common good' (for the general argument and exceptions to it see Alexander, 2006; Freise & Hallmann, 2014). Giving in to these forces may lead to giving priority to community outcomes above individual preferences. There is some resemblance between the orientation at the 'common good', which has to be instilled also into the communities targeted to make change viable, and "lifestyles" that generally shape individual behaviour and thereby outweigh idiosyncratic and self-centred preferences (Bliss, 1993). The trade-off between different utilities points to the need for approaching social impact from a collective rather than an individual standpoint.

2.1.2 Life satisfaction Approaches

Life satisfaction approaches mark the current end-point of a move from the more rigid judgement of benefits by the social welfare function to more subjective measures of target group and societal well-being. Shifts in this well-being through the actions of social purpose organisations could be interpreted as their social impact. Life satisfaction or happiness studies promote a stronger inclusion of the methods of psychology for measuring experienced utility (Diener, Helliwell, & Kahneman, 2010; Diener, Shigehiro, & Lucs, 2009). These methods, it is said, offer new and supposedly better ways of assessing the value people ascribe to intangible social conditions in particular, such as political autonomy (Frey & Stutzer, 2002a) or social relations (Elgar et al., 2011). The debate about happiness has led to a shift in the public perception of what makes a good life and what should be measured to assess it (Fioramonti, 2013). Phenomena studied as to their effects on human well-being are no longer restricted to economic dimensions, but comprise with increasing tendency democracy or volunteer work, but also terrorism (Frey, 2010).

The application of life satisfaction as a measure of well-being span from specific applications in health care (Fanshel & Bush, 1970; Gyrd-Hansen, 2003; Miyamoto & Eraker, 1985)[3] to the assessment of the 'good life' more generally, particularly on the national level (Graham, 2005; Tay, Chan, & Diener, 2014; Veenhoven, 2010). Kröger and Weber (2014) have recently proposed a conceptual framework for social impact measurement that is based on individual life satisfaction. They propose calculating social impact by comparing changes in the life satisfaction (LS) or domain-satisfaction (DS)[4] of an intervention

[3] The 'quality adjusted life years' applied in health economics are an intermediate construct in that they apply a cost-benefit rationale in the assessment of what a health intervention is 'worth', but use life-satisfaction as a measure for the potential benefits an intervention provides for a patient.

[4] DS is a breakdown of life satisfaction into domain-specific partitions such as satisfaction with health, education, etc.

group before and after an intervention relative to the regional, national or local LS or DS. Their proposition of using life satisfaction metrics in assessing social impact certainly brings focus into the otherwise wide landscape of indicators used. It also approaches impact by scales that have been tested across a variety of settings and found to be effective in assessing the state of and changes in individual well-being (Diener, Inglehart, & Tay, 2013). The focus introduced, however, also comes with some down-sides.

First and as posited above life satisfaction is fundamentally subjective and malleable by the circumstances—supposedly more so than other self-assessments tested by survey techniques. Individual expectations of a 'good life' are often dependent on one's life-circumstances as well as the situation others are in. In the Russian financial crisis of 1998 for example, both employed and unemployed people became happier with increasing rates of unemployment. This was explained by the fact that people can find comfort in seeing that others are not better-off or even at a disadvantage compared to themselves (Graham, 2005, p. 52). It shows that LS may be influenced by other people's life circumstances, independent of whether one's own situation has changed. While the subjectivity and sensitivity of context are exactly what LS measures are praised for when comparing living conditions across geographical location, it may negatively affect the accuracy of attributing impact to specific interventions by inviting many confounders.

In addition to the influence of context, LS is subject to habituation. Kahneman and Thaler ((2006), also (Diener, Lucas, & Scollon, 2006)) for instance point out that changes in LS are influenced by the novelty of a circumstance or the attention it gets. They found that LS is restored to a previous level, once the property has moved out the affected person's awareness. This held even for events as life-changing as paraplegia. Another issue in relation to the subjectivity of LS is that of deviant habits. It is revealed when we think about a

person who finds herself in an objectively bad situation, which is however a source of delight to said person. It applies where truancy is preferred to facing challenges in school or where people objectively 'suffer' from an addiction, but perceive it as a source of joy. Satisfaction is no good measure for effectuated social impact in such cases.

All of these effects show that we cannot always assume a uniform mechanism by which positive effects are translated into higher levels of LS and ergo a systematic comparison based on shifts in LS becomes difficult.

The second issue with LS lies in that it is an aggregate of a variety of 'feelings', almost equal to an index. This makes life satisfaction an all-encompassing instrument—again a property that may be appreciated in assessing life circumstances. Aggregation, however, may come at the cost of explanatory value, since individual aspects can become inseparable and layered. The same argument holds for most domain-satisfaction levels, since although narrower, health satisfaction for instance is still broad. Of course the explanatory value of every index is reduced by aggregation. However, there is a difference between indexes that are composed of several punctuated sub-measures and such that test an aggregate measure to begin with, that is cover many separate aspects by few or even a single question, namely: "How satisfied are you with X?" To put it pointedly: Using changes in LS or DS for assessing the social impact of an intervention is in some ways similar to asking for patients' health satisfaction when testing new blood pressure medication while disregarding the very change in blood pressure.

I am using the example of an intervention discussed by Kröger and Weber ((2014) in reference to (Suldo, Savage, & Mercer, 2014)) to outline the issue further: They discuss a "group wellness-promotion" intervention. They test the effect of that intervention by com-

paring the mental health satisfaction (DS) of a group of "disadvantaged students" who participate in the intervention against the national average level of that DS. It is a challenge to arrive at a reliable estimation of social impact in this setup, where a very specific group is studied against the general population with regard to one single measure. This applies in particular, since differences in mental health are modified by the immediate social environment, comprising numerous factors (see Norstrand, Glicksman, Lubben, & Kleban, 2012), which will not be distinguishable. This evokes questions about whether attribution between the intervention and the effects can be ensured. More important though is that it is impossible to find out how processes and intermediate outcomes might play a role by putting so much weight on the aggregate level.

2.1.3 The Capability Approach

Sen (1993a, p. 444) has argued that an orientation towards functionings, also referred to as a person's ability to function, would omit the flaws produced by assessing social welfare on the basis of individual preferences or utilities, while preserving the ability to judge "against the enforced deprivation of living standards." The combination of individual functionings is referred to as a "functioning n-tuple" and the set of such functioning n-tuples is called 'capability', from which a person has the freedom to choose one functioning n-tuple (see Sen, 1985), which then determines the person's standing and capacity in life.

The capability approach (CA) is widely used to judge social impact in the context of development assistance, more specifically in that of poverty alleviation (Alkire & Foster, 2011a; Alkire & Foster, 2011b). Such assessments are based on an understanding of poverty as a deprivation of a person's freedom to choose from a multitude of possibilities to function and thus a restriction of her capability, rather than a mere lack of income. Capabilities include a person's status in the community and her participation in community life,

her possibilities to influence political decision making, or her ability to appear in public life without shame (Sen, 1993b). In this regard Sen's approach seems more useful in thinking about social impact than CBA or LS, as it is closer to those 'social effects' that have been found lacking in current practice. Recently it has for example been used to gauge the effects of EU employment initiatives that push for social innovations to decrease barriers that marginalise certain groups of people (Jacobi, Edmiston, & Ziegler, 2017).

The main challenge of the capability approach derives from its aforementioned strength, namely its proposition of highly complex final outcomes (capabilities) that do not only resonate with the mission of social purpose organisations working towards them, but reflect what actually matters to the target group. The number, the relation and hierarchy of capabilities have been debated for long and Sen himself has acknowledged the approach's incompleteness in this regard (Sen, 1993b). Ever since, others have tried to work on a list of central capabilities and examples include the "ability to affiliate and interact with others" or the "ability to freely develop human emotions" (Nussbaum, 1993; Nussbaum, 2003). Although this effort has helped systematizing capabilities, the proposed ones remain hard to grasp. Supposedly for the difficulty of measuring the latter, applications of the capability approach often fall back on multidimensional, yet simple poverty alleviation measures, such as household expenditures or standardized physical health assessed by the Body Mass Index (Alkire & Foster, 2011a). This outlines difficulties in the operationalisation of measures when CA is applied.

Another issue may arise through CA's focus on individuals' "freedom of opportunity" (Sen, 2001). Increasing individual freedom is positive where people are deprived in their capabilities by the circumstances, as is often the case in contexts of poverty. A lot has been said about listing and prioritising freedoms and functionings at the intra-personal level

(Alkire, 2005). Sen's own assessment is that any such list cannot be predefined but must be subject to a "process of public reasoning" (Sen, 2005). Much less has been said about interactions and potential clashes between the freedoms of different individuals at the interpersonal level (Osmani, 2016), which become more central when we move out of contexts marked by deprivation. Increases in the capability of some people might decrease the capability of others. This can affect the assessment of impact, as illustrated here: Promoting the freedom of opportunity of an offender, for instance by trying to re-integrate the person socially, might restrict the capabilities of the person's fellow citizens, in particular in cases of recidivism. On the other hand putting him to jail might decrease his own future capabilities and the immediate capabilities of his kin, leaving a family without a son or a father.

Priorities need to be formed in such cases and the freedom of one party cannot be promoted while restricting the freedom of others. A "process of public reasoning" must come in here too. The level on which the process needs to make a decision is likely detached from that of individual capabilities and other frameworks would need to be used. Osmani (2016) argues that the normative framework of 'human rights' can be useful. While this is likely true in the context of human development, the framework is restricted in its relevance to the present example. Decision factors to be considered include: the distress caused to all involved parties; the potential effectuation of mind-shifts in the delinquent person, comprising a variety of sub-impacts; how social relations and constellations are affected etc. Available normative frameworks often do not offer guidance on how to prioritise impacts.

2.2 Summary – challenges in existing approaches

The previous chapters have clearly outlined the great merits of the existing options for measuring social impact, but they have also helped carve out four distinct challenges the approaches face. All four challenges may affect all approaches discussed, but each chal-

lenge is usually most prominent in one or two of them. The challenges, the approach(es) they are most prominent in, and their specific consequences are summarised in Table 2.

Table 2 Challenges in existing conceptual approaches

	Approach	Point of reference	Consequences
Measurement problem	CBA	Variables tangible and monetiseable	→ narrow set of variables
	CA	Variables complex and multi-facetted	→ simpler measures
Attribution problem	LS	Subjectivity: context and habits/habituation	→ confounding
Aggregation problem	LS	Priorisation of final outcomes	→ neglect of intermediate outcomes and processes
Foundation problem	CBA	Individual utilities or preferences	→ trade-offs at group level
	CA	Individual functionings and freedom of choice	→ clash between different individuals' freedom

The first prominent challenge is the 'measurement problem', which refers to difficulties in identifying, defining and capturing social impact (see also Anheier et al., 2014). It is most prominent in CBA and CA, but materialises in different ways in the two approaches. In CBA it mainly arises because of a striving for monetisation and a focus on 'tangible' or quantifiable variables (see also Marée & Mertens, 2012). As a consequence some relevant variables are mostly disregarded in the analysis. In CA the opposite is true: the measurement problem occurs not because the perspective on effects is narrow, but because it is especially wide. Capabilities are often multi-facetted and of a complex nature. The capability approach for instance uses abstract concepts, such as "dignity" (Nussbaum, 2003), which are hard to break down into measurable components. In other words, the operationalisation of capabilities can be difficult, which may make researchers focus on simpler yet insufficient measures.

The second challenge is the ‘causality or attribution problem’, that is the difficulty of establishing causal links between performed activities and the social impact produced (see also Anheier et al., 2014). It applies mainly to life satisfaction, which is affected by a variety of personality traits (Wu et al. 2009 discuss for example optimism; I have referred to habits and processes of habituation) as well as context conditions (for instance the situation of others). These factors, especially when combined, increase the probability of confounding and potentially decrease the reliability of the analysis or the effort that needs to be invested into controlling for confounding.

The third challenge is the ‘aggregation problem’ and most apparent in relation to life satisfaction. Life satisfaction is a composite measure that bundles the influence of many different factors (Elgar et al., 2011). By assessing the composite measure only, without a clear account of which single effects contribute to the variable, effectuated changes along intermediate outcomes get blurred. This also makes it hard to trace the process and pathways by which performed activities led to impacts.

The fourth challenge is ‘the foundation problem’, which refers to the grounding of the approaches at the individual level. It applies to all three approaches, but has been exemplified in relation to CBA and CA in particular. The orientation at the individual is in no way wrong, and it is necessary since individuals are usually the main group affected when we think about impact. The relative neglect of the group or collective level, however, can lead to difficulties in the analysis. While the issue in CBA lies in the different preferences people might hold towards a shared development (I have used the example of increased community connections), in CA it lies in different effects that might arise from the restriction or promotion of individual capabilities (re-integration programme versus conviction).

The measurement, attribution and aggregation problems depend mainly on the specific properties of the variables typically measured by each approach (monetary benefits, happiness, capabilities). The foundation problem in turn is grounded in the individual level focus of the measures (preferences and utilities for CBA and LS, or functionings and freedom for CA).

Some of the challenges discussed are interrelated. For instance complex measures as found in CA are not only hard to operationalise, their multi-facetted nature may also complicate attribution. The same holds for the application of aggregate measures. They make attribution more difficult through neglecting intermediate outcomes that are at closer proximity to the performed activity. However and as can be seen above, despite the connections each challenge is sufficiently specific and related to the approaches in different ways so as to form a separate category.

Now that we've identified the central challenges in the current conceptual approaches to social impact measurement, I move on to develop an alternative conceptual proposition for studying social impact, which if not able to rule out all existing shortcomings, can help moderate some of the challenges.

3 Conceptual advancement

This chapter explores how we can address the gaps in the present approaches to impact measurement. It first identifies where the stated problems come from by looking at various impact chains and develops a new approach by drawing on the literatures of performance measurement and economic productivity.

3.1 Measuring social impact – a new approach

Here I develop a new alternative to measuring social impact: the capital based approach. In response to the 'foundation problem' discussed in the previous chapter, I access social impact not from the individual level but lift the discussion to the macro level. The subchapters describe how the creation of social impact can be interpreted as a contribution to social productivity. Social productivity refers to society's ability to stimulate positive final outcomes for its individuals by enabling social transformation processes, which are similar to those happening in the productive system of the economy.

Departing from this, I outline how especially social purpose organisations contribute to social productivity by creating particular forms of capital for their target groups. I then go on to illustrate the forms of capital that are central to the activities of social purpose organisations and the assessment of their social impact. I argue that a systematic analysis of the change achieved in one or several capital forms can serve as the unit of measurement for the social impact produced by an organisation. The capital forms studied are derived from the goals that are most strongly pronounced in an organisation's mission. The chapter explains how this strategy may help mitigate the challenges identified in the previous chapter.

Social, cultural and political capitals are highlighted. Not only do they cover a large share of the potential impacts of social purpose organisations, they have also seen relative neglect as compared to for instance human capital (skills and expertise) which is more closely related to the concept of economic productivity. The significance and meaning of each capital within the new approach is discussed in turn. These portraits build a bridge to the methodology and the empirical work that will outline how the capitals can best be operationalised and effectively measured for assessing the social impact caused by three distinct interventions.

The capital based approach comes with several advantages: The focus on capitals moves the unit of analysis closer to the organisational activities performed, and thus gives priority to process and intermediate outcomes and reduces confounding at the same time. It also provides links to a magnitude of previous research, which aids the operationalisation of measures.

3.2 Social impact as a contribution to social productivity

In addition to its usefulness in clarifying terminology, the logic model referred to earlier (Weiss, 1972; Weiss, 1998) can help us understand why the measurement, attribution and aggregation challenges arise when assessing social impact. For this we need to regard whether and how the assessment of other types of impact differs from the measurement of social impact. The analysis I perform in the following needs to be universal rather than particular: Since different types of organisations have different types of effects in different areas of society, the question about which impact organisations create on society is best approached at the macro level. Rather than asking how specific organisations affect specific individuals, I begin with looking at how different ‘spheres of production’ affect society. This very approach to studying impact is contrary to that of the existing options, which

relate to impact via the level of individuals. As we have seen in the previous chapter this can lead to the 'foundation problem'. Taking an alternative mode of access will prove to be a potential response to it.

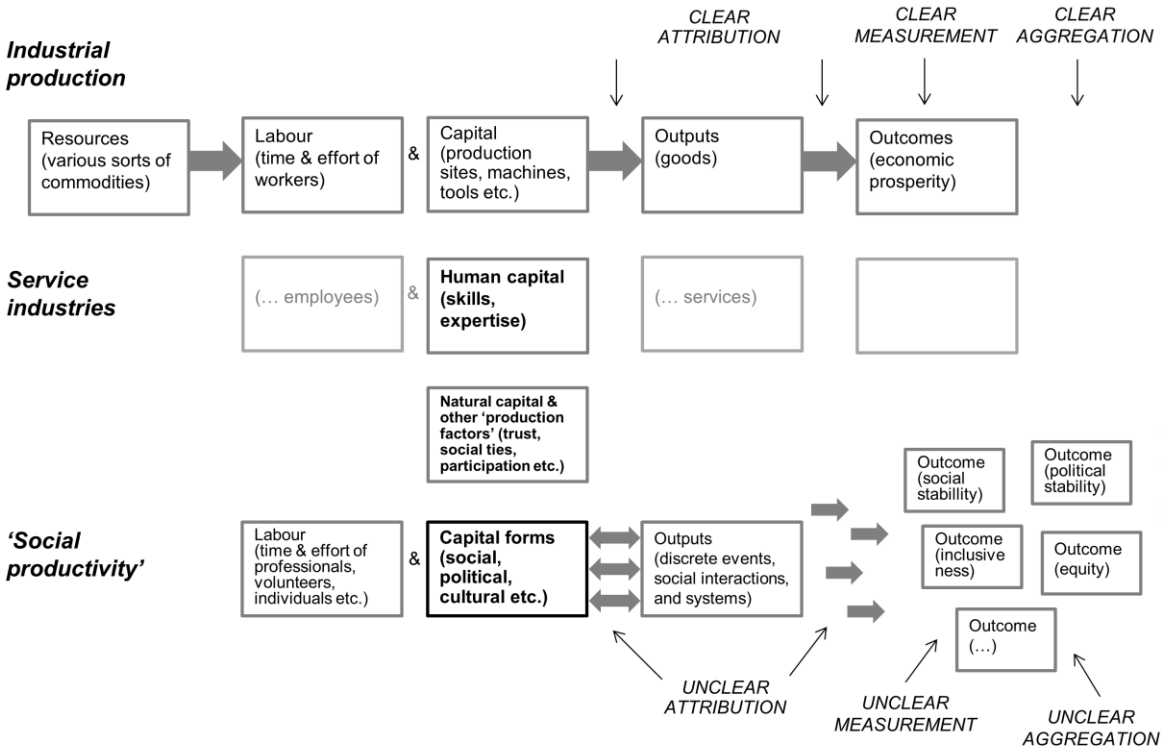
Let us first consider impact in a standard model of industrial production. In the latter the impact chain translates to: profit maximization (mission) – a production process, which involves but also creates resources, labour and capital (activities) – discrete consumer or investment goods (outputs) – a contribution to economic prosperity (outcome). Salter's (1969) study of "Productivity and Technical Change" is a useful reference for drafting a clearer picture of the intermediary steps in the productive system of the economy. What is more, it lays out the principles of productivity in a way that, if unintended, not only aids the study of the economic, but also that of the wider social production system. Salter shows how economic productivity depends on resources, labour, and capital as an embodiment of state-of-the-art production techniques.

Capital mostly comes in the shape of production sites, machines or tools and thus 'hard' and discrete factor units. Salter (1969, p. 17) relates to this in reference to Schumpeter's statement that the production function is "a planning function in a world of blue-prints where every element that is technologically variable can be changed at will" (Schumpeter, 1954, p. 1031). The industrial production process and all subsequent steps to follow are illustrated in Figure 1.

The classical view on economic production reaches its limits when we move from classical industrial production to today's modern service industries (also illustrated in Figure 1), where the significance of clear cut factor units declines and that of 'softer' and more abstract production factors increases. These factors include 'human capital', which refers to

education, skills, knowledge and expertise (Becker, 1964). It has transformed employees and workers, formerly treated as an anonymous mass of labour—merely accounted for by the number of man-hours worked, with only negligible further effects on productivity—to a highly valuable and adaptable production asset, where productivity is altered by individual skills. In parallel and driven by the sustainability agenda, nature has moved from being regarded as a mere resource to being regarded as capital, which possesses an individual reproductive capacity and an inherent value beyond its application in the production process (see Costanza et al., 1997; Hawken, Lovins, & Lovins, 1999 on natural capital).

Figure 1 From industrial production to social productivity



As illustrated in the third line of Figure 1, there are also other production factors that have become important in the modern 'socially embedded' economy: trust (Gluckler & Armbruster, 2003; Raiser, Rousso, Steves, & Teksoz, 2007) (for establishing inter-firm cooperation, but also customer loyalty); social ties (Jones, Hesterly, & Borgatti, 2005; Powell & Grodal, 2005) (for participating in network structures and levering connectivity); stake-

holder participation (Andriof, Waddock, Husted, & Sutherland Rahman, 2003; Clarkson, 1995; Freeman, Harrison, Wicks, Parmar, & Colle, 2010) (for ensuring social legitimacy); norms and values (Kolb, 2008) (for meeting citizen's ethical preferences, including for example that of sustainability).

All of this underlines that the traditional production factor of economic capital (physical assets, property, finance, etc.; see Hennings, 2005) is complemented by other production factors, which along with human capital and natural capital can be categorised as capital forms. Social networks and trust have been discussed as components of 'social capital' (see for instance Coleman, 1988; Putnam, 2000). Involvement and participation are aspects of 'political capital' (see for instance Huntington & Nelson, 1984; Verba, Nie, & Kim, 1978). Values and norms can be called 'cultural capital' (in reference to for instance Bourdieu, 1986). Natural, human, social, political and cultural capital, just as economic capital, can be regarded as 'productive assets' which are essential to any process of production. Without their existence transformation cannot take place and their neglect can lead to the complete breakdown of the productive system.

Firms have started applying various concepts of capital within 'socio-eco-efficiency' analyses of industrial production (Möller & Schaltegger, 2005; Schmidt et al., 2004). These tools at the firm level often relate to the capital approaches proposed by the World Bank as components of an extended social indicator system (Berger-Schmitt & Noll, 2000). In a similar way Stiglitz et al. (2009) refer to sorts of capital as important variables for assessing social progress in national accounts. Please note that the emergence of attention for a wider set of capitals occurred in parallel to the evolvement of interest in social impact in economic thinking as outlined in the introduction.

Let us finally think about society as a productive system as illustrated in the 'Social Productivity' line of Figure 1. While the outfits of the economic production and the social production system are essentially the same, the individual components and also the links between the stages differ. Let us consider the capitals involved. Despite the higher appreciation social aspects and impacts have gained in economic thinking, the main purpose of economic activity remains in generating financial surplus. The difference between classical industrial production and the modern economy is that economic activity is performed with an increasing stress on the preservation of social, political and cultural capital for instance and also by tapping these capital forms in the production process.

For society to be socially productive though, it has to actively create social, political and cultural capital. Social purpose organisations contribute to the creation of the capital forms other than money or physical assets in a major fashion. These capitals represent important goals in their production process. One could interpret social purpose organisations as 'capital producers' or 'suppliers of capitals'.[5] The capitals enable outputs within society embodied by events, interactions or systems: For instance they initiate a network of neighbourhood help in a community building context. They support the formation of actor coalitions in favour of certain political positions. Or they spur information exchange among citizens on discussion platforms in the areas of consumer protection or responsible consumption.

The societal outputs are harder to describe in well-defined units than economic products. What is more, there is an interaction between the formation of capitals and outputs with a reinforcing relationship between the two. Social capital is a pre-requisite for citizens to help each other, but a formalised system of neighbourhood help also enhances social capi-

[5] Social purpose organisations are of course not the only ones to contribute to the formation of capitals. Less formalised processes within civil society or families, but also policies or some firm activities do so too.

tal. This reciprocal relationship is marked by the use of two-sided arrows in Figure 1. As compared to economic production, there is often also a greater variety of capitals involved in forming an output. In addition to social capital, an inclination to care for others, and a corresponding normative mind set (cultural capital), is needed for establishing the system of neighbourhood help. The multilevel links between capitals and outputs is illustrated by the use of several arrows in the figure.

Eventually the societal outputs lead to higher level outcomes, such as: justice, equity, political and social stability, or inclusiveness. The link between outputs and outcomes is even more multilevel and scattered than the one between capitals and outputs, which is manifested in the more dispersed positioning of arrows and outcomes in Figure 1. When taken together the steps just outlined determine a societies' 'social productivity'. With social productivity I refer to society's capacity to function, that is its ability to enable positive final outcomes by means of social transformation processes in parallel to those happening in the productive system of the economy.[6]

The procedural description of transformations marks the critical importance of capitals in shaping social productivity. Since social purpose organisations contribute in a major fashion to forming these capitals, the assessment of the latter can serve as the unit of measurement for the organisations' social impact. I refer to this understanding as the 'capital based approach' or the 'capitals approach', two terms which I use interchangeably.

[6] Please note that this is not to be confused with 'social production theory' which offers a cognitively grounded approach to systematizing the contribution of instrumental goals to subjective well-being (Ormel, Lindenberg, Steverink, and Verbrugge (1999)), nor is it to be mixed up with the current policy use in the UK, where 'social productivity' refers to a greater recognition of citizen's preferences in public service provision (Kippin and Lucas (2010); Kippin, Hauf, and Sadique (2012); Commission on 2020 Public Services (2010)).

3.3 Challenges in assessing social impact within the system of social productivity

Beyond the insight on how to approach social purpose organisations' impact, which will still be spelled out in greater detail, the conceptual discussion and Figure 1 also help us understand the challenges, which occur when we try to pin down social impact within the system of social productivity.

We have seen that when we think about the impact chain of social productivity complications arise: A critical assessment problem is located in the difficulty of breaking down the above mentioned final outcomes (justice, equity etc.) into measurable components. One of the key challenges is that no equivalent to the market price of a consumer good exists, and even less so a standard accounting framework for aggregating outputs to outcomes, such as the one found in the economic production system (Nicholls, 2009a). As regards the latter, one has to bear in mind the decades, if not centenaries, it has taken for a socially recognized and largely uniform reporting standard to evolve (Burchell, Clubb, Hopwood, Hughes, & Nahapiet, 1980). If it took so long to establish an accounting framework for the 'economic value added', a comparatively uniform and straightforward construct at face value, we may not expect an easy answer on how to define, break down and measure concepts as intrinsically complex as social stability.

The complexity of final outcomes that social purpose organisations contribute to has been recognised in previous research (Clark et al., 2004, 3f.; Nicholls, 2009a, p. 758; Paton, 2003, 1ff.; Rotheroe & Richards, 2007, p. 32; Smith, 2010, 135ff.; Zappala & Lyons, 2009, p. 3). The fact that many outcomes appear unfathomable has made many organisations fall back onto measures of service utilization (Smith, 2010), process metrics on service delivery (Sawhill & Williamson, 2001) or other efficiency criteria.

The outcomes that matter here thus come with two particular challenges: measurement and aggregation.

Another critical assessment problem arises by the multitude of outputs, but particularly when we move from outputs to outcomes, where we meet difficulties when trying to establish a well-defined causal sequence. Difficulties in attribution have been described in a variety of contexts, where social purpose organisations can be found, for instance in complex health and social care interventions (Byford & Sefton, 2003) or in community development (Milton et al., 2012). Byford and Sefton (2003) outline how some health interventions are easy to assess, since there is a clear connection between the intervention (for example medication) and effects (reduction of mortality or symptoms, such as a decrease in blood pressure). They go on to compare such interventions to for instance "community mental health or social care services", whose effects for their multi-layeredness are much harder to trace. As Byford and Sefton report, this may not only lead to inaccuracies in linking the effects caused to performed activities, but also to difficulties in prioritising which effects are important at all.

Milton et al. (2012) look at a variety of community development interventions, which lever the difficulties just mentioned even further. This is because outcomes in community development may be either an endpoint or an intermediary step triggering follow-up effects. The question is: Is community development a goal in its own right, or is the formation of a community always a vehicle meant to promote something else? We thus see that the outcomes themselves, but in particular their connection to organisational activities and throughputs along the way come with a third challenge: attribution.

Readers will have noted that measurement, aggregation and attribution have been identified as problems occurring in the existing approaches to social impact measurement. The discussion here shows that these problems are only partly related to the specific focus and tradition of the measurement approaches discussed before. The challenges are also provoked by the very nature of the effects we are interested in when we talk about social impact. The following sub-chapters outline how the capital based approach relates to the three challenges we face, and also how it is positioned against the 'foundation problem'. The latter is less a property of the effects to be investigated, but arises in some situations due to the conceptual grounding[7] of the other approaches.

3.4 Properties of the capital based approach to social impact

Based on the insights above I propose a new conceptual definition of the measurement of social impact:

> *'Social impact' refers to the effects organisations produce on final social outcomes through their targeted formation of durable capital forms. Relevant capitals include: natural capital, human capital, social capital, political capital and cultural capital. The systematic assessment of changes in these capitals can serve as a measure of social impact.*[8]

This definition is marked by a strong normative grounding on two accounts, both relating to the level of analysis but in different ways. One argument relates to the level from which effects on the target group are approached. The other argument relates to the location of the analysis in relation to the impact chain.

First, my definition proposes to assess social impact from the collective rather than the individual level. It does so because we have seen that several problems emerge when collective welfare is measured by simply summing up individual welfare. The trade-offs be-

[7] Utilities and functionings of the individual.
[8] The theoretical reasoning and definition could likely be transferred to studying the destruction of capitals too. This research, however, focuses on positive social impact.

tween effects at the individual level and at the group level or the clash between the freedoms of individuals are only two cases in point. The centrality of the collective level is closely linked to the concept of social productivity, which vitally depends on the fruitful collaboration and interaction of groups of individuals. The orientation at the collective level does not neglect the individual or her freedom to act, yet stresses that as social impact always entails an orientation at a larger group of people, measures that are 'other-directed' to begin with offer benefits over such that target the individual in an isolated fashion. It will become clearer in the detailed discussion of capitals that these exhibit the required 'other-directed' qualities.

Second, the focus on capitals shortens the causal impact chain to be traced and helps define the key components to be measured in the analysis at the same time. The capital based approach to social impact relates to the key capitals that enable the eventual outcomes, in parallel to how the price of labour and capital could be used as a basis for the minimum price of aggregate outcomes. We can in fact identify close connections between specific capital forms and final outcomes. This is similar to how the differences in economic productivity between developing countries and industrialized countries can be explained by relative differences in sophisticated production machines (compare to Salter, 1969, p. 15).

A first example is political stability. It mainly depends on civic participation and the possibility for interest groups to assert their positions or at least make them heard. And societal division and a relative lack of citizen's political capital has led to the 'Arab revolution' (Mekouar, 2016) or the Euromaidan protests in Ukraine. Political stability also depends on other factors such as economic well-being, and austerity can be a source for social upheaval. Some authoritarian systems are viable although or rather because civic participation is suppressed. Yet, a political system that enables 'effective citizenship' (Linz & Stepan,

1996) will have higher chances to sustain itself than one which lacks participation and engagement. A second example can be found in Putnam's analysis (2000), which shows that 'individualism' and a lack of social capital does not only produce individuals who 'bowl alone' but threatens social cohesion and stability more generally. As a long-term consequence this may negatively affect economic productivity and political governance (Putnam, Leonardi, & Nanetti, 1993). A third example is the close connection between justice and equity in society on one side and shared norms and values (cultural capital) on the other, since as Rawls (2005, c1971, p. 60) put it: "[t]he desire to […] treat similar cases similarly, and to accept the consequences of the application of public norms is intimately connected with the desire […] to recognize the rights and liberties of others and to share fairly in the benefits and burdens of social cooperation."

I acknowledge that each of these links can be contested and numerous counter-examples can be found. One could for instance point out that some individuals will appreciate individualism more than community, and that community itself has different meanings depending on geographical setting and cultural context. Even more so does this apply to norms and values. And yet there clearly is such a thing as a common understanding of the pillars of social viability and progress, or 'social productivity' to apply the terminology developed here. Otherwise there would be no international agreement on the 'millennium development goals' (United Nations, 2012) or the extended 'sustainable development goals' (United Nations, 2015). These include the 'promotion of gender equality' and 'women empowerment', two goals which among other things point at shifts in mind sets or a higher degree of participation, and thus facets of cultural capital and political capital. They also refer to 'partnerships for development' that is community and joint action, which can be considered parts of social capital and political capital. Also, if there was no common understanding of what makes social productivity, we certainly wouldn't face an international land-

scape of non-profit and non-governmental organisations working towards a similar set of objectives (see Salamon & Anheier, 1992 for a classification).

As posited earlier, what exactly needs to be assessed in a social impact analysis should become clear when carefully analysing the studied organisation's mission. Some of the postulated goals can be captured directly by the capitals and their components, for instance whether social connections are fortified or community is being built (social capital), whether people show an increased level of respect for others (cultural capital), or whether they have increased their degree of pro-activity towards a certain cause (political capital). All of the latter can be assessed in the short or medium term. Some other goals may refer to follow-up effects of the latter. Examples comprise: improved health status as a result of decreased individual isolation; a life free of conflict and crime due to positive mind sets; individual or regional prosperity promoted through the actions of actor coalitions, etc. These are not only located at a higher level of aggregation, but will likely need more time to evolve, and are thus to be considered long term effects.

Given the resource intensity connected to a careful and in-depth analysis of capital formation, the effort that would need to be put into performing an encompassing analysis of effects in the future (timewise extension of the model) increases proportionately. This makes the realisation of such an analysis difficult, in particular for smaller organisations. What is more, it exacerbates the attribution and aggregation problems, and to such a degree that the perceived value-added of a broader analysis is undermined by the increasing insecurity of whether detected change can be linked to performed activities. If one had the ambition to pin this connection down reliably, analysts would likely need so big a sample size and so long a time horizon that it can hardly be realised in relation to single organisational activities. In this context it is important to recognise that the implicit claim of my approach

is that more of the capitals is better. Where organisations work with the explicit aim of promoting capitals, as referred to above, or even more so where they want to fix the lack of capitals, this claim is justified. It is hard to imagine cases where organisations would create too much social, political or cultural capital. Instead, where capitals are seen as vehicles for follow-up effects, the analysis of capitals specific to an organisation would need to be brought together with larger scale general research on the effects of social connections, a human environment marked by empathy and respect, or by a high degree of civic action.

Performing the latter step, however, will only be necessary where social impact assessment is performed with the aim of arriving at the absolute value created, and the latter would still depend on the level of capitals created. For judging the relative performance of discrete interventions—the primary purpose of performing social impact measurement from the stance of social productivity and with regard to its three functions portrayed above—it will be necessary to know how effective the assessed interventions are in creating capitals relative to other interventions or the absence of any intervention. What is more, if we don't have a robust account of the capitals created, the accuracy of analyses on follow-up effects will not only be deteriorated by confounding, but be based on defective assumptions.

A similar line of reasoning can be applied to how my approach relates to 'systemic change' that is so often sought after in view of today's social challenges. The pursuit of 'systemic change is embodied in the terminology of 'changing the world' (Bornstein, 2007; Elkington & Hartigan, 2008), 'paradigm shifts' (Elkington, 2010), 'sustainable change' (Nicholls, 2006), 'reconfiguration' (Nicholls & Murdock, 2012a), or 'catalytic change' (Waddock & Post, 1991). In some cases systemic change might be detected in a fundamental upwards shift in the capitals produced, as compared to the absence of an intervention or the presence of an alternative intervention. Where systemic change can only be assessed by a

broader set of outcomes or through extending the number of affected stakeholders considered (scope-wise extension of the model) the same logic as the one outlined in relation to long-term effects holds. Such assessments would benefit from a combination of specific organisational analyses and larger scale studies on follow-up effects. The more likely case is that systemic change cannot be judged solely by looking at produced effects, but requires an analysis of the organisational model or system by which these effects have been produced. The tackling of "grand challenges" is often tied to some element of innovation and performed within participatory structures (Ferraro, Etzion, & Gehman, 2015) such as 'innovation systems' (Wieczorek & Hekkert, 2012). In other words 'systemic change' is often spurred through system-based innovations, or through radical ideas such as reinterpreting as assets what is typically thought of as liabilities (Hockerts, 2015).

The intersection between the model of production and the effects produced is where my approach is closely tied to theory on organisational structure, functions, and processes which are essential to strategy formation in any kind of organisation (Miles et al., 1978). Through being closely tied to organisational activities and strengthening the role of processes and intermediary outcomes in the analysis of social impact, the capital based approach helps identifying which factors on the aforementioned levels enable the production of capitals and also where adaptions would need to be made to achieve greater impact. Capitals based reasoning, in principle, could even be applied to the specific value-added of certain actor traits or constellations, for instance if an innovation system enhanced coordinated and joint action or enabled the inclusion of new types of actors. Thus, my model and definition cannot only be extended horizontally, that is at the stage of outcomes with regard to addressing long term effects (time-wise extension), or in terms of the range of effects and affected stakeholders considered (scope-wise extension). It can also be extended vertically at the level of capitals and complemented by an analysis of the organisational

conditions enabling the analysed entities or groups to function as ‘capital producers’. Since the latter is essential to understanding how social impact is produced, but less so to the very measurement of social impact per se, links of the capital based approach to studying innovation, and to organisational theory and learning are focussed on in the discussion and conclusions chapters.

It remains to be stated that no matter from which angle and with which time horizon we look at social impact and the systems producing it, the capital based approach enhances our ability to analyse and compare the yielded effects and identify the processes by which they have been produced.

One of the further assets of the capital based approach is its grounding in a variety of social sciences and that the different capitals referred to are relevant to a large share of the fields of activity in which we find social purpose organisations. Thereby the capitals approach offers a grounded and interconnected set of dimensions and categories for the assessment of social impact. The beauty of it lies in the fact that research performed within the capitals framework can be related to research in a variety of disciplines ranging from psychology to macro sociology. It does not only enhance such research, but can also benefit heavily by drawing on it. Different capitals have found recognition in the literature and there is at least a minimum consensus of which variables belong to each of the capital dimensions. As will be seen in the conceptualisation to follow, but in particular in the empirical research, some elements of the capitals are long established. For the latter it is the academic discourse to date that defines whether a specific variable ‘counts’ as part of that capital. In other respects the capitals and their constituent components need to be re-drafted or drafted anew. The variables I employ are to be seen as propositions, which will have to stand their ground or need to be revised and extended in future research. Just as in standard perfor-

mance frameworks or financial accounting, metrics, designs and procedures as well as the competence of interpreting social impact measures need to evolve in an evolutionary process of establishing meaning (Ruff & Olsen, 2016). Tying this process to the canon of social science research, and a comprehensive framework, such as the capitals one, should prove highly beneficial.

This additional normative grounding of the capital based approach may expose it to critique from scholars of the philosophy of science that equal normativity with totalitarianism or non-objectivity (see for instance Høgsbro & Shaw, 2017 on Foucault's critique of the Critical Theory of the Frankfurt School of Philosophy as prescriptive judgement). This position is embraced by some scholars in fields relevant to social impact research: While Biesta (2007) in the context of education and Sund and Lysgaard (2013) in the context of sustainability acknowledge that both areas depend critically on normativity, they also stress that normative judgement needs to evolve in a democratic process that should be held separate from the research process. Biesta goes as far as claiming that evidence-based thinking mainly promotes efficiency and effectiveness and thereby crowds out participation, contestation, deliberation, and professional judgement, all of which change in education should be based on in his view. If his argument was taken literally, one might conclude that empirical evidence of educational effects is not only redundant but counter-productive and that the design and change of educational formats and practices should be based on discourse alone.

Next to social constructivist thinking in the tradition of Foucault, the unity of empirical research and normativity has also been challenged from the opposite direction. Max Weber (1949; 1988 (1920)) and David Hume (1978, 469f.) are often referred to when it comes to

value-free judgement in scientific inquiry.[9] Hume's empiricism and Weber's "instrumental rationality" (Weber, 1978, 24f.) have promoted a clear separation between descriptions of social realities and prescriptions of what should be. Mackor (2011, 70f.) claims that even legal doctrine, although it is trying to answer the question "what should I do" and is therefore sometimes referred to as a "normative discipline" (see Turner, 2007, p. 37 for the latter), "just like other practical sciences, […] does so by offering norm-descriptions and norm-contentions" that need to be distinguished from "norm-recommendations as normative claims".

The counter-view to the separation between empirical research and normative judgement has not only been promoted by the normative grounding of Critical Theory (Bohman, 1999), of which Habermas' (1984) work is only one example. Hilary Putnam (2004a) for instance has also challenged the "fact/value dichotomy". Spicer (2008, 57f.) posits that comparative social science research in specific areas, in his case in public administration, and all the vocabulary employed in it such as 'government' or 'citizen' is ultimately dependent on "the meanings that we attach to these words[, which] are tied to our notions of morality and what we think of as being human" and can thus not be performed in a neutral way. Kelly (2007, p. 468; emphasis in original) promotes the view that evidence is itself a normative concept, since "in offering evidence for a conclusion one is *ipso facto* providing reasons to think that that proposition is true". One of the most pointed statements on the matter has been voiced by Andrew Abbott (2015, 5f.), who not only argued "that all social science is in principle at least partially normative", but also that "a purely scientific social science - a social science without values - is not only impossible, but also a logical absurdity". The normativity of social science research has furthermore been highlighted in a large

[9] For their close connection and in order not to complicate things, I refer to non-normativity and value-freedom as equivalent, despite for instance Israel's Israel (1972) urge that norms and values should be distinguished.

number of contexts in relation to organisational theory and research methods that are relevant to the properties of the capital based view of social impact.

First, there is the normative strand of research on organizational learning, "which is sometimes referred to as research on 'the learning organisation'" (Robinson, 2001, p. 58 in reference to Argyris and Schön's work 1978). Related to this, Zajac et al. (2000, p. 429) highlight that the concept of 'strategic fit' in organisations is an example of "normative models of strategy formation". Some scholars argue that even in health-care, one of the prime environments of rational action, evidence-based decision tools "contain implicit presuppositions and values" and therefore do not only enhance the decision making process, but "also transform[ed] it in a morally significant way" (Molewijk, Stiggelbout, Otten, Dupuis, & Kievit, 2008, p. 415).

Second, normativity has established its part in the methods of the social sciences with examples being the 'normative case study' promoting "theories about the ideals we should pursue and the obligations we should accept" (Thacher, 2006, p. 1631) or 'normative action research' to generate "futures theories [that] identify ends and means for […] organizational development" (Babüroglu & Ravn, 1992, p. 19).

Third, normativity is inevitably tied to one of the most dominant contemporary theories of the organisation and its management, which as remarked previously is of high relevance to the study of social impact: stakeholder theory. Donaldson and Preston (1995) have highlighted the normative dimension it contains. Jacobs and Getz (1995, p. 793) have gone even further in stating that "stakeholder theory is most powerful as a normative model" and that "it loses force and clarity when presented primarily as a descriptive or instrumental model". Similar things might be said about any research on organizational performance,

which is always marked by directionality, that is a striving to improve the situation, and thus strongly linked to normative decision making (see Dess & Davis, 1984 on firm strategies to gain competitive advantage, Thomas, Clark, & Gioia, 1993 on the influence of 'sensemaking' in directing hospital management, or Dess & Shaw, 2001 on the effects of hiring practices on the productivity of networks within firms).

The capital based approach takes a middle ground between the two poles of full separation or full integration of research and normativity, with a tendency towards embracing the latter view. It denies that normative and positivist thinking should in general be analytically separated, as found for example in Meppem and Gill's (1998) discussion of research on sustainability. The approach is instead in line with scholars who argue the philosophy of science has to find a balance between radical empiricism on the one side and discursive practice on the other that may culminate in "radical anti-epistemological critique" (Ginev, 1999, p. 143). My work promotes rigorous scientific inquiry as a basis for judging the social impact created, but also the incorporation of public reasoning and organisational learning based on what ought to be. In view of social impact, the one doesn't work well without the other: evidence must give direction on what shall be achieved and how it could be done, just as it must test whether what should be achieved has actually been met. The one is inseparably tied to and can only evolve in combination with the other.

Both the composition of capitals as well as the functions of the capitals for social productivity will thus depend not only on future research, but also on socially constructed codes of valuation (Kjellberg et al., 2013). In this sense standards for the assessment of social impact are no more or less independent of social context than "[…] accounting systems, [which] as modes of organizational control, can arise out of the interplay of political processes both within the organization and at its interface with dominant external agents"

(Burchell et al., 1980, p. 22). Social impact criteria just as "accountings can emerge out of decisions rather than necessarily having to precede them" (also Burchell et al., 1980, p. 22). The evolution of standards will need to respect principles posed in view of science-policy interfaces more generally, including "debate about assumptions, choices and uncertainties, and about the limits of scientific knowledge" (van den Hove, 2007, p. 817).

As a first step in this process of sense-making and as later advanced in my empirical research, the following sub-chapter shows in more detail how capitals have been discussed previously and in which ways they matter within the capital based approach.

3.5 The different forms of capital

Various capitals, alongside economic capital, have found recognition in the literature. One of the first was 'human capital', which has been used to denote the abilities and skills that people acquire through education. After previous episodic use by Adam Smith or Arthur Pigou (see Goldin 2016 for a discussion) Becker (1964) was arguably the one who popularised the term and promoted the significance of human capital in the economic production process. He proposed the value of 'human capital' could be expressed as the discounted sum of future earnings in parallel to other production factors.

Bourdieu (1986) took a critical stance towards Becker's 'economised' and functionalist image of education and turned away from formal learning to the transmission of education in the family context. He was interested in how the formation of 'cultural capital' through this process determined the academic and material success of people from different classes. In his essay he furthermore referred to 'social capital' as the network of social connections held by people (see also Bourdieu, 1977). The concept was also prominently discussed by Coleman (1988) with a view on the interaction between human and social capital. 'Political

capital' was sometimes used by Bourdieu too in relation to his more well-known term 'symbolic capital' as a reputational dimension of politicians' standing, for instance within a party (Swartz, 2003). The use of the term in the latter tradition remains its main application (see for example French, 2011). It has, however, also been applied in relation to civil society and democratic participation (Bang & Sørensen, 1999; Bang & Sørensen, 2003; Li & Marsh, 2008). It may thus refer to politicians and citizens alike, and as I will argue to processes of engagement and influence outside the realm of politics, too.

'Natural capital', although used explicitly some time ago by E. F. Schumacher (1973), has mainly remained a generic term referring to natural ecosystems and stocks of natural resources. It has more recently experienced an upsurge in use and been applied in ecological economics (Hawken et al., 1999). Indicators relating to natural capital and ecology are furthermore found in assessments of the ecological and sustainability performance in firms (Huppes & Ishikawa, 2005; Möller & Schaltegger, 2005).

When we think about the activities of social purpose organisations (see for instance Salamon & Anheier, 1992 for a classification) all the above capitals would be of relevance. However, from here on I will focus on social, cultural and political capital. I do so for several reasons. Some of them relate to the properties of these capitals and another to the current practice in impact measurement, which I present each in turn.

First of all, the study of human capital and that of natural capital are more advanced than that of the other capital forms because of their centrality to economic production. Apart from that, natural capital simply exists and does not have to be created. Its production thus mostly equals its preservation and therefore differs from the other capitals. What is more, there are fairly objective measures available such as CO_2 reductions (see the literature on

emissions trading Rubin, 1995), or the size and scope of natural stocks etc. (Costanza et al., 1997). The same applies to human capital in that it is assessable by standard tests of skills or can be approximated by years and level of schooling (Hansson, 2008).

The other reasons for focusing on social, cultural and political capital lie in the current practice of social impact measurement explored in the systematic review of SROI mentioned before (Krlev et al., 2013). In the research we looked at the types of indicators used in SROI studies and categorised them. Table 3 provides an overview of the capital based and other measures we found. It gives examples of indicators subsumed under each category and shows the total number of appearances of the individual indicators across the 114 analysed SROI studies (multiple observations possible). The table clearly outlines that measures relating to social, cultural and political capital are observed much less frequently than measures of economic capital or closely related human capital (monetary benefits and skills), or measures of the quality of life (satisfaction).

Table 3 Indicators used in SROI studies

	Economic capital		**Human capital**		**Quality of life**	
#		*269*		*57*		*174*
Indicators	Jobs & income for individuals	88	Skills	57	Health (phys.)	58
	Reduction of public expenditures	69			Health (mental)	31
	Financial gains for organisations	65			Well-being	34
	Taxes & econ. stimulation	47			Ease at work	51

	Cultural capital		**Social capital**		**Political capital**	
#		*91*		*76*		*7*
Indicators	Self-esteem	46	Social contacts	45	Aware. raising	6
	Reduct. viol./discr.	30	Communities	31	Participation	1
	Independence	15				

Own illustration based on (Krlev et al., 2013).

Economic capital refers to financial benefits for the individual, the state and involved organisations. Indicators relating to it are observed 269 times, skills (human capital) another 57. The 174 quality of life measures relate mostly to health, but also to well-being more generally and ease at work more specifically.[10] Taken together benefits and satisfaction measures make up for 74 percent of applied indicators (N=500), whereas the capital based measures only account for 26 percent in total (N=174).[11] Thereby cultural capital appeared most frequently and referred to self-esteem, reduced violence or discrimination, and independence (91 times); social capital to individual contacts and measures of community (76 times); and political capital awareness raising and political participation (7 times).

The discrepancy between the usages of indicators is highlighted by two further observations. First, the occurrence of social, cultural and political capital was so low despite being highly relevant to the studied interventions. While they accounted only for 26 percent of total indicator observations, they would have been relevant to at least 45 percent of the analysed studies, which were located for example in the fields of life coaching, education, social integration, or housing. Second, the standard of indicators relating to the three capitals has been rated as particularly poor. Variables were often tested by means of single questions or episodic statements instead of a coordinated set of questions and scales, or a systematic analysis of qualitative data. This corroborates on a broad scale the lack of social, cultural and political aspects in SROI observed by others (for example Kehl, Then, & Münscher, 2012).

[10] Ease at work is comparatively high, because a considerable share of studies referred to work integration.

[11] Percentages indicated in relation to a total N of 674 observations. There has been an additional category on 'natural capital', which included 29 mentions in the areas of 'respect for nature' or 'environmental preservation'. It has not been displayed and included in the calculations here, since only 6 percent of all studies related to the environment at all, and if so they mostly referred to interventions promoting the 'appreciation of nature' or 'spending time outdoors' and not to ecological interventions targeting the reduction of waste or emissions for example.

The presentation of the practice of social impact measurement in combination with the observations on the properties of natural and human capital clearly shows that social, cultural and political capital are in particular need of development. Each of these capitals is conceptualised in turn in the following sub-chapters.

3.5.1 Social capital

In the literature we typically find two different outfits of social capital (Grootaert, Narayan, Nyhan Jones, & Woolcock, 2004): First, in reference to authors like Lin (2001) or Burt (2000) social capital embodies access to a diversity of capacities (be they information or support) acquired by virtue of relationships to other people. This rationalist interpretation of social capital (see also Becker, 1974; Coleman, 1988) implies that people strategically choose their relationships with the aim of increasing their acting capacity (Grootaert et al., 2004, p. 3), that is, "[…] construct their network with an eye to its returns" (Flap & Völker, 2004, p. 68). Second, in reference to Putnam (2000; 2002) social capital mainly refers to interactions between people, be they informal or formalised in associational life, with intended as well as unintended consequences for social productivity.

With regard to the activities of social purpose organisations, both versions seem to be of relevance. In relation to cooperatives, independent of the field in which people come together to achieve things in a community, the capacity-oriented view of social capital would be more relevant. Members of the cooperative would choose to enter the association in the hope of receiving support and help from others, while investing some of their own time in giving support (see van der Gaag & Snijders, 2004 for different types of support). A youth centre in contrast would relate more strongly to the second version, since frequent encounters with others and a vibrant community are more important than individual mutual support. The two versions are obviously not mutually exclusive, since good social relations

can help achieve things together and entering a venture jointly can build and fortify personal relations. Neither will an intervention affect social capital in an isolated fashion. Follow-up and related effects are likely, such as greater trust or more engagement of the individual inhabitants for the community.

Some scholars would see the latter as parts of social capital, since their definition of social capital is very broad. In the context of development studies Grootaert et al. (2004, p. 5) for instance list the following components: trust, group composition and group expectations, information flows, inclusiveness, empowerment and political control, and even well-being. And the list continues: various normative expectations, volunteerism, safety, and even the quality of government can be found (Narayan & Cassidy, 2001; Onyx & Bullen, 2000). Other scholars argue—supposedly in relation to the latter understanding—that the meaning of social capital has been overstretched "[…] in being social networks, but also norms, trust, and authority" (Flap & Völker, 2004, p. 68). Some go as far as saying the concept has become meaningless by being all and nothing (Fine, 2010).

Even Putnam (2004b) suggests that close correlations between individual components, for example between social capital and trust, might have been misinterpreted as signs for both belonging to the same construct. Putnam goes on to explain why he thinks this confusion might have arisen by referring to two different versions of trust. In one version trust refers to whether or not someone assesses the persons they interact with as reliable or 'trustworthy'. In another version trust refers to how 'good' or 'bad' a person believes their wider surroundings to be. This can relate to people in general, but also to institutions, such as the government. The main difference between the two versions of trust is that the first one results from immediate experience and direct interaction with others, while the second one is a more general perception influenced for instance by a person's upbringing or their social

status. While the ‘trustworthy version’ has a relational character, the ‘unwarranted version’ is more of a norm. Putnam’s take is that if any of the two, the first one would be a component of social capital, while the second is only closely related.

In reference to the other capitals to follow it seems advisable at any rate to focus on the relational aspects in social capital and hold some connected variables separate. After cultural capital and political capital have been discussed a summative sub-chapter will outline more clearly how the three capitals related to each other.

3.5.2 Cultural capital

Research on ‘cultural capital’, in some way or another, always refers to Bourdieu’s (1973; 1986) initial proposition of the concept of an educational equipment of knowledge and habits that helps people reach or preserve a certain status in life. However, we find some variations on this general theme in the literature.

Skills or ability have been held separate from knowledge about or competence in cultural activities by some scholars. DiMaggio (1982) for instance argued that formalised skills were part of ‘human capital’ (Becker, 1964), while cultural capital differed from the latter not only by being more informal, but also by relating to a specific kind of education, namely that of acquiring cultural habits. Cultural competence is enhanced through “high-brow” social activities (Lareau & Weininger, 2005) such as literary and musical education, and attending arts exhibitions or the theatre (DiMaggio, 1982; Khawaja & Mowafi, 2006; Sullivan, 2001 for a more recent application). Proponents of this view usually refer to Bourdieu’s (1973) work on “Cultural reproduction and social reproduction”.

Other authors see formal skills and cultural habits as more intertwined, and the latter as not restricted to "high-brow" activities only (for instance Lareau & Weininger, 2005; Lund-Chaix & Gelles, 2013). They mostly refer to Bourdieu's (1986) article "The forms of capital", which does not make explicit reference to high culture activities (Lareau & Weininger, 2005). Proponents of this view include a range of different "tastes, habits and styles" that define cultural capital (Small, Harding, & Lamont, 2010, p. 18).

In addition to debates about the components of cultural capital, there are also wider and narrower definitions of the supposed functions of cultural capital. While Bourdieu himself refers to it as a resource of power, a social selection mechanism and an indicator of status and class (wide definition), others focus on its function in propelling social exclusion and follow-up effects such as poverty (narrow definition; see for instance Lamont & Lareau, 1988 on both).

Independent of the different interpretations discussed in the literature, cultural capital is about codes, composures (and potentially skills) that enable some people to attain a higher status than others. Status is often assessed through their educational attainment (Sullivan, 2001). Thus, no matter what the content and functions comprise exactly, cultural capital is typically concerned with the exertion of power of some people over others and the ability of those with high levels of cultural capital to navigate life more effectively and attain more favourable social positions.

When looked at from the standpoint of social productivity rather from that of inequalities, as demanded by the capital based view, one particular issue seems missing in the present view of cultural capital: a focus on the more basic modifiers and determinants of culture embodied in norms and values, and codes of social conduct. It is certainly true that the high

culture or skills variables discussed above will be important determinants of individuals' position in life and the composition of societies. The viability of a society and its functioning, however, is moderated by more basic norm based factors too, such as pro-social behaviour, a feeling of self-worth, respect or tolerance. Only when a society can ensure that people identify with what is considered a shared mode of conduct, are conflicts or exclusion avoided.

When we think about the interventions of social purpose organisations, the prevalent version of 'cultural capital' is relevant as a measure of impact in some cases. Mentoring programmes for 'disadvantaged' students for instance (Lund-Chaix & Gelles, 2013) often do not only promote codified educational skills that would directly enhance school performance. Such programmes may also seek to enable the disadvantaged students to make new cultural experiences, acquire corresponding cultural manners or make themselves familiar with particular modes of cultural conduct, which can help them advance in life. Another goal is however more likely at the centre of such programmes: The purpose of mentoring usually is to make mentees discover and become aware of their abilities (both, skills and competencies), increase their self-esteem, and develop a sense of devotion and agency. In relation to the views held towards fellow citizens, important aspects would include: recognising the worth of others and exhibiting a positive attitude towards inter-personal interaction. Without promoting these components in the formation and transmission of culture, enhancing the other components seems of little use, if not impossible.

For the reasons just laid out, I apply a complementary 'extended' version of cultural capital in my research. I do so in order to grasp and systemise the cultural factors that enable individuals to function productively within a group or wider social setting, and more explicitly how social purpose organisations contribute to the corresponding attitudes and behaviour

needed for it. This concept of cultural capital is particularly important when examining the promotion of conflict-free behaviour, but also when analysing how inclusiveness towards minorities is fostered.

3.5.3 Political capital

There are essentially two different definitions of political capital. In the first one, the concept is placed in the tradition of political capital as used by Bourdieu (1986), with a distinct emphasis the exertion of power through political authority or 'symbolic' status (see Swartz, 2003). Nee and Opper (2010) for instance related to the first version when studying Chinese capitalism. They look at the mutual obligations embedded within political institutional structures, which help politicians, other individuals, groups or organisations to promote their goals.

Also drawing on the first version of political capital, Lin et al. (2014) investigate how Chinese firms improve their political standing and market position. They measure political capital by the political experience or party membership of firms' shareholders or top management. Philanthropic engagement of a firm and the resulting appreciation of such activity by those in political power are used as another measure of political capital. It is evident that whether or not philanthropic engagement can be interpreted as a means of exerting political influence will be very context dependent and that party membership does not equal effective political control. But it is becoming apparent that the term 'political capital' when applied in this context refers to an instrumental acquisition of assets to improve one's power position. Consequently this version of 'political capital' has quite a narrow focus.

The second version of political capital is preoccupied with the democratic functioning of society. It refers to the discursive exchange between a large variety of constituents. It can

be promoted via formalised engagement within parties and in official democratic posts (Kauppi, 2003), but also by informal means of enabling citizens to have an influence (Mair, Battilana, & Cardenas, 2012). Political capital when analysed with a view on the functioning of political systems, does not only turn to the creation of effective political support for (few) constituents, through financial power or political brokering (see Nee & Opper, 2010, p. 2116). Instead it refers to the recognition of political rights and claims that enable a variety of actors, including minorities, to make themselves heard and participate in political processes (Almond & Verba, 1989; Putnam et al., 1993; Sørensen & Torfing, 2003).

Therefore, a high level of political capital, locally or nationally, cannot be assessed by the dominance of certain political actors, but rather by the breadth (diversity), depth (profoundness of issues) and frequency (episodes of and continual involvement in political dialogue) of political participation and actors' ability to assert their positions, across an entire population. It can be expressed by degrees of civic political involvement or levels of political transparency for instance (Huntington & Nelson, 1984; Verba et al., 1978). Political capital can also point to engagement in political debate, or participation in policy formation (Sabatier & Jenkins-Smith, 1993; Weible, 2006).

With regard to the capital based approach and the activities performed by social purpose organisations, it seems advisable to relate to the core elements of engagement and influence by individuals or groups of individuals found in the second version. It also seems advisable, when necessary, to detach the concept from the sphere of 'politics', which otherwise restricts the application of political capital unduly. Rather than referring only to engagement and influence in the political process, political capital can be used to refer to these aspects in basically any kind of process, since the general principle of being active

remains unaltered. Political capital could thus relate to the joint engagement of different actors in regional development. It could also relate to advocacy or discourse based agenda setting in specific thematic areas, for instance in consumer protection; or to pro-active problem solving and the exertion of influence within other contexts. I shall henceforth refer to this broader version of political capital, whenever I use the term while not denying its application within the realm of politics.

3.5.4 Relating the capitals

My use of the terms social, cultural and political capital as described above strongly resonates with previous research on the concepts across disciplines. I have, however, also introduced some specific notions or new interpretations against the background of the function of capitals within the capital based approach and with regard to the activities of social purpose organisations. In contrast to Bourdieu for example my interest lies in investigating how capital forms enable collective productivity, not in how they might be used to exert power or improve one's social position at the expense of others. For this very reason I had to extend the meanings of the capitals in some ways. I have for instance shifted the focus in 'cultural capital' from a specific emphasis on cultural skills and competencies that influence social status to include the most basic modifiers of cultural conduct (extended definition). I have also extended the meaning of political capital in that I have detached it from the realm of 'politics' to comprise aspects of participation, engagement and influence in other settings too (including that of politics of course). In contrast to this, I have narrowed down the meaning of social capital to its relational components in order to enable an analytical separation between it and the other two capitals.

This results in a conceptual framing of the three capitals that is useful within the system of social productivity developed and with regards to the activities of social purpose organisa-

tions that promote different capitals in their target groups. All three capitals enhance social productivity but each is different in how it relates to the individual and the social setting she is placed in:

- Social capital: Relations, connections and exchange between individuals
  (comprises: familiarity, intimacy, interaction, mutual support etc.)
- Cultural capital: Positioning of individuals towards others
  (comprises: self-image, image of and appreciation for others, values, norms etc.)
- Political capital: Engagement and activity of individuals within social settings
  (comprises: activism, involvement, influence etc.)

In connection to the preceding discussion we see that trust and other 'norm-related' constructs, such as tolerance, open-mindedness or integrity should be held separate from social capital, which is relation focussed, and integrated into cultural capital instead. Political capital in turn forms a separate category in that it is neither about the connections between people nor their positioning towards each other but the intentional activity of individuals towards a specific goal.

I should remark that this conception has informed the empirical research to follow, but it has also been shaped by the latter in an iterative process. Some of the reasoning and justification for the conceptualisation given here will be highlighted by insights gained during the empirical research and reported in the empirical results. The conceptual drafting and the empirical insights will be critically acclaimed in the conclusions. Before we move on to the methodology that guided the empirical research, it needs to be reflected which advantages the capital based approach might offer.

3.6 Advantages of the 'Capital Based Approach'

The promotion of capitals by social purpose organisations, and thus their social impact, can be assessed by operationalising the portrayed capital forms and accounting for shifts in the levels of capitals held by target groups. The capitals approach hence does not deviate from the practice of the other approaches to impact measurement in targeting individuals in the measurement process. However, the character of the variables assessed differs. Capital based measures are not grounded in individual utilities or functionings and then aggregated. As the conceptualisation showed capitals are held *between* individuals, exhibited by individuals *towards* others, or employed by individuals *within* social settings. Capitals are thus individually measured but collectively grounded and might thereby help to ease some of the complications referred to as the 'foundation problem' in the review of the literature.

The capital based approach also promises progress on the other three challenges: First, it eases the measurement problem by drawing on previous research from a variety of social sciences. Some of the available resources have already been cited, but more specific uses and previous empirical applications will be presented in the methodology chapter to follow as well as in the three empirical studies. Second, it decreases the attribution problem by defining a unit of measurement (capitals) that is more closely located to the organisational activities, which have advanced these capitals, than to the final outcomes at the wider level of society. Third, the intermediary character of capitals does not only increase their tractability, but also makes the approach more fit for studying the process of creating impact and the elements involved. This provides deeper insights to the 'how' and 'why' of social impact. It thereby helps address the aggregation problem.

The capital based approach represents an alternative way of framing social impact that comes with the further advantage that several dimensions of impact (norms, social con-

tacts, engagement, skills etc.) are connected by a common terminology and a joint conceptual thread. The following chapter on methodology will not only show how the approach offers points of reference for prioritising a certain capital form in the study of particular organisational activities, but also how it enables the study of interrelations, reinforcements or trade-offs between multiple capitals. The methodology chapter will furthermore outline how the conceptual approach can be operationalised for empirical application, including considerations on how it can be combined with various research designs.

4 Methodological framework

The methodology chapter serves two purposes: First, it presents how the theoretical model developed in the previous chapters can be translated into empirical research. It serves as a general methodological commentary for the empirical work performed and presented in the subsequent chapters. It outlines how to link up to previous research to identify specific measures for the capital forms in focus. It also discusses how to revise existing measures and develop new ones, if necessary. The chapter outlines the fundamental principles of the research and leaves the exact details for the individual empirical chapters.

Second, the chapter develops an analytical framework that can be used to structure the creation of different forms of capital by an organisation. The framework is universally applicable across different types of organisations, interventions, and fields of activity. It includes a procedural description, which leads analysts of social impact from the mission of the organisation under study to one or several forms of capital as measures of its impact.

Taken together this provides a rounded impression of the empirical research to follow and the strategies that have been applied in executing it.

4.1 Measuring capitals – breadth and focus

Many social purpose activities will target several types of capital and all capitals are potentially relevant to the analysis of these activities. To test the capital based approach we need to study the occurrence and interplay of capitals within different settings. In other words, for a comprehensive testing, we need a ‘multi-sites and multi-capitals’ empirical setup, which as will be seen below has in fact been realised. While enabling breadth, the capital based approach also provides focus.

The very logic of applying capital formation as a measure of social impact, helps calibrating the research focus on the type(s) of capital that are most substantial to the activity in question, since the aim is to move the assessment right to the point where it is most central to decision making. To put it differently: the capital based approach helps to make sure that the supposed effects which are closest to the core of an organisation are effectively identified and tested in the measurement. The focus derives from a match between the different capital forms and the intervention's main goals as formulated in a mission statement for instance, which is especially vital to social purpose organisations (McDonald, 2007). It should mostly be possible to focus on one particular capital in studying the social impact of an intervention, while considering some of the other capitals or variables from other approaches alongside, such as monetary benefits, satisfaction or capabilities.

For some organisations and interventions it will be harder to focus the analysis than for others. Some might necessitate the testing of a very broad set of capitals and other variables. This applies in particular to projects that are composed by a variety of different intervention strategies such as the 'millennium development villages' (Clemens & Demombynes, 2011) or organisations like Bangladesh's BRAC (2017), whose widely postulated aim is to "empower people and communities in situations of poverty, illiteracy, disease and social injustice." But even in such cases, for the rigorousness of measurement, we can differentiate by programme components to relate the individual intervention to the respective types of capital it fosters.

When following this logic, education efforts would *mainly* be assessed by the human capital concept, multi-stakeholder brokerage by the social capital created, gender equality initiatives by changes in the cultural capital of women and men, and advocacy for the participation of minority groups in decision making by the formation of political capital. A cur-

rent example of the advantages of using capitals to bring focus to the analysis, is Mair, Battilana and Cardenas' (2012) typology of social enterprises by the capitals they highlight in their mission. As stressed above, this does by no way mean that other capitals shouldn't occur in the assessment of impact, but only that other than the primary capital should be treated as complementary to the latter and relatively more inconsequential decision criteria in judging the impact achieved.

It needs to be remarked that studying capitals promoted by an intervention might come with surprises and unexpected insights on the measures relevant for grasping the main impacts effectively. All of the empirical studies performed have started off with a focus on one capital as derived by the (supposed) main goals of the organisation. The studies have considered other capitals and further variables in a complementary fashion though and have generally been held open to the inclusion of new aspects along the analysis. This was vital, since some effects that had been allocated to one specific capital in the beginning proved to be incompatible with it in the analysis and therefore had to be re-allocated to others. I also encountered that variables and capitals emerged out of the analysis as being of high importance, although they had initially only been at the periphery of my attention. It is therefore important to strike a balance between the breadth and focus of the analysis. This will be discussed in depth in the discussion chapter following the presentation of the empirical results.

4.2 Research design versus the conceptual approach

Before moving on to the portraits of the three empirical studies realised within the research, I want to clarify the difference and connections between the conceptual approach to social impact applied in the empirical studies (here: the capital based approach) and how it relates to research designs. I do so, since the relation between the two is not always clear

and weaknesses of a study emerging from the design are sometimes confused with those arising from the conceptual approach and vice versa.

To control for inaccuracies in the measurement in basically any assessment of intervention effects, scholars are trying to control for confounding factors. Thus, in order to effectively link performed activities and effectuated changes, social science scholars mimic the experimental research designs of the natural sciences. 'Randomized controlled trials' (RCTs) are regarded as the gold standard, followed by 'quasi-experimental' or 'observational' designs and then by case studies or other qualitative methods (see Glasziou & Heneghan, 2009 for an overview). Observational studies may be 'cross-sectional', that is based on a single 'one-off' observation, or they may come in a 'cohort' design which rests on several observations over time. Cohort studies are thus longitudinal and effectively a time series of cross-sectional observations. Observational studies just as RCTs usually work with an intervention and a control group to make sure that the impacts measured (here: the increase in capitals) is not due to some unobserved factor affecting all people in a similar way. Because groups are not randomly allocated in observational studies, one group might differ in some properties from the other right from the beginning and these differences might have distorting effects on the measured impacts. Therefore 'matching methods' are often applied in the analysis to statistically account for and level out differences between the groups (Imbens & Wooldridge, 2009).

In principle all of these study designs can be applied with each of the conceptual approaches discussed. And no matter what the conceptual approach, the more sensitive the subject, or the more vulnerable the target group, the more restricted is the ability to apply rigorous experimental designs. Byford and Sefton (2003) for instance discuss the practical and ethical limits to randomization in the field of social care for reasons of safety. They point out it

would be unethical to return to or leave a young person in a family, where the youngster would be subject to serious neglect. What also needs to be taken into account with regard to social interventions, is that qualitative designs might be superior to quantitative designs for the richness of data the former provide (Byford & Sefton, 2003). What is more, the complexity and context dependence of social interventions, as compared to the testing of new drugs for instance, may generally inhibit the reliability and usefulness of applying experimental designs (Deaton & Cartwright, 2016). It could be said that the more complex and multi-facetted an intervention, the more useful is it to combine designs and gather both qualitative and quantitative data to make an informed judgement of the impact caused.

As regards the various challenges emerging when studying impact within the system of social productivity that I sketched out, study designs are restricted in their capacity to make a difference. The study design can only moderate the attribution problem and has no influence on the measurement, aggregation or foundation problems, since it does not alter how caused effects are defined, broken down and measured. The conceptual approach in contrast, as has been seen before, can either exacerbate or mitigate all four problems (attribution, measurement, aggregation and foundation). It defines: (1) the main dimensions of interest and the relevant variables to be assessed to capture these dimensions; (2) how individual variables should be processed, for instance combined to form a separate category or an index, or how they should be transformed into another 'currency', for instance through monetisation; and (3) how the resulting measures should be interpreted, for instance held against the amount of resources invested into the intervention.

For the above reasons, one needs to be careful when providing a critique of a study. Relating back to the systematic review of SROI, illustrates this well: The fact that there have been so few ex ante-ex post observations (which could be performed with any design) has

nothing to do with the fact that SROI is a cost-benefit approach, but simply with the effort of realising rigorous designs. In contrast, the fact that many 'social effects' have been neglected in the analyses is due to the application of a cost-benefit logic and the resulting focus on monetisation, and has little to do with the exact study design applied.

We've now seen how conceptual approaches and methodological designs are linked. It has become clear that the critical lever for enhancing the knowledge and evidence produced in a social impact assessment lies in the conceptual approach. I have, however, also shown that a poor research design may undermine every conceptual approach and that an effective study would seek a methodology that best fits the targeted intervention. In the following I show how the capital based approach has been combined with three different designs to cater to the particularities of my three empirical cases. I outline the general design of each study along with some of the individual capital components and other variables applied therein. I also discuss some points of reference that have been used to screen for measures applicable to the capitals of interest. However, the specification of the design as well as of the exact composition of measures is performed in the separate empirical chapters.

4.3 The studied interventions and points of reference for capital measures

The purpose of the empirical work was to test the capital based approach empirically. In order to do so in a comprehensive way the research needed to be broad in scope—all three capitals should be sufficiently tested—as well as deep in the analysis of each of the capitals. At the same time the research was drafted so as to be able to study potential trade-offs or reinforcements of the capitals within the empirical case. This led to the 'multi-sites multi-capitals' research setup outlined above. One case each was identified to study social, cultural and political capital in-depth, while also regarded other capitals and some additional variables within each study as demanded by the nature of the intervention. The dif-

ferent interventions, together with these research objectives generally asked for a mixed quantitative and qualitative research design. All three cases are situated in Germany. Each of the cases is described briefly below, after the primary form of capital investigated has been listed together with the research question guiding the analysis and the intervention presenting the empirical context.

(1) *Social capital*: What is the impact of 'community oriented housing models' on the social relations of elderly people living there?

(2) *Cultural capital*: What is the impact of a programme on prosociality in schools on the constructive attitudes and behaviour of pupils?

(3) *Political capital*: What is the impact of a political web-platform on the political activity of users and the interaction between citizens and politicians?

The first study is situated in the field of elderly care. The analysis performed in the chapter is concerned with the social impact of 'community oriented living' facilities for elderly people. The facilities I studied are run separately by two private faith-based foundations. The models differ from 'assisted living' and related concepts by actively promoting contacts, activities and exchange between the inhabitants of the facility and with people outside it. Community oriented models have a more weakly pronounced service component than assisted living and instead foster individual and group activity and self-initiative among their residents. Due to these particularities, the supposed main social impact of community oriented models lies in a higher degree of social capital among their tenants than that found in assisted living. This was quantitatively assessed in a cohort study. The data was gathered in a person-assisted survey that was repeated after one year (longitudinal) to do both, verify the measured impacts and extend them in some regards. The design was furthermore matched and addressed an intervention group situated in community ori-

ented models and a control group of people living in assisted living. Alongside the other capitals several satisfaction based measures have been assessed (health and life satisfaction, and satisfaction with the living conditions).

The second study is located in a school context. It examines the social impact produced by a programme which means to revert anti-social and strengthen pro-social attitudes and behaviours in pupils. The programme is therefore investigated as to its effects on the cultural capital of the participating pupils. It is called 'RespAct' and performed by a social enterprise active in the field of youth work. It is provided in a condensed 2-day format in schools. Since the analysis seeks to trace change processes, it is mainly qualitative and uses in-depth interviews with pupils to find out not only if there had been an increase in cultural capital through the intervention, but also how it has come about. Some quantitative measures from a paper-based survey are assessed as complimentary evidence to cross-check effects. The views of teachers have been assessed to complement insights gained by interviewing and surveying pupils. In addition to the other capitals, skills acquisition (human capital) and satisfaction with the programme have been considered.

The third study deals with democratic functioning and politics. It measures the social impact of a political web-platform, which attempts to increase political transparency, spur direct and open communication between politicians and citizens, and sensitise the public for lobbyism. The platform investigated is 'www.abgeordnetenwatch.de' (English equivalent: www.votewatch.eu). In resonance with its goals, the platform's main social impact is expected to lie in the creation of political capital. This is assessed by an online survey among users of the platform, both citizens and politicians (cross-sectional design). In the absence of a proper control group, the effects of the platform on political and other capitals have been analysed by means of comparison to the levels of similar variables in the general

German population as found in responses to the European Social Survey (ESS). Data from ESS have also been used to find out in how far the users of abgeordnetenwatch differed in their characteristics from the general population to put the results into perspective. Variables other than those relating to the three capitals have not been assessed because of the restricted scope allowed for by the survey.

As regards the operationalisation of the capitals, and thus the drafting of the questionnaires and interview guides, I have drawn on one of the key assets of the capital based approach: its grounding in a variety of social sciences. The following paragraphs provide some references to how the three capitals have been recognised in previous research. Please note that here is not the place for a nuanced description of all capitals, since each conceptualisation will be performed in detail within the methods section of each individual study. I therefore only outline the potential links to previous research I had identified when drafting the studies and proposing them to the targeted organisations.

It will become apparent that previous applications have often been limited to the state of capitals at the macro level of society or the micro level of individuals. They have mostly not occurred in relation to their active creation by an organisational intervention. Please also note that I have only been looking at studies, which had analysed the levels of or shifts in the capitals in specific groups or society at large. This excludes a large body of literature on the occurrence or instrumental employment of certain capitals within organisations (see for instance Tsai & Ghoshal, 1998; Zheng, 2010 on social capital).

A considerable body of literature has developed concerning the measurement of social capital, be it in relation to regional communities in industrialised countries (Lillbacka, 2006; Onyx & Bullen, 2000; Stone, 2001) or in developing countries (Grootaert et al., 2004;

Grootaert & van Bastelaer, 2002; Narayan & Cassidy, 2001). Although empirical research is usually situated at the macro level of society or larger regional communities and employs secondary panel or cross-sectional data, most of the methods applied can be transferred to more specific groups of people and how their social capital is affected by organisational interventions.

Political capital has been discussed in relation to political participation and the recognition of minority interests (Almond & Verba, 1989; Putnam et al., 1993; Sørensen & Torfing, 2003). Insights into its measurement at the national level (Dalton, 2008; Kalaycioglu & Turan, 1981) can provide guidelines for its application in organizational social impact measurement. Another variable of political capital is assertiveness in promoting and enforcing political positions. This relates to 'advocacy coalition frameworks' (Sabatier, 1998; Sabatier & Jenkins-Smith, 1993; Weible, Sabatier, & McQueen, 2009) that are used to trace actor influence in policy formation.

As stressed earlier cultural capital has been drafted in my research with a focus on basic cultural attitudes and behaviours grounded in societal values and oriented at normative standards. My view has also been focussed on how it contributes to social productivity and not to "class reproduction" (see Goldthorpe, 2007 in relation to Bourdieu's work for the latter). In the former outfit it is closely connected to individual "identity" and its formation, which is intensively debated in psychology (Davis & Marin, 2009; Wernet, Elman, & Pendleton, 2005). The subject is somewhat delicate due to its vagueness (Mruk, 2006, 37ff.). Nonetheless, it has been discussed as an operational variable for research (Abdelal, Herrera, Johnston, & McDermott, 2006). For operationalising the cultural capital concept, the methods and findings of values surveys (Esmer & Pettersson, 2007) are very useful.

Despite a vast repertoire of previous research that can be drawn from to fill the capital based approach with life in its empirical application, none of the identified resources are ready-made and can simply be applied in the measurement of social impact. Available scales are mostly either too narrow and the number of questions to high (psychological investigations of a very specific construct) to be practical, or too broad and the number of questions too low (national surveys) to ensure encompassing measurement. What is more, available instruments usually will not match the specific setting or target group of interest. Existing instruments thus offer some guidance, but need to be adapted significantly in any case and often instruments need to be drafted anew with only weak connections to pre-existing ones. How to go about this and how to move from organisational mission and goals to a fully-fledged research design, including instruments and an analytic strategy, will be discussed in the following. The procedural description is directly related to the steps I performed in testing the capital based approach, but it is written in a way so as to refer to the steps involved in measuring social impact more generally.

4.4 Analytical procedure

The considerations made above have been integrated as vital components into an analytical procedure, which comprises guidelines to follow in performing empirical research on social impact. The procedure within the methodological framework sketched in this chapter comprises seven steps or guiding questions:

(1) Decide: Is the capital approach needed and useful?

(2) Choose: What are the main capital(s) and further variables of interest?

(3) Draft: Which research design is most appropriate?

(4) Compose: What kind of insights should be produced and what kind of data will be needed (quantitative, qualitative or mixed)?

(5) Design: Which individual variables shall be applied and how will they be measured?

(6) Calibrate: Is the measurement repertoire appropriate and practically applicable?

(7) Display: How can results best be communicated and displayed?

Each of the corresponding steps is discussed in detail in the paragraphs to follow. References are made to how the procedure can be applied in studying interventions outside the realm of those that I have selected.

4.4.1 Accounting for complexity and research interest

First of all, anyone attempting to analyse the social impact created by a specific organisation, or rather one of its interventions, has to decide whether or not the complexity of the intervention's (1) 'theory of change' or 'logic model' or 'impact chain'[12], and (2) operational strategy (see Ebrahim & Rangan, 2010) demand a capital based approach, or more generally the assessment of impact at all. Certain interventions won't require my approach or any of the others discussed, since they can be assessed by simpler means.

Ebrahim and Rangan provide a rationale for deciding on the type of variables the assessment of effects caused by some organisational activity needs to focus on. They argue unless both, the impact chain and the operational strategy, that is the ways in which an intervention is executed, are complex, organisations will not have to address social impact at all. Their effects can be captured reasonably well by simply looking at 'what the organisations have done'. In contrast to Ebrahim and Rangan I use 'activities' and not 'outputs' to denote 'what the organisations have done'. I do so in order not to confuse the impact chain outlined in relation to 'social productivity' in my theory development on one side, which

[12] All three terms are commonly used and refer to the same: The supposed causal links between an intervention and its social impact (see for example Weiss (1998); Social Impact Investment Taskforce (2014)).

refers to the processes happening within society, and the 'production chain' of individual organisations as 'capital suppliers' to that system of social productivity on the other.

Table 4 gives an overview of the possible constellations and examples of where they could be found. It highlights that only when both dimensions are complex would we need to think about impact, and the capital based approach as one way to assess it. In all others, focussing on activities would mostly be sufficient.

Table 4 Differences in the complexity of interventions

Impact chain	Operational strategy: Focussed	Operational strategy: Complex
Complex	Activities *Example:* *Drug withdrawal*	Impacts *Examples:* *Violence prevention* *Empowerment activities* *Community development*
Focussed	Activities *Example:* *Food kitchen*	Activities *Example:* *Immunisation initiative*

Own illustration in relation to (Ebrahim & Rangan, 2010).

A focussed impact chain and operational strategy are found for instance in a food kitchen. There is a clear need (hunger) and a simple intervention (providing food). To judge the effects of the latter, it would be sufficient to look at how many meals have been given out. The same logic applies to other interventions, such as disaster relief or the life-saving interventions of sea rescue services. They can be assessed by the number of shelters provided, or the number of lives saved. A focussed impact chain and a complex operational strategy are also fairly easily identified. Think about an immunisation campaign. It will be

more complex than in a food kitchen to reach out, inform and attract people (operational strategy), but the impact of the intervention is quite focussed in that performing the vaccination directly provides protection from the disease and thus a health gain. Water hygiene initiatives would be another example in this dimension. They could be assessed by the number of chlorine tablets distributed, but people need to be convinced to use them once they have them.

It is harder to find a good example for a focussed operational strategy in combination with a complex impact chain. I disagree with Ebrahim and Rangan that interventions which promote "rights and freedom" or "good governance" belong into that category. In my view those are complex at both levels and should thus be assessed with an interest in impacts and not merely in activities. The best example I could think of are drug withdrawal programmes. Although the intervention needs to be multi-facetted and it can be hard to achieve effective behavioural change in patients, many such programmes work with a rather standard programme of how to promote abstinence and thus pursue a focussed operational strategy. However, there are alternatives to these standard models that have only been recognised recently, which focus on 'controlled consumption' of illegal drugs instead and challenge the status quo in many ways (Bergmann, 2014). They modify both the operational strategy and the impact chain and thereby move such interventions to a double complex state.

There are also other 'in between cases' with a tendency of being more complex in both regards: HIV/AIDS campaigns promoting safer sex for instance share more commonalities in terms of operational strategies with the immunisation or sanitation campaigns, but will need to effectuate impact through a more complex process of change in the individuals targeted and their surrounding environment.

Other interventions are inarguably complex on both levels. A violence prevention programme, independent of whether it targets the family or some wider social group, will have to seek out and intrude the social setting at risk and try to effectuate change of a behaviour that might have deep root causes. The empowerment of women in deprived or traditional societies will need to span many different levels, from self-determination in various regards to the ability and acceptance of women to participate in public life. In a similar way, yet with a different target, community development initiatives will need to become active at different levels and employ diverse modes of intervention.

The difficulty of finding good examples for interventions that are focussed in at least one dimension outlines that the question about social impact is justified in relation to most social purpose organisations and the work they do. Please note that there may be assessments whose primary interest is not in impact, but in other aspects such as cost-effectiveness, the professionalism with which the intervention has been executed or other dimensions of performance. Analysts of social impact will have to decide whether the capital based approach must be at the centre of the investigation or whether it serves a complementary function and by which other measurement components it should be accompanied—for instance satisfaction measures, cost accounts or potential interaction effects such as negative externalities. Once this decision has been made, the following procedural steps help draft and perform the corresponding measurement.

4.4.2 Defining scope and measures

Second, the impact analyst has to identify which capital(s) are central to the investigated intervention. In this it is important to consider a range of stakeholders and the different effects they might benefit from, as highlighted in basically every guide to social impact measurement (Nicholls, Lawlor, Neitzert, & Goodspeed, 2009; Schober & Then, 2015;

Social Impact Investment Taskforce, 2014). It is equally important though, not to let this effort distract the focus of the analysis as proposed above. Finding focus should be guided by the two rationales of: (1) delimiting the range of the stakeholders under study and (2) delimiting the range of (follow-up) up effects considered. With regard to the first aspect direct beneficiaries should be at the centre of attention. This derives from the main conceptual argument of the capital based approach, namely that the analysis of impact should be located at the point that is most meaningful for answering whether and which effects have been achieved. If an intervention is not effective in making a difference for their target group, its overall impact is seriously impeded.

With regard to the second aspect, the measurement components have to be tied to the core goals articulated in the intervention's mission. This is a claim I have raised above, but it merits more detailed explanation: Since social impact is a way of measuring the performance of social purpose organisations in contributing to social productivity, any measurement pertaining to it needs to refer back to the place where the supposed main dimensions of performance are defined. The danger otherwise is that the variables to be measured are manifold and at such a distance that they exponentiate the identified assessment problems.

I have illustrated previously that the capital based approach is capable of covering a large share of social purpose activity and that it can account for primary and secondary effects, including the occurrence of multiple capitals within one intervention. Although an intervention's potential effects might be broad and entail a series of follow-up effects—all contributing to final social outcomes— the decision of whether and how an organisation creates impact can be focussed on the critical enablers within the system of social productivity, the capitals defined in the last chapter. Most follow-up effects are varied by the degree to which capitals have been enhanced.

It is to be remarked that delimiting the range of subjects and focussing on organisational goals for deriving the exact capital measures are in line with each other, since direct beneficiaries will usually be highlighted as the core target group in any organisation's mission. This can be illustrated by drawing on the mission statements of the investigated organisations. Abgeordnetenwatch (2017) for instance presents itself as a "[…] web-platform that enables citizen requests to politicians" and furthermore "documents the voting behaviour of politicians and their income earned outside their political mandate." It does so in order to "[…] increase participation and transparency in politics."[13] In addition to its goals (participation and transparency), it defines citizens and politicians as its target groups. The RespAct programme (CamP Group, 2015) in turn is meant to "empower children" (target group), more specifically "[…] to strengthen their self-esteem on the physical, the social, and the emotional level" and to teach them " […] the skill of mastering even difficult situations without relying on violence within the class-community."[14] Given such statements, the study of RespAct should focus mainly on the pupils affected and neither their school as a whole nor their families. There might be subsequent effects on those stakeholders, but we first need to know with certainty how the intervention affects the pupils participating in it.

Once the core stakeholders and capital components likely promoted by an organisation have been identified, we can think about how changes in capitals within one group or differences in capitals between groups can be assessed reliably.

4.4.3 Choosing the appropriate research design

In a third step, the research design needs to be drafted. As already mentioned experimental designs may clash with feasibility or ethical considerations, which is why impact analysts

[13] Translated by the author.
[14] Available in English.

may have to rely on observational research designs. These exist in a variety of outfits, which need to be selected based on how effectively they are applicable to the specific context. Controlled designs, that is such that compare capitals between the intervention group and a control group, require several specifics from the research context:

First, there needs to be a meaningful control group, that is a group of people as similar as possible to those participating in the intervention, except for the very participation. The control group may and should be subject to another intervention if two alternative interventions are to be compared as to their impact (for example community oriented housing and assisted living). Second, the recruitment of a control group needs to be feasible, that is there needs to be sufficient ground to assume members of a potential control group will be willing to participate in the study. With regard to Abgeordnetenwatch it was nearly inconceivable that a sufficiently high number of random citizens that did not use the platform would choose to answer my survey. This is why comparisons have been made to secondary data from ESS, although the latter did not provide a perfect fit with my questions. Third, the insights achieved through the comparison should be valuable. No controlled design was chosen in the study on RespAct for instance, since little was to be gained by asking pupils, who had not participated, about their pro-social attitudes and behaviours. In particular since qualitative methods would not have been able to arrive at an exact level of the cultural capital components pupils held in general. A reasonably meaningful estimation of impact was expected from asking those, who participated about their perceived effects and changes in their composure.[15] Fourth, it needs to be economic to realise a controlled study given the size of the organisation and the scope of the intervention under study.

[15] See Edelstein (2016) for a similar strategy if applied as a complementary element to a quantitative study only.

If one of these criteria does not hold, within group comparisons before and after the intervention with careful account of potential confounding factors are an alternative. They might generally be the design of choice where a multitude of confounding factors could be responsible for an effectuated change, for instance in fast changing environments or where interventions are hard to delineate strictly from other social influences, which are likely to vary across contexts or even individuals. This applies for instance to any normative education of children. Every child has a different position within the larger setting of the school it goes to, has an individual family background, is embedded in different peer groups or associations. The social settings and changes over time are likely more variable for children than for other target groups. Even if children or groups of children are similar at face value, they are likely more affected by confounding than other individuals and reliable inference of impact from group comparisons is harder. Where such circumstances apply, it seems furthermore advisable to use a rather short time frame between pre- and post-assessment, since time might lever confounding so much that a longer duration deteriorates the precision of the analysis more than it offers to demonstrate the sustainability of effects.

On both points, context and changes over time, educational research on bullying in school offers some illustrative insights. There typically is a significant "age decline" in reports about being bullied from the age of 8 to 16, that occurs without any specific intervention. Already some time ago Smith, Madsen and Moody (1999) have found the greatest support for two potential explanations for the observed decline: (1) when children are younger there are more peers older than them attending school, who could potentially bully them; (2) children only come to acquire the social skills and assertiveness needed to deal with bullying as they grow older. Thus, social constellations and learning within these settings matter and change is occurring although there is supposedly no directional force steering the influences. If we don't assume that the learning process occurs in a linear fashion and

in the same way for each and every individual, there might be differences caused by (changes in) group composition or (episodes of) external social influences that are hard to control for, even by quasi-experimental or experimental designs.

Finally, when pre-post observations are practically hard to realise, for instance when it would require too much time to form a sufficiently large panel for the study, an option is to make use of retrospective accounts. I have done this in the study of the community oriented living models, since the numbers of new inhabitants joining the targeted facilities was so low that it would have taken years to assemble enough observations on people who had just moved in. Instead, inhabitants have been included in the study independent of the length of time they had lived in the facility. In order to ensure that capitals in the one model were not only higher due to the different disposition of people choosing that model, participants were asked to assess their general social connectivity throughout their life. The latter served as baseline accounts for the capitals investigated (inclinations to hold higher or lower levels of the capitals) in order to contextualise effects and to compare the two different groups. As outlined hereby, combinations of any of the above elements are possible. A longitudinal controlled study might still make use of some retrospective accounts. After decisions on the design have been made, we need to think about what types of data will need to be generated and assessed.

4.4.4 Composing instruments

The fourth step is to compose instruments to generate different types of data, namely quantitative or qualitative data. Quantitative data in the context of measuring impact will mostly be generated by primary surveys or drawn from secondary data sources, while qualitative data will emerge from original interviews or observations. Quantitative methods may appear more valid and reliable, since some of the applicable measures might have been tested

extensively in previous research across a variety of contexts. They also appear to be more objective by their use of numbers. However, King, Keohane and Verba (1995, 4f.) for instance promote the argument that "[…] differences between quantitative and qualitative traditions are methodologically and substantively unimportant" and that both, if rigorously applied, are able to arrive at "valid causal inference." They make the last statement in relation to historical research. In terms of explanation, qualitative methods offer more in-depth insights into a matter and are more easily adaptable during the research process. The impact analyst can probe issues right away and flexibly narrow down or extend the range of questions during one interview and carry the adaptions made to the next. Questionnaire items in contrast can be adapted only after a larger number of questionnaires have been completed and analysed as to gaps or inconsistencies in the knowledge they generate.

Qualitative methods are particularly useful under three conditions. First, they offer to generate insights where we can say only little about the supposed effects of a specific intervention, that is where we have little experience with it or related activities. Second, they offer room for differentiation and probing where effects differ remarkably across different types of beneficiaries. In the study of RespAct for instance I differentiated between pupils with higher and lower behavioural problems. Third and most importantly, qualitative methods should be applied where we seek direct insights into change processes or dynamics in capital formation to make an informed judgement about their creation—and thus where a steady state comparison of the investigated capitals across groups does not offer enough of such evidence.

Longitudinal quantitative data of course also enables us to follow changes in capitals over time, but in contrast to qualitative data it is hardly able to uncover the underlying mechanisms or understand and describe the interactions between effects. In consequence, qualita-

tive data is decidedly better at establishing links between the range of performed activities and the increase in capitals. In other words it offers to address the additional question of how changes have come about rather than only the one if they have occurred. This very link can be pivotal to organisational learning and the (re-)design of intervention components. Maxwell (2012, 31f.) mentions causal inference is better ensured by quantitative testing. But at the same time he makes the case for qualitative methods. He does so by relating to research in educational contexts from back in the 1980s and 1990s that suggests the effects of educational research on educational practice had been restricted by the fact that most of it was quantitative and lacking practical context and appeal. In relation to the study of RespAct we will see that this circumstance might not have changed much since then.

After having identified which types of data will offer the best evidence, we enter the most important phase in the measurement of social impact, which is the selection, adaption or design of the capital and complementary measures to be applied.

4.4.5 Refining measures

The first step in the development of measures is a thorough differentiation of the individual capital components. The exact measures, just as the different capitals themselves, are derived from what is stressed in the organisational mission. Depending on that mission the capitals can take on different outfits. In particular cultural capital, in referring to a large variety of norms and values, can accommodate a variety of aspects. One of the goals of RespAct for example is increasing self-esteem. But cultural capital could also refer to tolerance towards minority groups and the embracement of diversity.

Depending on the distinct focus the drafting of individual measures has to relate to different points of reference in existing research. Standard and validated scales from psychology, sociology, economics or political science can be a useful resource, independent of whether the applied methods in the measurement of social impact are quantitative or qualitative. In case a qualitative approach is applied, questionnaire items must be translated into a (semi-)structured, and in rare cases unstructured, interview guide (Gubrium & Holstein, 2001). If the study is quantitative no such conversion needs to be made. However, it is mostly useful to shorten, adapt, revise or combine existing sets of questions to match the specific context of the analysed intervention and the focus in the knowledge to be produced. One of the reasons for performing some revision is the practicality of the measurement. It has to be feasible, in particular for smaller or less resource rich interventions. It often also has to be repeatable over time, for instance for being implemented into periodical social reporting (Nicholls, 2010a), or for replicating the analysis by increasing participant numbers, or under varying context conditions.

In some cases measures will have to be drafted anew, since no previous research exists that can be applied directly, which is the case when the impact assessment is great in depth, large in scope, particularly original, placed in a specific context or located at a level typically not addressed in previous research in an adequate way. Although there is not always a perfect fit between the goals of the measurement and previous research, it is very likely that some guidance can be found in the questions and scales applied in psychology (see for instance Mruk, 2006 on the vast array of theories and measures of self-esteem alone) or in large social surveys (for example European Social Survey, General Social Survey, World Value Survey, and the many national surveys). Whenever new questions and scales are developed or existing ones revised, it is advisable to test their reliability and validity in a pre-test, no matter whether the measurement is quantitative (Cronbach & Meehl, 1955;

DeVellis, 2016; Murphy & Davidshofer, 2005) or qualitative (Golafshani, 2003; Merriam, 1995; Morse, Barrett, Mayan, Olson, & Spiers, 2002).

4.4.6 Pre-testing and executing the study

Sixth, pre-tests also serve another, very central function, namely that of establishing a focus on beneficiaries' perspectives. The way measures are assessed has to be meaningful not only judged against organisational mission, but also based on the viewpoints of the target group. This principle is embodied for instance in "user-oriented evaluation" of social service provision (Beresford & Carr, 2012; Martin, 2011). It ensures that the performed analysis is 'material', that is relevant, to the served target group. Materiality towards target groups as well as other stakeholders, such as investors, is an imperative in social impact measurement (Social Impact Investment Taskforce, 2014; Somers, Nicholls, & Nicholls, forthcoming). While not centred on individuals conceptually but collectively grounded, the capital based approach still addresses individuals directly. Thus, impact analysts need to find out whether the derived and developed measures work well with regard to their perceived importance to the affected target group.

For developing a good understanding of the latter, the analyst has to engage with these groups. Since discussions with the individuals targeted by the research might bias the effects measured later, the researcher has to find individuals who are reasonably close to the situations the target groups find themselves in. Factors to be considered include similarity on a variety of dimensions: institutional settings, age composition, socio-economic status etc. A study placed in a school setting, could simply test measures with other pupils and would also best include the opinion of teachers on the relevance of the instruments, their practical applicability, the meaningfulness of the analysis and potential ambiguities. Pre-tests can and should involve quantitative and qualitative means of assessment, that is an

explicit testing of survey items but also a one-to-one discussion of the developed questionnaires, interview guides or other means of assessment. The questionnaires and interview guides applied in each of my studies have been pretested and discussed with: seniors living in assisted living; school children in secondary school and teachers; colleagues using and not using abgeordnetenwatch.de.

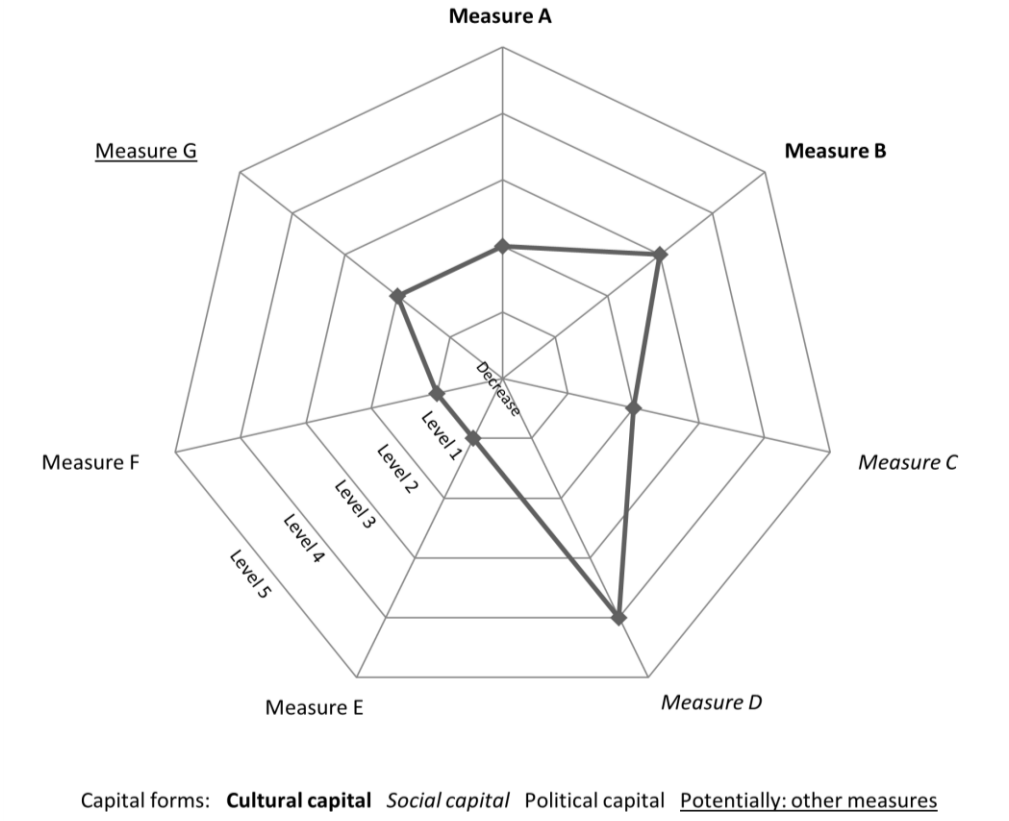
Where it is very difficult, hardly possible, or extremely resource intensive to test the assessment with beneficiaries or groups similar to those to be studied, it is the more important to engage closely with the organisation under investigation. The staff of the organisation will know better than the analyst about the characteristics of specific user groups and their potential views on the organisation's work and how it might affect them. This is indispensable for developing a measurement concept that is comprehensive and concise at the same time. When engaging with the organisation, it is however important to maintain a neutral distance between the researcher executing the study and the studied organisation, which will have a biased standpoint. At any rate the impact assessment should be iterative in character and allow for the revision of measurement components based on insights gained during the research process. This includes the revision of complementation of existing measures just as the inclusion of new ones in case new capital dimensions turn out to be relevant. In relation to my research this has not only happened within, but also across cases, as will be displayed in detail in the discussion of the results across the three cases. The presentation of the results of the analysis is the last aspect to be considered.

4.4.7 Analysing data and presenting results

Seventh and eventually, analysts need to think about how to display the results. The problems of monetisation (see again for instance Marée & Mertens, 2012) have been discussed and it was therefore dismissed in my studies. Dashboard approaches are currently favoured

in social-environmental reporting over the composition of an index, for the greater informational content the former bear (Fioramonti, 2013; Porter, Stern, & Loria, 2013). After screening for available alternatives, I have decided to use radar charts for that matter (see for instance Schmid, Schütz, & Speckesser, 1999). As will be seen, they are useful for summarising, evenly scaling, and portraying the states or changes in different capitals. At the same time they offer detailed information on capitals and their individual components. Capitals, if manageable in numbers of individual components, can be displayed in one radar chart, or different radar charts can be used for different capitals. Figure 2 is an exemplary illustration of a radar chart exhibiting the levels of several capitals (marked with different fonts), and their individual components (measures A-G), in relation to one group (dark line). Radar charts can however also be used to portray capital changes within a single group for pre-post observations on that group, or compare states and/or changes in capitals between two or more groups in case of a controlled research design. Besides their fit with different research designs, radar charts can be used to differentiate between different types of beneficiaries within a group.

Figure 2 Radar chart: exemplary display of capitals measured



The charts maintain a critical degree of complexity in communicating the results of the impact measurement, while easing their interpretation. The scaling contained in the charts is adaptable to different ranges and intervals. It can be both ordinal and verbally described or metrical. It can potentially also incorporate decreases in the capital components (negative social impact, illustratively marked by "decrease" here). Finally, radar charts can also include other (complementary) measures, for instance such that relate to the other conceptual approaches like satisfaction. All of these traits make radar charts apt for reporting the results of impact analyses. Before moving on to the later themselves, I give an appraisal of the methodological framework sketched out here.

4.5 Properties of the methodological framework

While my conceptual approach is new, my methodological framework might be criticized for relating strongly to the established practice of programme evaluation in terms of research designs. Don't we need new methods to go with a new conceptual approach? And don't we reproduce the deficits in past practice by applying traditional methods? I argue the answer to both questions is no and I do so for several reasons: We have seen earlier that existing attempts of measuring social impact are far from applying traditional designs rigorously (see Savedoff et al., 2006 on deficiencies in the state of practice in development assistance). Outlining how empirical research could improve in this regard, might be considered an innovation on its own. At the same time research designs and the principles of good scientific inquiry, when applied too rigidly may not be able to account for the specificities of social interventions (see Byford & Sefton, 2003 on the limitations of quantitative research in social care or Deaton & Cartwright, 2016 on restrictions to the usefulness of RCTs). The methodological framework outlined tries to strike a balance between the two extremes. It has a number of specific properties and (potential) consequences, which tie it to common standards and at the same time, differentiate it from the latter.

First, it helps calibrating and focusing the assessment and therefore represents a consequent formalisation of the capital based approach. Second, it avoids wasting resources through the aforementioned focus and considerations on when to apply the capital based approach or impact measurement at all. At the same time it enables organisational learning by tying organisational mission and impact measurement closely together. Third, it highlights the importance of rigorous research designs, but through the distinct focus on capitals and their individual components, keeps the realisation of the design manageable. Fourth, the framework gives equal importance to qualitative and quantitative methods and

thereby enables a comparison of the levels of capitals just as a tracing of changes in capitals. Fifth, it explicitly taps existing research and contributes to a systematisation of the applied measures. Sixth, the framework does not only contribute to the assessment of individual interventions, but also helps develop benchmarks across different interventions in the same area. Due to its integrated approach it fosters the development of common standards and comparability without the need for converting measures into a 'single currency', for instance through monetisation. What makes monetisation particularly attractive at present is often the lack of classification and structured display of impact measurement components.[16] Seventh and lastly, the analytical framework highlights the procedural and iterative character of the measurement and the necessity of adapting and complementing measures as new insights evolve during the process. The procedure can therefore be described as reflective.

A clear downside of the methodological framework I propose and use in this research are the transaction costs it involves, which will be a real issue for many social purpose organisations (Sawhill & Williamson, 2001). In particular quantitative studies that need one-to-one questioning (person assistance), such as the one performed in the living facilities for the elderly, may require middle ranged five digit budgets. Another difficulty is the effort involved and the expertise needed for performing the screening of applicable measures, their adaption and application. The good news in this regard is that there are many resources available that can be made use of. How this has been done and which insights it

[16] It is to be remarked that monetisation is not impossible even within the framework. Capital measures could be translated into monetary values by means of "income compensations." The monetary value of "non-market goods" (here: the capitals and their constituting components) could be assessed by various utility-based estimation techniques of how much income people would be willing to forgo for some higher unit of the non-market good (see Dolan, Fujiwara, and Metcalfe (2011) on the specifics of such techniques, and Then et al. (2012) for an exemplary application). Please note that the criticism towards preference-based measures, which I discussed earlier, would of course prevail if such a conversion strategy were applied.

has produced will be demonstrated in the presentation of the three empirical investigations performed.

In order to provide a better overview of the empirical work, I have not integrated the cases into one empirical chapter, but present three separate chapters. The individual chapters focus on the knowledge generated on the studied social purpose organisations' impact. An in-depth outline of the specifics of applying the capital based approach within the analysis across all three cases is provided in the discussion chapter. It illustrates how the empirical application has pressure-tested my conceptual proposition, which insights it has generated on the different capitals and their constituting components, and how it has enhanced the performance of social impact measurement overall.

Please note that before the research on each empirical case is presented in detail in the separate chapters, I give a short overview of some key characteristics of the study and some of the main insights gained for setting the scene.

5 Social impact of community oriented housing

This chapter deals with the results of the investigation of community oriented housing (CO) with a particular focus on the social capital of the inhabitants of such models.

Community oriented housing is a new living arrangement for elderly people. It tries to create a setup that is similar to people's original living environment by promoting interaction among inhabitants. It is supposed to create supportive relations between people and encourages individuals to be active. The empirical study examines whether the models really have an effect on inhabitants' social network, mutual support, participation in activities of the facility etc. This is tested in comparison to 'assisted living' (AL) as a standard model of living for the elderly, with a more pronounced focus on service provision than on self-activation. 260 people living in either of the two models have participated in a person-assisted survey in two consecutive years. Aspects considered in addition to social capital comprise the level of trust (cultural capital), or engagement in taking over communal tasks (political capital).

The performed analysis finds that community oriented models, with few exceptions, create higher levels in the assessed capitals, with the promotion of social network strength to and instrumental support from neighbours as the most prominent effects. The results do not only provide insights into the individual variables but also point to connections and differences between the latter through the joint conceptual framework applied. The main difference between community oriented models and assisted living does not lie in a higher absolute level of the measured capitals but in their greater spread among inhabitants. Community oriented models can thus be considered more inclusive. This is of particular relevance against the challenges of individual isolation in old age and its negative effects.

5.1 Background

Assessing social impacts, in particular those that seem to defy measurement, is promoted as an area of highest priority for research on long-term care provided by Home and Community-Based Services (HBCS) (Wysocki, Butler, Kane, & Shippee, 2013). The idea of fostering HBCS has two sides to it. First, it is a policy goal motivated by the ambition to save costs, which are usually higher in institutional care. Second, HBCS is also supposed to be preferred by the consumer to more institutionalised care (Wysocki et al., 2013). Especially at low to medium levels of impairment, multi-generational housing, where older and younger people co-reside, and other 'community oriented' living facilities (Kehl & Then, 2013) present a new alternative to standard 'assisted living' (Hawes, Phillips, Rose, Holan, & Sherman, 2003; Zimmerman, Sloane, & Eckert, 2001), where the focus is more on formalization and a 'service component'.

Community orientation responds explicitly to the European policy of strengthening society's self-regulating capacity and social cohesion (Crepaldi, Rosa, & Pesce, 2012, 32ff.). The term 'social capital' strongly relates to this, as it denotes the bonds and interactions between people and their participation in social life (Coleman, 1988). With regard to the field of care for the elderly, "low levels of social intercourse" (Ames, 1990, p. 674) are likely to have adverse effects, and depression is only one example. Hanson et al. (2013; 2012) highlight there is a growing evidence base which shows that social environments (see Barnett & Casper, 2001, p. 465 for a definition) affect e.g., the mobility or health of elderly people. However, "[…] few studies specifically address the complexity of community participation at the person, neighbourhood, and societal levels" (Hanson et al., 2012, p. 1).

Research on social capital has traditionally focussed on family or friendship but can be fruitfully transferred to the field of care (Rogers et al., 2011). Comparable research approaches are hardly found. There are some that address social capital in care explicitly (Rogers et al., 2011; Vassilev et al., 2013). The cited studies investigate how social networks can facilitate the (self-)management of long-term conditions. However, the authors focus on how social networks affect the process of care and not on how social capital is promoted in different care settings. Zimmerman et al.'s (2013) investigation relates more closely to the latter idea. The authors have systematically reviewed studies on care in nursing homes versus residential care/assisted living for patients with dementia. However, the studies contained in the review relate to the social environment—organisations, structures, processes involved in providing the care—but do not analyse the bonds between people. What is more, measuring direct health outcomes (e.g., pain, sleep quality or mortality) and psycho-social outcomes (anxiety, behavioural symptoms) in relation to wider social contexts may invite many confounding variables.

In my study instead, I focus on variables that are more directly linked to the living models instead, such as relations between neighbours, time spent participating in joint activities, or the amount of individual initiative taken for the community of neighbours. Data was collected from three organisations: One that only operates community oriented models (provider A), one that only operates assisted living (provider B) and one doing both (provider C). Elderly people have been questioned in a paper-based and person-assisted survey (effectively they were interviewed and research assistants noted the answers). In total, 26 sites have been visited by the research team.

In the context of housing for the elderly both versions of social capital I identified in my theory section will be of relevance: capacity-building and social interaction. This derives

from the two main functions that community oriented housing is supposed to have (Kehl & Then, 2013; Netzwerk: Soziales neu gestalten, 2009): First, CO shall activate inhabitants and enable them to address their needs, through their own capacity but also by drawing on the support of others. The support given and used may span different categories, for instance instrumental and emotional support, or companionship (van der Gaag & Snijders, 2004). This function is in line with the 'capacity-oriented view' of social capital (Flap & Völker, 2004; Grootaert et al., 2004). Second, CO is meant to stimulate community life. The hope is that people will benefit from many social contacts and frequent encounters with others, mutual interaction and exchange. For the second aspect, the 'interactive view' of social capital is more relevant.

Regardless of which version applies and as also outlined earlier, I restrict 'social capital' to the "relations, connections and exchange between individuals", which comprise "familiarity, intimacy, interaction, [or] mutual support" (see theory chapter). I delineate trust from the latter due to its character of a norm that might be enhanced but is not necessarily tied to existing social relations. In addition, I allocate variables that identify activism or agency to the category of political capital, which comprises for example self- and collective efficacy or engagement within the facility. Several satisfaction variables, for instance in relation to life in general or health status, are considered to complement the analysis. Before I move on to explicate the methods applied in more detail, I outline the general differences between the two residential models.

5.2 Specifics of community oriented housing

The two models, community oriented living (CO) and assisted living (AL), differ in a number of respects. The traits distinguishing CO and AL, which are summarised below have emerged from discussions with executives and senior managers of the studied provid-

ers and verified by a survey targeting the personnel running the facilities. The identified differences are located in six different areas:

(1) Meeting place: CO facilities perceive and present themselves as a meeting place for many different people instead of a mere living facility for the elderly.

(2) Open events: The majority of events and activities taking place in the facility are not restricted to the tenants but explicitly invite participation from people living in the vicinity.

(3) External offers: Events and other activities offered in the facilities are often convened by or in cooperation with external organisations (church organisations, associations etc.).

(4) Pro-activity of the tenants: Activities taking place in the facility are mainly shaped by the inhabitants, not the provider. The inhabitants do not merely participate in activities, they are proactive in making them happen.

(5) Tenants' influence: Also beyond events in habitants in CO have greater influence in determining the character of the facilities and shaping community life. They are encouraged to bring in their ideas and to realise them. This involvement is informal and continuous rather than based on formalised mechanisms. CO facilities all have a tenants' council, but this is not a criterion distinguishing them clearly from AL, which also often have one.

(6) Promotion of self-initiative: The facility managers foster and demand self-initiative from the inhabitants in many regards, for instance in problem solving or helping each other out. Self-initiative is more pronounced in CO than services delivered by professionals.

There are other areas, in which the living models do not differ remarkably. The general level of working connections with other organisations in the vicinity can be quite pronounced in both CO and AL (please note that it leads to more joint events in CO though, as outlined in number 3 above). The attempt of providing a good fit between incumbent tenants and new ones is another shared trait.

The hypothesis to be tested here is that due to the specific traits CO has over AL, community oriented models have an impact in that they enhance the formation of inhabitants' capitals, social capital in particular. This is based on the assumption that the above traits will spur the capacity-oriented and interaction-based functions of social relations outlined in relation to existing theory. Whether the hypothesis can be confirmed is to be tested by the following methods.

5.3 Methods

5.3.1 Setting

One part of the CO facilities visited was located in the south of Germany (provider A), the other one in the north (provider B). The comparator, namely AL facilities, were also located in the south (provider C) and north (provider B). The number of inhabitants for each of the four 'sub-groups' was roughly equal. All facilities in the north had the same provider, who operates both models (CO and AL). This was in contrast to the south, where two providers were involved, each one operating a single model. There was some variation in the size and foundation dates of the different facilities. The variation was, however, quite equally spread across the two living models, so that comparability was not impeded.

It needs to be remarked that provider A, who operates on a CO model has an explicit focus on multi-generational housing, where younger people, mostly families, live alongside el-

derly people. The CO facilities of provider B do not have this explicit focus and the facilities are inhabited by elderly people only. However, provider B also tries to foster the multi-generation aspect through the above-mentioned measures such as opening the facility to the surrounding environment, but also through locating and operating child care facilities at close proximity to the living facilities for the elderly.

Since living facilities for the elderly are usually not very big in size, ranging from about 40 to little over 100 dwellers in the present case, it was necessary to include a number of different facilities (26 in total; both models included). Participants were approached via the organisation running and maintaining the respective facilities and the research has been sponsored financially by the participating organisations.

5.3.2 Participants

The research was characterized by a low degree of intrusiveness and enabled inhabitants to voice their concerns, it was challenging but not too hard to recruit and retain participants. To be included, inhabitants had to meet at least one of the following criteria: (1) be 60 years of age and older (6 cases below this threshold); (2) be 'severely disabled' (schwerbehindert), which is defined by an officially testified degree of disability of above 50% (youngest participant: 50 years old, disability at 80%); (3) have a 'care level' rating, ranging from 1 as the lowest level to 3 as the highest (youngest participant: 47 years, care level 2).

In 2015, 339 people participated in the survey (CO: 132 = 73 at provider A + 59 at provider C; AL: 207 = 104 at provider B + 103 at provider C), while only 260 people participated again in 2016 (CO: 110 = 59 + 51; AL: 150 = 74 + 76). The loss in follow-up was mainly due to two reasons: (1) Casualties among the inhabitants, (2) missing motivation to partici-

pate a second time. The drop-outs due to the second reason were particularly severe at one facility of provider B, where a difficult personnel situation due to long-term sickness of the facility's manager had led to considerable dissatisfaction among the inhabitants, who for that reason chose not to support the investigation again.

The participation numbers show that inhabitants in AL, where the decrease was almost 30%, were more likely to drop out of the study than those in CO (decrease of participation by about 15%). Since we might suppose that people equipped with lower capital levels will have a higher inclination to drop out than those with higher levels and there is no sound reason to assume the reverse, it is more likely that the measured effects of CO against AL will be underestimated than overestimated. All analyses have exclusively been carried out in reference to the inhabitants who participated in both years to be able to include a longitudinal component in the analysis.

5.3.3 Analytic strategy

In principle, inhabitants could have completed the survey (a paper-based questionnaire) by themselves. However, the inhabitants actually needed assistance in completing the survey due to its length and complexity, which is why they have been assisted by researchers in doing it ('interviews'/completion of individual questionnaires took 40-75 minutes).

Participants have responded to the survey twice, once in 2015 and once in 2016. The survey was repeated for two main reasons: First, the second survey should test the stability or relativity of the results from the first round. Longitudinal data allow for better causal attribution than cross-sectional data. Second, the second survey implemented a number of new questions which emerged as promising follow-up issues from the first survey, or were meant to clarify or probe issues that could not be answered in the first round in a satisfacto-

ry fashion. As will be seen in the discussion of results, the second survey has served both purposes well.

5.3.4 Accounting for group characteristics: matching and regression

As outlined in the general methodological commentary, observational controlled studies need to ensure that the differences between the groups detected in the measurement are not due to different inclinations or other characteristics of the groups. Matching methods can help control for confounding and so can regression models. Both techniques are combined in my study, which thus applies a so called "doubly-robust approach" (Imbens & Wooldridge, 2009). How they have been applied is outlined in the following sub-chapters.

5.3.4.1 Initial differences between groups

I gathered information on many variables that characterized the two groups. These were used as control variables to filter out (matching) and account for (regression) confounding. I chose a wide range of variables in order to make sure that there was no hidden mechanism pre-determining the allocation of one type of participants to one group and a different one to the other. In other words, a mechanism that made the populations of the two living models fundamentally different.

Among the control variables were some socio-demographic items that are typically used in social science research, such as age, sex, education, income, health status (here for instance: the number of chronical diseases, or the degree of disability). I have, however, also included further variables with a relevance more tailored to my investigation. Most important in this regard were questions (partly more-dimensional scales) on the predisposition of inhabitants in terms of their social connections or their trust in people. By control-

ling for these aspects, I ensured that CO models were not inhabited by people who had a much higher level of social connectivity or trust than those living in AL to begin with. Predisposition had been measured by asking people to answer the applied questions/scales in retrospect of their life. While all these data helped increase the robustness of the analysis, they also increased its complexity.

Overall 22 control variables were used in the analysis, all of which could theoretically influence both, the allocation into the intervention and control group, and the outcome variables. They can be grouped into five categories: (1) socio-demographics; (2) family; (3) facility; (4) health; (5) predisposition in relation to the outcome variables (see Table 5).

Table 5 Matching variables and differences

#	Category	Variable	Values	B	Sig.
1	Socio-dem.	Gender	female-male		
2		Age	4 categories	-0.500	***
3		Members of household	Dummy (1-2)	1.096	*
4		Years of residency	Years		
5		Education	Dummy (0-1)		
6		Income per person	5 categories	-0.430	***
7	Family	Kids & Siblings	#	0.193	**
8		Family (wider) freq. of meetings	#		
9		Upbringing by single parent	Dummy (1-2 parents)		
10	Location	Region	Dummy (North-South)		
11		Locality	Dummy (remote-central)		
12		Distance previous residence	Dummy (< > 20 km)		
13	Health	Health-related Quality of Life	Index 0-1		
14		Disability High (>50%)	Dummy (severely disabled yes - no)		
15		Care Level	Three levels: 0/1/2		
16		Chronical Diseases	Dummy (< > 2 conditions)		
17	Predisposition	Activity Preferences	Scale, 4 items, 1-5		
18		Trust	Question, 1-5	0.783	*
19		Self-efficacy	Question, 1-5	0.658	**
20		Support for individuals	Question, 1-5		
21		Engagement in groups	Question, 1-5		
22		Social contacts (quant and qual)	Scale, 4 items, 1-5		

Remark: */**/*** Significance at 5/1/0.1 percent level.

Both applied scales on predisposition, namely the one on the "activity preference of the person" and the one on "the previous social connectivity of the person" (see items 17 and 22 in Table 5) have been tested for consistency by means of factor analysis and Cronbach's alpha (Cronbach & Meehl, 1955), which ensured that scales were reliable (alphas: activity preference=0.77 (item 17); social contacts=0.76 (item 22)).[17] The "social connectivity" consists of a combined measure taking into account quantity (number of contacts and frequency of meetings) and quality (importance and intimacy of the contacts). The same procedure was used to test whether a combination with other variables would have made sense. However, this was not the case, which is why trust, self-efficacy, support for individuals, and engagement for groups (items 18-21, all single questions) have been maintained as separate items. Furthermore, health-related variables (items 13-16) were correlated, but not to a degree that would have justified combining them. In line with the predisposition variables, the index (questions lent from EQ-5D, Herdman et al., 2011) on heath-related quality of life (HQoL) in 2015 was considered the baseline, against which HQoL in 2016 as a complementary effect measure was compared to detect potential effects.

Table 5 does not only display the variables and how they have been measured ("values"), but also documents the direction and the magnitude of the difference of CO versus AL (regression coefficient B). Only the differences that proved statistically significant are reported and significance levels are indicated.[18] All analyses presented here and in the results section have been performed by using SPSS Statistics 24 and the R plugin for SPSS (R version 3.2.5). Differences in the 22 variables were detected by application of a binary logistic regression on the membership in one of the two groups as the outcome variable.

[17] See Table 29 in the Appendix for more details.

[18] All matching variables have been tested for multicollinearity and all variance inflation factors (VIF) were clearly under the threshold of 3, which even under conservative assumptions is considered a very first hint at multicollinearity O'Brien (2007).

Only six variables showed a significant difference: age, number of households with two members, income, size of the immediate family (own siblings and kids), trust and self-efficacy PRE.

5.3.4.2 Matching procedure

The goal of the propensity based matching procedure (Austin, 2011a) that followed was to eliminate all existing differences between the groups.[19] The probabilities of belonging to one or the other from the performed logit, also called propensities, were used for this purpose. Applying a standard 'propensity score matching' would have reduced the number of total observations by statistically pairing one individual from the intervention group to one or more individuals in the control group with a similar propensity (caliper matching; see Austin, 2011b). Instead of using the latter approach, all observations have been preserved but weighted by the inverse of their individual propensities to level out differences.[20] The benefit of this 'inverse propensity weighting' (IPW) (Austin, 2011a) is that loss of information is avoided. In particular the reduction of the overall variance of the data can have negative effects on the significance of the performed tests, which is particularly important here due to the relatively low number of observations. In order to supress too heavy influence by single cases some weights have been truncated, which should improve the precision of follow-up analyses (Lee, Lessler, & Stuart, 2011). A truncation point of 10 has been chosen, which lay at the 97th percentile. IPWs have then been normalised.[21] The maximum normalised weight was located at around 4, which is consistent with typical weighting procedures (compare to European Social Survey, 2014a).

[19] Applying 'coarsened exact matching' (CEM) (Iacus, King, and Porro (2011)) has also been considered but dismissed due to the magnitude of variables used. The large number of variables in combination with the strictness of the matching per variable, by which CEM is marked, would have reduced the number of inhabitants too much. It seems more suitable with rather small number of variables.

[20] Those in CO were weighted by 1/(propensity score), while those in AL received a weighting of 1/(1-propensity score).

[21] Initial weights have been divided by their group arithmetic mean, so that the total number of cases when weighted remained at 260.

After IPW the differences between the groups disappeared entirely. The logit when repeated did not yield significant differences anymore, neither in those where differences had been detected previously, nor in the other variables. As an additional criterion for the quality of the weighting Nagelkerkes R^2 as a measure of model fit had decreased from 0.248 at the beginning to 0.011 after the weighting. Since our presumption is that the control variables used are not good in predicting group allocation, in other words that groups are not different, a low level of the R^2 is preferable. As yet another test of the effectiveness of the weighting standardised mean differences (SMDs) between the two groups were calculated for each control variable. They were all below 0.1, which serves as a conservative border of imbalance in the data (McCaffrey et al., 2013; Rubin, 2001; Stuart, Lee, & Leacy, 2013). The low SMDs suggest that the groups are very similar after the IPW. Finally, this similarity has also been verified through displaying density plots of the variables across the two groups. The latter showed that the groups had become more similar after the IPW.

5.3.4.3 Group characteristics after weighting

For their literal equality, the groups' characteristics are outlined here by a single figure on each of the items listed above. The exact figures are displayed in two separate tables (Table 6 and Table 7) for the different format of the variables.

Table 6 Control variables after weighting (part 1)

#	Category	Variable	Values	AL		CO	
				N	%	N	%
1	Personal traits	Gender	f	118	78.7%	85	77.3%
			m	32	21.3%	25	22.7%
2		Age	<70	12	8.0%	9	8.2%
			71-80	62	41.3%	48	43.6%
			81-90	60	40.0%	43	39.1%
			>90	16	10.7%	10	9.1%
3		Members household	1	120	80.0%	85	77.3%
			2	30	20.0%	25	22.7%
5		Education high	No	115	76.7%	83	75.5%
			Yes	35	23.3%	27	24.5%
6		Income	0-1000	21	13.9%	16	14.4%
			1001-1380	30	19.9%	14	12.6%
			1381-1700	38	25.2%	29	26.1%
			1701-2040	14	9.3%	28	25.2%
			>2041	48	31.8%	24	21.6%
9	Family	Single parent	No	19	12.7%	14	12.7%
			Yes	131	87.3%	96	87.3%
10	Facility	Region	North	82	54.7%	63	57.3%
			South	68	45.3%	47	42.7%
11		Location	remote	72	48.0%	57	51.8%
			central	78	52.0%	53	48.2%
12		Distance prev. res.	< 20 km	101	67.3%	73	66.4%
			> 20 km	49	32.7%	37	33.6%
14	Health	Disability	< 50%	94	62.7%	66	60.0%
			> 50%	56	37.3%	44	40.0%
15		Care level	0	116	77.3%	84	76.4%
			1	30	20.0%	20	18.2%
			2	4	2.7%	6	5.5%
16		Chronical conditions	< 2	86	57.3%	59	53.6%
			> 2	64	42.7%	51	46.4%

The following can be said on personal traits: (1) about 80 percent of the respondents are female; (2) about 10 percent are less than 70 years old, 40 percent are between 70-80, 40 percent between 80-90, and 10 percent older than 90; (3) about 80 percent of the inhabitants live alone; (4) about 25 percent have acquired a higher educational degree (high school or higher vocational education); (5) the average duration of residency is 7 years; (6)

almost 35 percent of the participants have a monthly income below 1380.00 €, 25 percent up to 1700.00 € and another 40 percent above 1700.00 €.[22]

The family traits of the inhabitants are these: (7) about 13 percent of the people have been raised by a single parent; (8) on average respondents have 4 siblings and/or kids (closer family); (9) they meet with their wider family (for instance at family festivities) about 4 times a year.

Table 7 Control variables after weighting (part 2)

#	Category	Variable	Values	AL	CO
4	Personal traits	Length of residency	Years	7.27	6.97
7	Family	Siblings&kids	#	4.16	4.15
8		Freq. meet family	# per year	4.04	4.32
13	Health	HQoL	Index 0-1	0.76	0.75
17	Predisposition	Activity preference	Scale	3.18	3.20
18		Trust	Question	2.80	2.82
19		Self-efficacy	Question	4.34	4.31
20		Support for individuals	Question	4.00	4.01
21		Engagement in groups	Question	3.11	3.17
22		Social connectivity	Scale	3.71	3.70

The facilities in which the respondents live exhibit the following traits: (10) about 55 percent are located in the North of Germany; (11) about 50 percent are rather remote, the others central (judged by the providers‘ staff as well as the research assistants, who conducted the survey—this does not mean that some are located in an urban, others in a rural context, it is more about the available infrastructure of the immediate vicinity; (12) for almost 70 percent of the tenants their last residence is located at a distance of 20 km or less.

[22] The categories applied here represent the lower deciles of the German income distribution from the European Social Survey (2014c).

The inhabitants‘ health traits can be described like this: (13) on average the health-related quality of life HQoL is at 0.75 (on an index from 0-1); (14) about 40 percent of the inhabitants have a ‘high degree of disability’ (> 50% disabled); (15) about 75 percent do not have a care level, while about 20 percent have care level 1, and almost 5 percent have care level 2; (16) about 55 percent of the inhabitants have 2 or less chronical conditions.

In particular with single questions on the predispositions of inhabitants (on a 1-5 scale from “very low”-“very high”) it is usually not advisable to indicate averages, due to the fact that categories are not necessarily equidistant. This is why I only indicate a general tendency for all of them, although activity and social connectivity have been measured by a multi-dimensional scale: (17) the inhabitants’ activity preference (activity versus passivity) is „medium“; (18) their predisposition as regards trust is also „medium“, but at a somewhat lower level; (19) the self-efficacy level is slightly above the „high“ threshold; (20) the inhabitants‘ estimation of how strongly they’ve usually supported other individuals is also „high“; (21) the one relating to their engagement in groups at a “medium” level; (22) the inhabitants’ social connectivity (a combined measure taking into account the number and frequency of contacts as well as their importance and intimacy) tends towards a “high” level.

5.3.4.4 Regression analysis

After the matching procedure, effect tests have been carried out with the IPW weighted data.[23] As already mentioned a “doubly robust” approach has been chosen for the analysis (Imbens & Wooldridge, 2009), whose purpose was to further increase the reliability of the analysis. Instead of simply analysing mean or median differences by means of t-tests or

[23] Effect/outcome/dependent variable are used interchangeably to refer to the measures by which differences in CO as compared to AL have been detected. The illustration of effects and their discussion take exactly this direction: „How much higher or stronger is variable X in CO than in AL?”

non-parametric tests, a regression has been performed for each individual effect variable (to be presented further below). Each outcome variable has initially been tested for the influence of the living facility (binary: CO or AL) by a model additionally comprising all 22 control variables that had also been used in the matching. All independent variables have entered the model as main effects, that is without interaction terms. The model was subsequently reduced by maintaining only the variables which were statistically significant. This has been done to test the robustness or sensitivity of the detected effect and base the documented effect results to be reported later on the leanest underlying model. All measured effects, both in size and in their significance, were remarkably insensitive to the reduction of control variables throughout the analysis and thus very robust.

All regressions were performed by means of Generalized Estimating Equations (GEE) for two reasons: First, most of the dependent variables were not normally distributed. Count data (number of persons or time intervals) usually came in a negative binomial distribution, with a large share of persons at the level of 0 or with low values in the variable. For instance, substantial shares of inhabitants did not receive any support in their daily life by their neighbours at all. Second, not all observations are independent. 20 percent of inhabitants live together with a partner and partly both had been interviewed. Third, the main interest lay on population-average comparisons rather than individual level inferences. GEEs accommodate for these peculiarities (see Homish, Edwards, Eiden, & Leonard, 2010 on partly dependent observations and Hubbard et al., 2010 on GEEs more generally). The regressions were run with robust estimators for the covariance matrix, which might lead to an underestimation of effects while making the results more reliable.[24]

[24] Analyses have also been re-run with flexible estimators not indicating major differences to the application of robust estimators, which is an indication that there are no major and undue distortions through the use of the latter (King and Roberts (2015)).

The distributions of all variables have been checked before running regressions. Metric data have been brought into a normal distribution through transformations where applicable (the trust and "satisfaction with the living facility" variables have been log-transformed). QICs ("quasi-likelihood under the independence model criterion") as a measure of the goodness-of-fit of the model applied within the GEE (Pan, 2001) were regarded to make sure that the reduced models used for reporting results were superior to the initial full model, whereby lower levels indicate better model fit.

Before I performed the regression analyses all missing data had been estimated by a multiple imputation procedure with five estimations. The effects reported in the next sections all refer to the combined results merging the original data with the five imputations. The number of observations is therefore always complete, with no missing data remaining. Before imputation the data set was visually and statistically tested for patterns in the missing data. The result of this analysis was that missing data were Missing Completely At Random (MCAR). This applied to the variables that were available for 2015 and 2016 and those only available for 2016[25] alike (Little's MCAR test for 2015-16: Chi^2=5660.435, DF=9898, p=1.000; for 2016: 2132.158, DF = 2177, p=0.750). Data imputation was therefore applicable.

5.3.5 Effect measures

In order to differentiate effects clearly, I have split the applied capitals into several components. The structure of capitals tested is reported in Table 8. As can be seen the explication to follow relates to 12 components, which partly have been measured by a single question, partly by a multi-facetted scale; one has also been assessed by a composed index. The paragraphs below describe the individual measurement items applied and give references to

[25] See next sections for specifics.

the previous research the measures have been derived from. An explicit portrait of the applied scales and questions is given in Table 27 and Table 28 in the Appendix.

Table 8 Capital forms and secondary variables

#	Capital	Category
1	Cultural	Trust
2	Political	Engagement
3		Collective efficacy
4		Influence
5		Self-efficacy
6	Social	Individual connections
7		Group cohesion
8		Support received
9		Support given
10		Participation
11	Secondary variables	Satisfaction with living situation
12		Life-satisfaction
13		Health related quality of life

It is obvious that social capital variables are most important and by which other variables they have been complemented. In addition to political and cultural capital, several satisfaction-based measures have been assessed.

Among the primary variables for social capital in my investigation is the inhabitants' social contact network (Flap & Völker, 2004; Grootaert & van Bastelaer, 2001; Lillbacka, 2006; Lin, 1999; van Duijn, van Busschbach, & Snijders, 1999). My measurement includes (1) the number, (2) the frequency, (3) the intensity and (4) the importance of contacts and thus not only the quantity but also the "[…] quality of the constituent ties" (Adler & Kwon, 2002, p. 24). The network strength is the most complex measure applied and needs some deeper explanation.

The strength of the inhabitants' social network (NWS) was measured by an index. This index was derived from a list of persons, the respondents were asked to name, with whom

they interacted most. This list comprised family members and friends as one category and neighbours as another. In addition to naming the persons, respondents were asked to report: how often they usually were in touch with the respective persons; whether they met in person or communicated via telephone; and how important as well as how intimate they assessed the specific contact to be. Based on this information the index was composed according to the formula below. The number of persons (x) was multiplied by two components: (1) the frequency of contacts with the person, whereby contacts by telephone were weighted by 30% and personal contacts by 70% (t=telephone, p=personal)[26], and (2) the importance (imp) and intimacy (int) of the relationship to the person, equally weighted.

$$\sum_{i=1}^{n} x_i * (0.3 * t_i + 0.7 * p_i) * \ (0.5 * imp_i + 0.5 * int_i)$$

The values for imp_i and int_i were derived from a 1-5 scale (1=low importance/intimacy, 5=high importance/intimacy). The frequency of contacts was initially measured in days per month. Based on the distribution of values, the data was then transformed to a 5-point scale too, where frequency was indicated as 1=very low (≤ once per month, that is seldom), 2=low (≤ 3 times per month, or every other week), 3=medium (≤ 8 times per month, or 1-2 times per week), 4=high (≤ 16 times per week, or up to 4 times per week), 5=very high (> 16 times per week, that is up to daily).[27] To increase usability of the index the two digit and

[26] There is long-standing research that suggests that telephone and personal contact are complementary rather than substitutive (Wellman (1979)). It also suggests though that people living close by more rarely use the phone to communicate and that intimate relationships cannot exist without some in-person contact. To strike a balance between these two poles, in particular in view of the close communities between neighbours, both have been included in the calculation but personal contacts received a considerably higher weight.

[27] This division was guided by the distribution of the data that suggested cut-off points. It is also oriented at previous research, yet more fine grained in two ways: it is able to differentiate in the direction of intense and very frequent contacts (especially relevant if close family is nearby and for neighbour contacts) as well as in that of more sporadic contact (relevant when the reason for moving into CO or AL is that close family has moved far away). Wellman (1979) instead differentiates only between "weekly and more often" and "twice per month or less."

partly three digit numbers were divided by 100, resulting in a range for NWS to family and friends in between 0-1.37 and 0-0.60 for NWS to neighbours as will further discussed in the results.

With regard to the social network, I disregard the division into bridging and bonding social capital (Putnam, 2000), since it is not instrumental to judging the social impact caused in this specific setting. Another important social capital variable the study examines is mutual support, more specifically (1) instrumental support, (2) emotional support and (3) companionship which "[i]n the social support literature [... have] gained some tradition" (van der Gaag & Snijders, 2004, p. 163; also Cohen & Hoberman, 1983; Cyranowski et al., 2013; McCallister & Fischer, 1978; van der Poel, 1993, 54f.). The level of support provided by a network is more central to people's physical and mental health than the mere size of the network (Lubben et al., 2006; Norstrand et al., 2012, p. 302). Mutual support is therefore also tested (Elgar et al., 2011; Pope, Miller, Wolfer, Mann, & McKeown, 2013; Stanca, 2009). This is done reciprocally, that is in view of how much support elders receive and how much they provide themselves. Also, the 'sense of belonging' (Hagerty & Patusky, 1995; Norstrand et al., 2012) is of importance in studying residential models and has therefore been included. Finally, participation in communal life has been measured simply by accounting for the time individuals spent on communal activities.

Other factors associated with social capital such as membership in voluntary organisations (Bjørnskov, 2006b; Elgar et al., 2011; Lillbacka, 2006) have not found explicit recognition in my measurement model, but tested as an exploratory dimension as regards links between the 'inside community' and the surrounding quarter.

For the reasons laid out initially and in my theory chapter trust or trustworthiness, just as other norms, are components of cultural rather than social capital. I have included trust in its 'faith in people' version (Morrone, Tontoranelli, & Ranuzzi, 2009 in relation to Rosenberg, 1956; Rosenberg, 1957).

Self-efficacy in turn, identified for instance by Lillbacka (2006) or Onyx and Bullen (2000), is of specific value where support structures influence individuals' ability to help themselves. 'General self-efficacy' (Schwarzer & Jerusalem, 1995) has therefore been included, but not as part of social capital as the cited authors have suggested, but as part of political capital, along with a measure of 'collective-efficacy' that refers not to the individual problem solving capacity, but to the capacity of a group or community (Patras & Klest, 2012). The same goes for engagement in the sense of inhabitants taking over communal tasks and being involved in decisions and actions shaping the facility and community, which therefore have also been subsumed under the political capital dimension.

The primary variables are assessed by a range of established scales as used in the previous work cited above.[28] Some secondary variables provide links to alternative measurement approaches, specifically satisfaction based measures, for comparative reasons. These include health-related quality of life (Wolfs, Kessels, Dirksen, Severens, & Verhey, 2008), several generic satisfaction items in relation to the living environment (Wright, Hazelett, Jarjoura, & Allen, 2007) and a question on overall life-satisfaction (Diener et al., 2010).

[28] In analogy to the procedure for the control variables, when scales have been applied instead of single questions or counts, these scales have been analysed by confirmatory factor analyses and Cronbach's alphas have been calculated. The scales in the versions applied in the analysis proved all sufficiently reliable (alphas: trust=0.85; self-efficacy=0.82; sense of belonging=0.77; satisfaction with facility=0.83). All alphas are given in relation to 2016, yet with little variation over the two years. See Table 27 and Table 28 in the Appendix for more details.

The focus on different capitals as laid out above has been confirmed in the empirical testing of the applied scales, both on the outcome and on the predisposition level. For example, the measures for trust could not be reasonably combined with the scale asking for the quantity and quality of social contacts (pre-disposition). Both the factor analysis and the testing for Cronbach's alpha suggested that trust was to be treated as a separate category. The same goes for the capacity of the inhabitants to solve problems in a joint effort, which was initially considered to be part of the communal sense of belonging and group identification. Details on this will become clearer in the presentation of the results.

Before we move on to the presentation of results it is to be noted that most variables had been assessed in 2015 and re-assessed in 2016. However, some variables were added to the analysis in 2016 in case an issue needed clarifying and thus additional questions or prompts for some additional probing questions had emerged in the preliminary analysis of 2015 data. The analyses on both years comprise an N of 520 observations, while those that refer to only one year comprise 260 observations. The years(s) the measured effect refers to is indicated explicitly in the presentation of results. For all analyses that related to measures assessed in both years, a binary variable "year" was used as an additional control variable in the regression model.

5.4 Results

Overall 21 individual outcome variables belonging to the 12 categories are displayed and discussed here. Some further variables had been tested, but are not illustrated here in order not to overload the presentation of results. Some of those that have not been pursued further are mentioned in the discussion. All effects contained in Table 9 refer only to the influence of the living model. The influence of control variables in the reduced model, including differences between the two years in which data have been gathered are not report-

ed for the purpose of a clear presentation. A more detailed documentation of the analysis can be found in Table 30-Table 32 in the Appendix.

In addition to the year(s) to which the results relate Table 9 contains the following information on items listed by capitals and their constituting components:

- "regression results" composed of: (1) regression coefficient B; (2) exponentiated regression coefficient Exp (B), except where not specifically useful;[29] (3) statistical significance of the effect;
- "translated effects" in different formats: (4) depending on the type of data either a translation of the effect into percentage points (for metric data, scales and indexes), or likelihoods/odds ratios (for ordinal data, single questions), or numbers and amounts (for count data);[30] (5) the type of measure (scale, index, number of persons, h per week).

In addition to the above the table indicates (6) the approximate "level" of capital components in AL and CO indicated by the mean or median; (7) and only for count data, the "share of non-zero values, meaning the share of inhabitants that the variable applies, and (8) the observed minimum and maximum values across the two groups.[31]

[29] Exp(B) indicates odds ratios instead of effects sizes for metric variables (scale, index), unless the latte have been log transformed (see trust and satisfaction). Effect sizes were however at the centre of interest and odds ratios only assessed for variables treated as ordinal (singe questions).

[30] Due to the distribution of the count data expressed in time, Exp(B) indicates units of 30 mins, which were then transferred back into h/week.

[31] To prevent undue influence from outliers the maximum value of count variables has been capped at a certain threshold after which only few and rather extreme values occurred. The corresponding individuals were allocated the capped maximum in the respective variables. The capping point for all variables but "spending time with neighbours" (item 15) was approximately at the level of the 98th percentile. For the spread in answers to item 15 and the extremity of the higher values, the threshold lay at the 94th percentile in the latter case.

Table 9 Overview of effects CO

#	Cap.	Category	Item	Year	Regression result			Translated effects				Level			Non-zero[+]		Observed	
					B	Exp(B)	Sig.	%	Like.	Quant.	Measure[x]	AL	CO		AL	CO	Min	Max
1	CC	Trust	Trust LOG	15-16	0.063	1.065	***	6%			Scale	3.42	3.64	M			2.50	4.95
2			Trust # NB	16	0.620	1.859	**			2.0	Pers	0	2	MD	41%	53%	0	10
3	PC	Engagement	Performing tasks	15-16	0.448	1.565	*			0.75	h/week	1.00	1.00	MDp	32%	47%	0	6
4		Collective efficacy	Problem solving	15-16	0.828	2.289	***		2.3		Question							
5		Influence	Inhabitants' influence	16	0.613	1.846	*		1.8		Question							
6		Self-help	Self-efficacy	15-16	0.064			2%			Scale	3.65	3.69	M			1.30	5.00
7	SC	Connectivity	NWS Fam&Fr	15-16	-0.074		**	-7%			Index	0.42	0.35	M			0.00	1.37
8			NWS NB	16	0.364	1.439	**	44%			Index	-	-	-			0.00	0.60
9			# NB friends	16	0.607	1.835	***			2.0	Pers	1	3	MD	64%	85%	0	15
10		Group cohesion	Sense of belonging	15-16	0.229		**	6%			Scale	3.88	4.11	M			1.25	5.00
11		Individ. sup. received	Support NB instrumental	15-16	1.114	3.046	***		3.0		Question							
12			Support NB compan.	15-16	0.675	1.965	**		2.0		Question							
13			Support NB emotional	15-16	0.355	1.427			1.4		Question							
14			# NB inst	16	0.652	1.920	***			2.0	Pers	0	2	MD	40%	73%	0	8
15			# NB compan	16	0.609	1.839	***			2.0	Pers	1	3	MD	58%	75%	0	11
16			# NB emot	16	0.503	1.653	*			1.5	Pers	0	0	MD	22%	40%	0	8
17		Individ. sup. given	Individual support h	16	0.413	1.511				0.75	h/week	1.00	1.00	MDp	27%	49%	0	6
18		Participation	Participation in events	15-16	0.387	1.472	**			0.75	h/week	1.00	2.00	MD	66%	85%	0	9
19	Sec.	Satisfaction	Satisfaction LOG	15-16	0.041	1.042	**	4%			Scale	3.75	3.91	M			2.58	4.95
20			Life satisfaction	15-16	0.580	1.786	*		1.8		Question							
21		Health	HQoL	16	-0.003			0%			Index	0.74	0.73	M			-0.14	1.00

Remarks: */**/*** Significance at 5/1/0.1 % level; [x]=Range of 1-5 for all scales and questions, except life satisfaction (0-10); [+]=Share of people with non-zero values in the variable; MDp=Median only for cases with non-zero values; M=Mean; MD=Median; CC=Cul. Cap.; PC=Pol. Cap.; SC=Soc. Cap.; Sec.=Second. variables; Like.=Likelihood; Quant.=Quantity; NWS=network strength; NB=Neighbours; compan.=Companionship.

Means are indicated only for metric data, and such that are normally distributed, in order to give an impression of the general level of variables in the living models and not only the difference between the two. This was not reasonable for the variable network strength (NWS) to neighbours due to the distribution of the variable.[32] For the count data, due to their negative binomial distribution, medians were reported as a measure of the central tendency. No such level has been reported for the ordinal variables, due to missing equidistance.

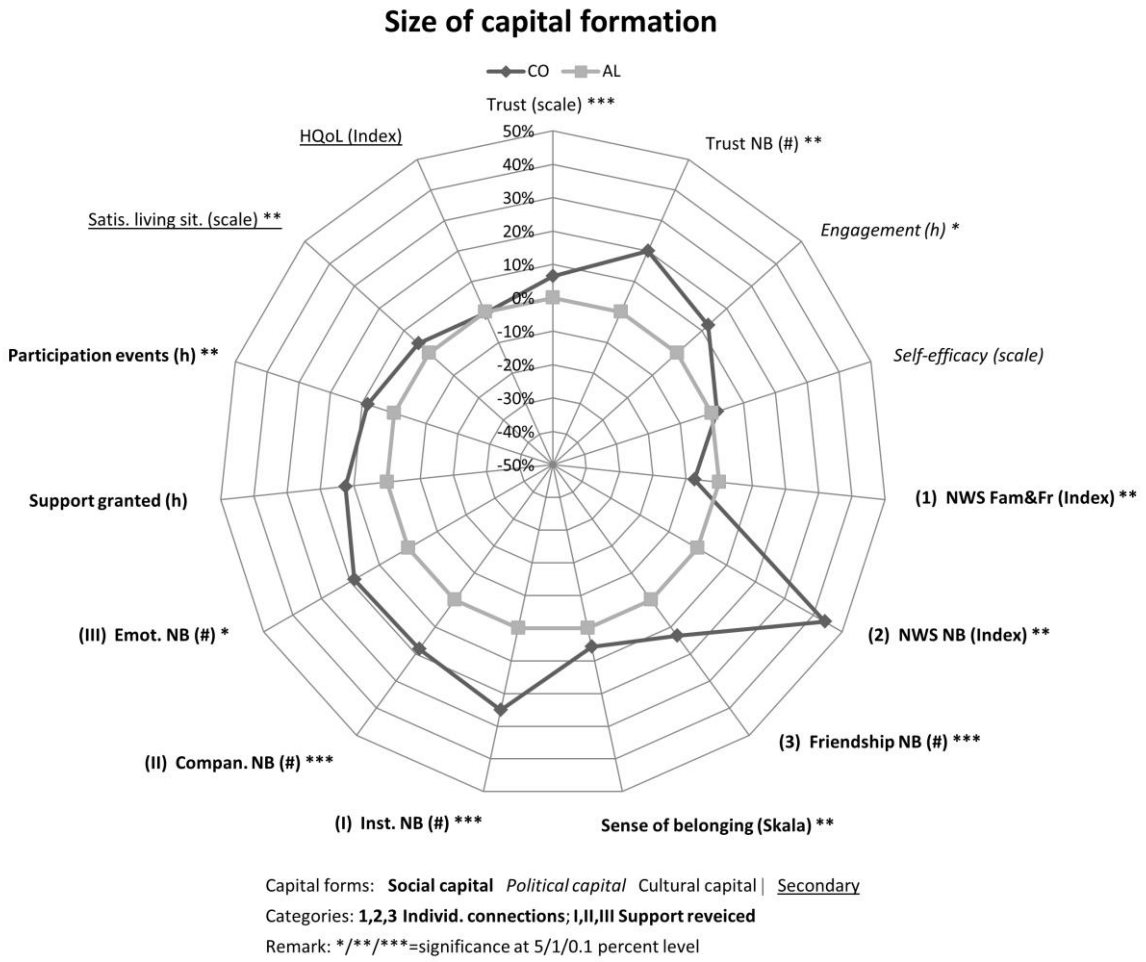
For the variables marked by "15-16" the reported results refer to the combined data from both years, the ones marked by "16" only to one year. Almost all variables newly introduced in 2016 relate to a quantification of the relationship to neighbours (number of neighbour contacts across several categories; items 9, 14-16). The question on the "influence of inhabitants on the facility" (item 5) has been added to complement the "problem solving capacity" at the collective level (item 4), which had originally been thought of as part of the "sense of belonging" (item 10) but couldn't reasonably be integrated.

As highlighted in the general methods chapter, there is also another way of displaying results in a comparative fashion: using radar charts (see Schmid et al., 1999 for examples for their application). Their strong point is that they are capable of illustrating effects across the different capitals and their components. As an information system of a 'dash-board' style, they contain more information than indexes, yet present measured effects in a way that is easily accessible to decision makers. They do not only have the potential to outline impact and thus performance of the investigated organisation across all relevant levels (as defined prior to the impact measurement), but also enable comparisons between discrete

[32] Distribution NWS to neighbours: Tweedie distribution with excess numbers of zero, so called "compound Poisson-Gamma distribution."

alternatives: Figure 3 illustrates the capital building and satisfaction effects of CO in comparison with AL.

Figure 3 Size of capital formation CO



The figure comprises only the 15 scale-based and count data (number of persons and units of time) and not the single questions from which only odds ratios could be derived. Different font styles are used to denote the different capital forms and secondary variables. In two cases within social capital the components are assessed by several measures: The first one is "individual connections" (components marked by Arabic numbers 1-3) and the second, "support received" (components marked by Roman numbers I-III).

Due to the difficulties in indicating the levels of a central tendency in the groups caused by the distribution of some of the data, all exhibited differences are provided in relative, not in absolute terms. Thereby AL is used as the comparative baseline level from which the values in CO deviate by the amount of the individual differences reported in the Table 9. The differences between CO and AL are indicated in percentage points, whereby differences in the scale-based data are calculated and reported in relation to the defined maximum value of the applied scale and those in the count data to the observed maximum value (also see Table 9).[33] The radar chart also reflects the significance of the diagnosed differences.

The table in conjunction with the figure outlines that CO produces higher levels of capitals across almost all components. However, it also shows that some of the differences are considerably more substantial than others: "Network strength to neighbours" and "instrumental support" stand out against the others. In the following I present and explain the measured effects as contained in detail and will show that most differences derive not necessarily from a higher level of capitals in CO in absolute terms, but mainly from a wider spread among inhabitants as indicated by the shares of non-zero values in Table 9.

5.4.1 Cultural capital

Inhabitants in CO have more trust in others than those in AL. The difference is highly significant, but with the effect being at 1.065 it translates into a difference of only 6 percentage points (CI95 Exp(B)=1.033-1.098, p=0.000). On the scale the means in the groups lie between a "medium" and a "high" level. The difference as regards the "number of neighbours" that inhabitants of the two living models "fully trust" is more striking. The differ-

[33] Here is an example of how CO and AL are compared in the radar chart: The time spent on "participating in events" (item 18) is 0.75 hours higher in CO than in AL and the observed values range from 0-9 hours. The level in CO is thus by about 8 percentage points higher than in AL in relation to the full scale. Since AL represents the baseline in the chart (is located at the level of 0 percent), the line of CO will be at +8 percent above AL on the percentage scale ranging from +50 to -50 percent (an even spread has been chosen purely for reasons of effective illustration).

ence of 1.859 indicates that people in CO have two more neighbours they “fully trust” than those in AL (CI95 Exp(B)=1.235-2.798, p=0.003). The median shows that people living in AL have a tendency not to trust any of their neighbours, whereas those in CO trust about two of their neighbours. It needs to be remarked though that this low level, especially in AL, derives from the fact that only 41 percent of the inhabitants have at least one neighbour they can trust, while the share in CO is at 53 percent.

5.4.2 Political capital

The inhabitants’ “engagement” within their facility, measured by the time spent weekly for executing communal tasks is higher in CO by 1.603 or about 0.75 h per week (CI95 Exp(B)=1.035-2.367, p=0.034). Examples for the latter comprise: administrative tasks; organising or helping with events; decoration and care for the facility buildings, including the garden; welcoming and showing around new inhabitants or other groups etc. As remarked previously, time-based data had to be transformed into counts of 30 min. units. The result suggests that the variable is higher in CO by 1.6 30 min. units, that is 45 minutes or 0.75 hours. The effect is significant at the 5% level. Due to the excess zero values in the variable, even the median is not a very helpful indicator for the level of the variable in the living models. If we look only at those, who are engaged at all (47 percent in CO and 32 percent in AL), the median is at about 1h/week in both models (MDp). It is thus by the share of people, who are engaged and not by the level of those who are engaged that the two living models differ.

The “collective efficacy” was assessed by the capacity for “joint problem solving” (2015 and 2016) and “inhabitants’ influence” in shaping the facility (2016). Due to the ordinal nature of the measures, the effect is expressed as an odds ratio. The effect sizes of 2.289 and 1.846 indicate that inhabitants in CO have a 2.3 times higher likelihood of solving

problems in a jointly (CI95 Exp(B)=1.452-3.611, p=0.000) and a 1.8 times higher likelihood of a higher individual influence than in AL (CI95 Exp(B)=1.105-3.085, p=0.019).

There is no statistically significant difference in the level of "self-efficacy" (here: the ability of the persons to help themselves even in difficult situations) across the two models. The difference suggested by the regression coefficient of 0.064 translates into 2 percentage points when related to the scale ranging from 1-5. The means are hence very close to each other and located between a "medium" and a "high" level.

5.4.3 Social capital

In the analysis of network strength, "NWS family and friends" relates to data from 2015 and 2016, while "NWS neighbours (NB)" relates to 2016 data only. The reason for it is the following: The preliminary analysis of data from 2015 yielded higher support from neighbours for CO (see below), yet a lower NWS NB compared to AL (although the difference was not statistically significant). This seemed counter-intuitive, which is why the survey in 2016 contained an explicit question on neighbour contacts. As a result this led to a stronger NWS NB in CO than in AL. The reason for this shift seems to be that in 2015 the inhabitants in CO only listed the most important neighbour contacts and thus not all substantial contacts they had with neighbours, while those in AL tended to list all neighbour contacts.

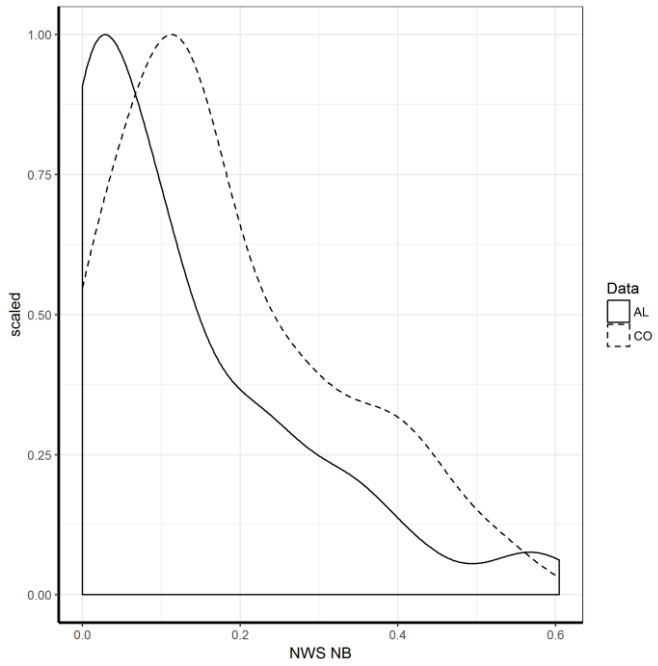
In fact, in the 2016 data there was a much stronger increase in the number of neighbour contacts mentioned by CO dwellers than those mentioned by the inhabitants of AL. In order to make sure that the difference did not originate from contacts newly established between 2015 and 2016, the respondents were asked to indicate whether the neighbour mentioned was an 'old' or a 'new' contact. This clearly showed that the difference detected in 2015 was not representative but due to different perceptions by the inhabitants of the two

models of what they should list. Due to the obvious misperception in 2015 this data was disregarded and only the one from 2016 used.

The analysis yielded a coefficient of -0.074 (CI95 B=-0.127 to -0.022, p=0.006), which means a weaker NWS to family and friends in CO by roughly 7 percentage points. The means, which are quite indicative in this case, since the variable is normally distributed, are located at about 0.42 in AL and 0.35 in CO. The maximum observed value is 1.35. Since 98% of the data however lie below 1, the index is effectively scaled from 0-1.

At the same time the effect of living in CO on NWS NB lies at 1.439 und thus roughly at 44 percentage points above AL (CI95 Exp(B)=1.124-1.843, p=0.004). Both NWS effects are significant at the 1% level. Due to the compound Poisson-Gamma distribution it is not reasonable to report a central tendency for NWS NB. However, we can illustrate the difference between the two groups by means of a density plot.[34]

Figure 4 Network strength to neighbours



[34] This is a scaled density plot, where the maximum density in each group is at a level of 1.

Figure 4 illustrates that CO has a substantially lower share of people with no or very low neighbour contacts. Once we move to the higher levels of NWS, we furthermore see that CO lies above AL, and at a constant distance. The difference detected by the regression is thus composed of these two effects.

A complimentary question about how many neighbours the inhabitants would consider "close contacts or friends" was added in 2016. The coefficient of 1.868 indicates that this number is by about 2 persons higher in CO than in AL and the difference is highly significant (CI95 Exp(B)=1.417-2.465, p=0.000). The medians imply that the inhabitants of AL have about 1 neighbour they consider a friend, while this number lies at 3 in CO. 85 percent of the inhabitants in CO have at least one neighbour they consider a friend, while this is the case for only 64 percent of the people in AL.

When it comes to the cohesion between inhabitants on a more general level, more specifically the "sense of belonging", the difference is 0.229 and thus about 6 percentage points (CI95 B=0.051-0.408, p=0.012). The means suggests that both models are at a high level in this regard.

Differences in the area of "support received" are more pronounced. Three different types of support have been considered: (1) "instrumental support" (support in daily life activities, helping someone out), (2) "companionship" (spending time together), (3) "emotional support" (discussing worries, fears and difficult situations). These were tested in two different ways. First, on a 1-5 scale, on which respondents should indicate the degree of support

they received (ranging from "none at all"-"very high"). Second, by indicating the number of neighbours they could draw on in particular situations ("situational support").[35]

Instrumental support comprised different activities such as assisting with shopping or with housework, giving people a ride by car to the shops or the doctor, or sharing knowledge and giving advice. But it also came in more subtle forms such as looking after others by calling them or looking in on them regularly, for instance when someone became sick. Some lower level support, such as watching a neighbour's flat while they were on holidays or minor support acts such as bringing the newspaper by were also of some if lower importance than those listed first.

The likelihood of receiving a high level of instrumental support lies at 3.046 and is thus about 3 times higher in CO than in AL and the difference is highly significant (CI95 Exp(B)=1.898-4.888, p=0.000). The likelihood of strong companionship is located at 1.965 and thus about 2 times as high (CI95 Exp(B)= 1.301-2.967, p=0.001) and that of receiving high emotional support at 1.427 and therefore about 1.4 times as high (CI95 Exp(B)= 0.951-2.140, p=0.086). The last effect, however, is only significant at the 10% level. In combination with the rather low effect size, both of which are of course related, suggests that there is no big difference when it comes to emotional support.

As regards the number of neighbours that the inhabitants can draw on in terms of situational support, we see the following picture: With a coefficient of 1.920, people living in CO have 2 more neighbours available that could offer instrumental support (CI95 Exp(B)= 1.370-2.691, p=0.000) and with 1.839 about 2 more neighbours that offer companionship

[35] On (1): Taking care of an urgent task while you were sick; on (2): Spending time together frequently, although not necessarily in long or intense meetings; on (3): Talking about fears or in mood of depression.

(CI95 Exp(B)= 1.415-2.390, p=0.000). Both effects are highly significant. Emotional support is available from about 1-2 more neighbours—the coefficient is 1.653. The latter effect, however, is only significant at the 5% level (CI95 Exp(B)= 1.019-2.681, p=0.042). The limited size of the effect in combination with the comparatively lower level of significance is mirrored by the fact that the medians do not detect a difference and that it is at zero for both groups. This is likely the case because in AL only 22 percent of the inhabitants receive any emotional support at all; in CO the share is larger, but still only at 40 percent. The differences in medians in the other two categories are clearer. While the central tendency for tenants in AL is to have no neighbour available for instrumental support, people living in CO can draw on 2 and they have about 3 neighbours available to spend time with, while the level in AL is at 1 neighbour.

Overall, the main benefit of CO models seems to lie in instrumental support. Alas, the complementary question introduced in 2016 about the amount of time in which neighbours received instrumental support (and thus other neighbours gave that support) on average, has not been recorded consistently, which is why it has been excluded from the analysis.

Against the detected importance of instrumental support, the illustration of "support given" by the respondents themselves focuses on this very aspect, and in units of time—another question introduced in 2016 but posed and answered consistently in contrast to the one on support received. The analysis detects a coefficient of 1.511 and thus a higher individual investment into granting support of about 0.75 h per week. The effect, however, is merely significant at the 10% level and thus statistically insignificant (CI95 Exp(B)=0.917-2.491, p=0.105). The implication of an effect is supported if we look at the respondents, who grant any support at all. While the level in AL is only at 27 percent, CO lies at a level of 49 percent. As per the engagement variable, the difference between AL and CO is not ground-

ed in the higher intensity of support of those giving support, but the wider base of people offering any support at all. If we look only at those, who offer support, the median in both groups is at a level of about 1 h/week (see MDp).

In the area of "participation", I looked at the amount of time spent attending events of the facility ('internal' and 'external' ones; the latter with participants from outside the facility). The amount is around 45 minutes higher in CO than in AL, the coefficient being 1.472 (CI95 Exp(B)=1.168-1.855, p=0.001). The difference is reflected in the median, which is at 2h/week in CO and only 1h/week in AL. What is more, 85 percent of the inhabitants in CO participate in events, while only 66 percent do so in AL.

5.4.4 Secondary variables

In terms of satisfaction three different aspects have been assessed: (1) "satisfaction with the living situation" (comprising the ability to interact with neighbours and people from outside; events; support received; and overall impression); (2) "general life satisfaction" (single question); (3) "health related quality of life" (index).

With a coefficient of 1.042 the satisfaction with the living situation is about 4 percent higher in CO than in AL (CI95 Exp(B)=1.011-1.075, p=0.008). A look at the means shows that the general level in both models is close to a "high" level on the scale. The difference in general life satisfaction judged by the likelihood of being at a higher level is at 1.786, which means that it is about 1.8 times higher in CO. The effect's significance is only slightly lower than the previous one's (CI95 Exp(B)=1.142-2.792, p=0.011). Although the measure is typically treated as metric in social surveys, since it is assessed on a scale from 0-10, life-satisfaction has been interpreted as an ordinal variable here to match the procedure for all other single-question items used.

Finally, the HQoL assessed in 2016 has been compared to the baseline level assessed in 2015 in order to detect any potential effects on the health of the inhabitants. It is not surprising that no effect has been detected (CI95 Exp(B)= -0.051-0.045, p=0.908). One year seems too little time for any potential health effect of the higher social capital to show. If any such effects exist, they would need to be detected by repeated longitudinal assessment, since it is hardly possible to establish a baseline of health by retrospective statements of the individuals that reaches back before they have moved into the facility.

5.5 Discussion

In the discussion, I highlight the main insights produced by the capital based investigation of community oriented housing and refer to some complementary perspectives and some issues that have not been pursued further. I also outline limitations of the performed research before drawing conclusions.

5.5.1 Main effects produced

We have seen in the results that there are only few areas where no or very small positive effects of CO over AL have been found. There have been no differences in self-efficacy, nor in HQoL. There have been positive yet small effects in trust (6 percentage points) and in satisfaction, whereby the difference in general life-satisfaction (likelihood of a higher level being 1.8 times higher) seemed more relevant than the difference in the satisfaction with the living situation (4 percentage points).

The meaning of the difference in general life satisfaction is partly relativized by the fact that the condensed regression model revealed a high influence of HQoL 2015 as a measure of baseline health on the general life satisfaction in both years (B=3.087; CI95 1.918-4.256, p=0.000). The interdependence of the two satisfaction based measures is not surpris-

ing, yet relevant for interpreting the data. As long as no clear link has been established between the living facility and the inhabitants' health status, the difference in life satisfaction, which seems to rest on the latter for a large part, is to be interpreted with caution.

Another aspect that relativizes the use of life satisfaction to denote effect differences in the present case is based on anecdotal evidence gathered in the questioning of participants. Although it could not be verified systematically, it is still important in this context. Especially older generations, who had grown up in the immediate aftermath of the war or even experienced it first hand, sometimes articulated that they could be nothing but satisfied with their lives at present if compared to the past. This could be one of the factors pitching up the average level of life satisfaction in both groups. The mean and median are at 7.8 and 8 in AL, and at 8.4 and 9 in CO (not listed in Table 9 for reasons of coherence with the other variables analysed by ordinal regression). This compares to an average life satisfaction in Germany of around 7 (OECD, 2016, p. 127). These circumstances, at least partly, call into question the applicability of life satisfaction with this particular target group.

At any rate, the differences in political capital, except for self-efficacy, and in social capital are much more pronounced than those in satisfaction and cultural capital. On the latter, it is to be remarked that in line with my conceptual reasoning in the theory chapter, the number of neighbours that inhabitants fully trust (relational trust), a measure in which the difference is more prominent, may serve as a bridge between the more general level of trust in others and the inhabitants' social capital.

This is also implied by the consistent difference of two people across several measures. In CO there are 2 more neighbours that (1) inhabitants trust, (2) they consider friends, (3) they can draw on for instrumental support, and (4) they spend time with. The available data

is not sufficient to verify whether this difference relates to the same people throughout, and of course these are central tendencies only with variation across individuals, but the observation outlines connections between the assessed measures. In contrast to ‘relational trust’, trust in its ‘unwarranted version’ rather clearly belongs into the separate norm-based category of cultural capital, as it is not enhanced to a similar extent as several social capital components.

As concerns political capital, the higher likelihood of greater problem solving capacity (2.3 times as high in CO as in AL, and highly significant) and a higher degree of engagement for the community of about 0.75 h/week, are most relevant. The latter difference translates into about 3 h/month, in which inhabitants are actively engaged within and for the neighbour community. The 1.8 times greater likelihood of a high individual influence in the CO facilities points into a similar direction. It seems advisable to study collective efficacy with a more fine-tuned repertoire of measures, though. As mentioned previously these perspectives only emerged when they turned out to be separate from the social capital components assessed.

When it comes to social capital, the sense of belonging is stronger in CO, but the difference to AL is only 6% and thus rather small. This suggests that more abstract measures may not be a good indicator of actual social connectivity, and support for that matter, and vice versa.

Differences in emotional support seem to play some role, but a subordinate one. More important are the detected differences in relation to interpersonal interaction. The latter is pronounced at two levels, the individual one (companionship is twice as likely to be at a high level and inhabitants can draw on two more inhabitants in CO than in AL) as well as

in the group or community level. For instance, people in CO participate in events by 0.75 h/week or 3 h/month more than those in AL.

The strongest indication of higher social capital though, is the NWS to neighbours, which is more than 40 percentage points higher in CO than in AL. NWS as a measure takes into account both the quantitative and the qualitative dimension of social connectivity in that it considers the number and frequency of contacts as well as their importance and intimacy.

NWS is not an isolated factor though, but greater connectivity seems to convert into higher social support—the likelihood of receiving instrumental support is 3 times higher in CO and people can draw on two more neighbours in this regard than in AL. The last aspects further gains importance when considering that only 40% of inhabitants in AL receive any instrumental support from their neighbours at all, while this is the case for 73% of tenants in CO.

5.5.2 Complementary and follow-up issues

There is some indication that network strength is not only an issue within CO facilities but also towards the outside. An additional question in 2016 has been on how well respondents thought (1) the facility as an organization and (2) themselves to be connected with the surrounding neighbourhood and quarter. Based on the accounts of inhabitants, the likelihood of stronger connections to the outside was about 3 times higher when it came to links between the facilities' provider and stakeholders in the neighbourhood (organisational level), the coefficient being 3.054 (CI95 Exp(B)=1.887-4.942, p=0.000). It was about 2 times higher on the individual level, referring to links the participating inhabitants reported having to other people in the neighbourhood, the coefficient being 1.915 (CI95 Exp(B)=1.120-3.272, p=0.018).

Another complementary perspective is the time inhabitants spend on performing activities in external organizations, associations for example. Related to only the association that the respondents listed first (in case they mentioned several organisations), the time spent on performing such activities yielded a coefficient of 1.763 and was thus almost 1 h/week higher in CO than in AL (time spent assessed in units of 30 minutes; CI95 Exp(B)=1.216-2.556, p=0.003). The investment of time into such activities, so the regression revealed, is dependent on the inhabitants' predisposition expressed in the control variable "engagement in groups", but the predisposition's effect (Exo(B)=1.347, CI95 Exp(B)=1.162-1.561) was smaller than that of the residential model. 51 percent of the inhabitants in CO reported any such activities at all, while only 37 percent in AL did.

At this point, none of these two indications—neither the connectivity to the quarter nor the activities in associations—provide a clear image of the dynamics between the internal and the external sphere in CO. These dynamics have not been at the centre of this research, but would deserve more attention. It is certainly possible that other factors, such as previously established connections play a bigger role here than the residential model. Unfortunately, a question on whether people had established new connections or whether these were longstanding, and a second on whether activity in the associations had been spurred by the living environment, lacked too many responses for being included in the analysis. And even if they had been answered comprehensively, there had not been enough follow-up questions to uncover the underlying mechanisms. For instance, one would have to qualify whether the activity in groups was more of a passive, participatory and relational character (social capital) or more geared towards exerting influence on the organisation or even the quarter by taking an active role in the association (political capital). These probing issues would need to be given attention, if the analysis were extended in this direction.

Another complementary perspective has been taken in exploring how CO performs under a cost-benefit rationale. The human resource costs involved in operating four of the visited facilities have been analysed in an exemplary fashion to test this. It turned out that CO models have the effects detected in my research while being operated at no higher personnel costs than those found in AL. The calculated costs of running CO models were in fact lower than those found in AL. Since there was considerable variation of costs within the CO model though, with a substantial advantage of provider A over provider C, and also for reasons of confidentiality, the results are not presented. What can be said with some confidence, however, is that given a comparable financial investment, CO models will yield a higher level of capitals.

5.5.3 Un-pursued perspectives

Aside from the follow-up issues that could not be explored in greater detail, some initially intended perspectives have not found their way into the presentation of results. For instance, care support provided by inhabitants to neighbours with a care level only occurred in very few exceptional cases and thus was not a relevant comparative dimension.

Being in closer contact with younger people, be they younger adults, youth or children, also was of no major relevance. A significant difference between CO and AL had been detected in this regard, in so far as the inhabitants of CO were in touch with 2-3 more young people across the above categories than those in AL (Exp(B)=2.595, CI95 Exp(B)=1.068-6.304, p=0.035). However, the share of people to whom this was relevant at all was very low: 9 percent in AL and 17 percent in CO. The intergenerational aspect, at least according to this measure, was therefore not a meaningful analytic dimension.

Finally, NWS was not only assessed in relation to family and friends on one side and neighbours on the other, but also to professionals, both within and outside the facility. Neither of these dimensions, however, was of any relevance for more than 10% of all participants, which is why they have been disregarded in the analysis.

5.5.4 Limitations

The presented research has some limitations, most notably the one that selection bias cannot entirely be ruled out. Both the matching and the regressions have been performed in order to filter out or account for differences in effects that might be caused by confounding factors. It cannot be ruled out with certainty that there is no omitted variable that has an important influence on the formation of capitals. Two particularities of the research restrict the severity of this problem though. First, a large number of control variables have been assessed to account for potential confounding of many sorts. Some of them were explicitly tailored for assessing predispositions, which could be of major relevance for the outcome variables. Especially with research on secondary data it is almost impossible to arrive at the same depth and breadth of relevant control variables. Second, the initial logistic regression revealed that the two groups were very similar to each other to begin with. What is more, the two predisposition variables in which differences were detected, namely unwarranted trust and self-efficacy, were in the areas in which only small differences could be detected on the outcome level, suggesting that they might not have mattered considerably even if no matching had been performed.

There are also some limitations to the generalisability of results. Although community oriented models will share the general characteristics outlined that distinguish them from assisted living, interpretations of the model may differ. It cannot be said with certainty how much of the detected effects derive from very particular actions that may not reflect suffi-

ciently in the general description of the community oriented model. The study would need to be replicated in a similar setting to verify whether or not effects are bound to the specific organisations under investigation here.

5.6 Conclusion

This empirical study has demonstrated the use of considering the formation of capitals when assessing social impact. The capital based approach has helped understand the effects promoted by community oriented models by assessing variables that for the most part do not focus on the benefits arising for the individual inhabitant, but instead are consequently located at the intersection between the individual and the collective level. It has also delivered insights on the interconnections between the assessed variables and enabled an identification of the most relevant ones. By allowing for differentiating effect measures within a coherent framework it has furthermore enhanced insights on how effects are composed, as summarised in the following.

My findings suggest that community oriented living models for the elderly are superior in creating cultural, political and most importantly social capital than assisted living.

Two observations are particularly important with regards to the measured impacts: First, it is remarkable that the measured level of dependent variables in CO is consistently higher than in AL, if by small amounts in some cases. This suggests that there is not a trade-off between the capitals in this context, but rather a mutually reinforcing relationship, which is stronger for some combinations of variables than others. COs seem to be particularly potent in strengthening networks, mutual instrumental support, participation (passive and mainly relational; social capital component) and engagement (active and mainly geared at

influence; political capital). They are less so in increasing satisfaction, the ability of self-help, cultural capital (here: trust), or emotional support.

Second, the detected differences, in particular those in the count data, do not mean that the inhabitants of CO exhibit higher levels of capital in absolute terms. When we look at the levels of people exhibiting a positive value in the studied capitals at all, we find that those levels are rather similar across the two groups. Thus, people in CO do not generally live with their best friends, at least not more or less than people in AL. Instead, it is the wider spread of capitals in CO that the relative difference between the models stems from. This means that fewer people are at zero or low levels in the considered variables. In consequence, we can say that COs have inclusive impacts in that they leave fewer people outside the community of inhabitants. This is of relevance in relation to the political regulation of such models, in particular in view of the problems social isolation in old age may cause on many levels, for instance on relation to health status (Meyer & Schuyler, 2011). It also has important implications for how organisational practice in the field could be reshaped to arrive at greater social impact.

6 Social impact of a school-based intervention

In this chapter I discuss the results of my empirical investigation of the 'RespAct' programme and how it affects the cultural capital held by pupils. Thereby cultural capital refers to the values of individuals, but not exclusively to those that promote educational attainment or distinction but to all that make up the normative equipment of individuals that enables or disables them to position themselves in a productive stance towards society.

RespAct is meant to promote pro-social behaviour in schools, for instance self-esteem and the competence to deal with conflicts. These and related variables are the primary targets as components of cultural capital. The investigation of the impact of the programme has however also looked at social capital (interpersonal connections) and political capital, which here refers to the promotion of individual and collective action and individual influence within the school class.

I have interviewed pupils and teachers of four different school classes that have participated in the programme as to whether it had an impact on the mentioned variables. The chapter shows that effects of RespAct vary for different types of pupils. It finds that RespAct is better at activating passive pupils, who are shy or outsiders in the class, than it is in effectuating change in pupils with high behavioural problems. This has implications for the settings the programme can best serve in. Another factor of use in calibrating the programme, is the finding that the programme's impact in creating social capital was greater than in fostering cultural capital.

6.1 Background

The RespAct programme, aims to help address an important challenge in schools: conflict, destructive behaviour, and educational as well as social frustration of young people. Incidence of violence in schools has been moderately declining on average in Germany (Baier, Pfeiffer, Simson, & Rabold, 2009, 92f.; analysing figures from 1993 to 2007). Nonetheless, social conflicts in schools remains an area of primary interest—not least because of the potential damage caused in later life (Scott, Knapp, Henderson, & Maughan, 2001; Ttofi, Farrington, & Lösel, 2012). Violence prevention programmes or others that work towards improving this situation are a fairly new phenomenon in Germany, but programmes and evaluative studies have become more common (Schick, 2012).

Most of the existing research examining such interventions is marked by two traits: First, almost all performed studies provide direction on whether programmes work, but not on why and how they work. Most existing research is preoccupied with the outcome level and neglects the processes occurring and therefore offers little possibility for understanding organisational performance or enabling organisational learning. Second, existing studies apply various measures from psychological research with little systematisation from the social sciences of the dimensions of social impact achieved.

Both observations can be outlined in more detail in relation to current studies found in the literature. A review of the latter shows that there is a paucity of research simultaneously addressing ‘the if’, ‘the why’ and ‘the how’ of the social impact caused by programmes promoting pro-social behaviour in children and youth. It has been remarked before that qualitative research designs could be a means for doing so. However, such designs that explicitly look at impact are rarely found. Instead qualitative research is used to describe how teachers react to community violence (Maring & Koblinsky, 2013) or relate to chil-

dren (Stuhlman & Pianta, 2002), which factors affect the implementation of evidence-based programmes at the community (Allison, Edmonds, Wilson, Pope, & Farrell, 2011) and at the national level (Spiel, Wagner, & Strohmeier, 2012), how they should look like in the eyes of pupils (Cunningham, Cunningham, Ratcliffe, & Vaillancourt, 2010), or how ethno-cultural context is relevant to prevention programming (Affonso et al., 2010; Helm & Baker, 2011).

Interventions themselves are almost exclusively studied by quantitative methods (see for instance Eisner & Malti, 2012; Wilson & Lipsey, 2007). Even articles dedicated to the question "what works and what does not, for whom and under which conditions?" (Guerra, Boxer, & Cook, 2006), explore this in reference to quantitative studies and disregard the very questions they set out to answer. Although they are treated as a more objective way of capturing effects (see for example Lösel & Stemmler, 2012, p. 218), quantitative designs have specific limits in generating insights. A systematic review of the numerous evaluations of the US programme 'Second Step'[36] and its cross-country adaptions (Schick & Cierpka, 2013) provides evidence in this regard.

Schick and Cierpka (2013) outline that despite mixed evidence, most studies found that children had become more competent at problem solving and anger control, and that empathy had increased. Thus, the programme had altered attitudes. However, studies mostly could not find significant or only weak differences in behaviour. One of the reasons for such findings might be that there is a difference between the reproduction of knowledge (do pupils know about problem solving strategies) and its application (how do they react in

[36] Second Step is one of the most widely spread and studied programmes (Schick (2012)). It is typically integrated into the school curriculum and runs over several school years. It tries to develop several skills through practical exercises, for instance empathy, problem solving skills, or frustration tolerance. It is provided in adapted versions from preschool to middle school levels. The programme's German version is called 'Faustlos' and has also been studied several times (Schick and Cierpka (2003); Schick and Cierpka (2006); Bowi, Ott, and Tress (2008)).

a concrete situation?). For instance there is also a huge difference between knowing about healthy lifestyles and leading a healthy life. There is in fact some rare qualitative evidence that suggests problem solving strategies will not help much if the surrounding environment's belief in appropriate social behaviour is counter to what the strategies demand. If fighting is embraced as a legitimate means in this belief, problem solving strategies or emotional regulation might not be able to prevent or revert such anti-social behaviour (see Farrell et al., 2007; Farrell et al., 2008, 408f.). This calls for more studies that are better contextualised, that is can grasp the intervention and its supposed effects in relation to the individual and its environment or "ecology" (Dahlberg & Butchart, 2005). This is difficult in purely quantitative studies. Another downside of quantitative designs is that they hardly provide insights on which exact elements of the programme have mattered most, not to speak of how they mattered.

While the former issue is one of missing depth in the analysis, there is another issue with the scope of indicators tested and variables applied. Although the programme is explicitly about "[…] the promotion of empathy, impulse control and anger management" (Bowi, Ott, & Tress, 2008, p. 509) a wider range of psycho-social effects are often tested. Bowi et al. themselves for instance refer to a "significant increase of verbal skills" (2008, p. 519; quote translated by the author). Schick and Cierpka (2005) in turn have highlighted effects on anxiety. For both the centrality with regards to the programme's goals is not obvious. Other variables often assessed, that are at least at some distance to the intervention itself, comprise school performance, school participation or knowledge (Wilson & Lipsey, 2007). Given the wide range of effects considered it is of no big surprise that Wilson and Lipsey (2007, S130) in view of the results of their meta-analysis of available evidence conclude that "[…]"[s]chools seeking prevention programs may choose from a range of effective programs with some confidence that whatever they pick will be effective."

With the way I approach the investigation of RespAct's impact I mean to promote analytical depth through using a qualitative design, and a focussed and systemised measurement through relating to relevant capital components as well as the connections between them. My main anchor for the analysis is 'cultural capital'. For its definition I relate to my conceptual work, where an 'extended version' of the concept has been laid out, which comprises the "[…] basic modifiers and determinants of culture embodied in norms and values, and codes of social conduct." As I have further outlined shared modes of social conduct are needed to prevent conflicts and enable productive interaction. Of the components that meet this definition pro-social behaviour, with a particular focus on negating violence, is most important here, but a feeling of self-worth and some related components are relevant too. These are complemented by social capital that points to the relations between people as well as political capital, that is group initiatives or individual agency is promoting issues or solving problems

6.2 Specifics of the programme

RespAct (RA) is performed by CamP Group. The organisation has previously been known as 'Boxgirls', a social entrepreneurial organisation primarily active in youth work in Germany, but also in Africa. The organisation's original aim was to build up the physical, emotional and social strength of girls in challenging social circumstances. RA builds on these experiences and uses material from this original field of activity, but has been extended both in terms goals and target groups. The main rationale behind choosing this programme for study is that the transfer from learnings made in the youth work, notably the use of boxing as a tool for empowerment, and the fact that the organisation has been recognised as innovative in their practices (Ashoka Germany, 2010) suggest that its means for promoting cultural capital are particularly effective.

RA mainly takes place in schools. It has initially been designed for pupils aged 8-14, but has since been adapted and is now also applied to older pupils in higher level schooling, including vocational schools. The programme website states that RA "[…] empowers children and youth to become active in remodeling their surroundings and to protect themselves from violence" (CamP Group, 2015). Among the (further) effects the programme aims to trigger are the following: Anti-violent and fair behaviour; problem solving skills; self-esteem (physical, social and emotional); empathy; mutual respect; group cohesion; individual engagement for the (class) community; collective efficacy (joint engagement in the promotion of issues or the solution of problems).

In order to produce effects on these dimensions RA applies a variety of modules, which are predominantly based on two pillars: (1) Boxing exercises (positioning, distance, moves and punches, pad work); (2) video work (conceptualisation, shooting and cutting of a film on a self-selected subject in the context of the programme's goals). These two are combined with other sport elements (for instance ball games or the testing of physical strength) and exercises related to trust and self-revelation (for instance falling backwards and relying on being caught by others or talking about individual anxieties). All parts of the intervention are connected by recurring reflective discussions about experiences during the programme and its goals. RA is being performed by two trainers at the participating schools and usually takes two full days. On the first day the pupils make some first experiences with boxing, develop ideas for the films and perform discussions on subjects related to RA. On the second day they perform advanced boxing exercises, produce and present the films. The trust games and other sport elements are performed flexibly during the two days.

An intervention of two school days, and thus around 10 hours of exercises and activities, at first seem like a relatively short amount of time dedicated to levering pupils' cultural capi-

tal. However, when compared to other programmes, this is well within the range of a typical duration. The middle-school programme of Second Step (2017) for example comprises 25 lessons of 25 minutes each per grade 6-8, yielding a total learning time of 10.5 hours per school year. RA is thus not less intense than other programmes in terms of the time invested into learning. Other programmes, which are often integrated into the school curriculum, may have the advantage of repetition over time, but RA could be more intense due to the 'experience character' of the condensed format.

6.3 Methods

6.3.1 Setting and participants

My research relates to two classes in vocational schools in a major German city (both for pupils without a degree or without placement for vocational training) as well as to another class in a vocational school (with the same type of pupils) and one class in a regular secondary school, both in another major German city. Participating pupils were between 14-20 years of age. The four interventions have taken place in the spring and summer of 2015 and at the beginning of 2016. They were executed by four staff members of CamP Group in groups of two and in varying composition, three of the trainers were female and one male, all of them in their early to mid-twenties.

Several rationales have guided the selection of participating classes. First, pupils in vocational and secondary schools rather than in primary schools have been chosen, since from what we know their beliefs and behaviour are generally less malleable than that of younger ones (Roberts & DelVecchio, 2000). Effectuating change in individuals will likely be harder to achieve with older target groups, since they will have become more manifested in their beliefs and actions. This has specific implications when combined with the specificities of how pro- or anti-social behaviour develops in young people. I have previously men-

tioned that there is an "age decline of reports of being bullied" (Smith et al., 1999, also Nansel et al., 2001). Thus, when conflicts and other anti-social behaviour are still observed at later ages, even though they might be less frequent than at younger ages, they likely indicate a more profound problem. In consequence, if the programme is found to be effective with older pupils, it is more likely that principles can be adapted to younger age groups than the other way around.

A second but related rationale was that the locations of all the schools should be marked by some level of social tension and conflict, to see whether and how RA is effective under challenging circumstances. All schools are located in districts of the cities that would generally be considered more 'disadvantaged'. Third and in contrast to homogeneity in terms of the challenges presented by the environment, the exact location of the schools, the pupils's age and their educational achievement should be marked by some variance. This decision was taken to be able to identify general impact patterns which manifest independent of context conditions, or to spot variations of impact along these factors. Two of the included classes were marked by low educational motivation of the pupils and high rates of school truancy, while these issues were more moderate in the other two. Fourth, priority was given to having at least as many males as females in the study, boys typically display a higher degree of disruptive behaviour (van Lier, Vuijk, & Crijnen, 2005). However, overall the ratio of male to female pupils should nearly be 1:1 in order to account for effects on both sexes.

The participation of pupils per class is summarised in Table 10. It also contains the share of male pupils, the average age of all pupils (in years) and the share of pupils with at least one parent of non-German origin (in percent).

Table 10 Socio-demographics of participating pupils

		Class 1	Class 2	Class 3	Class 4	Total
RespAct & both surveys	N	8	8	9	21	**46**
	Thereof male	*2*	*3*	*8*	*13*	***26 (57%)***
Interviews	N	7	5	5	6	**23**
	Thereof male	*3*	*2*	*4*	*3*	***12 (52%)***
Immigrant background	Yes (%)	75%	63%	78%	76%	**73%**
Age	Mean	18	19	17	15	

As will be explained in greater detail below a larger number of pupils have participated in the 'before' and 'after' surveys than in the interviews conducted, which is why both are indicated. As mentioned above, somewhat more than half of the participating pupils were male. The average share of pupils with immigrant background (at least one parent of non-German origin) was at about 70 percent and the mean age per group lay between 15 and 19 years.

6.3.2 Analytic strategy

When considering pro-social behaviour as an aspect of cultural capital, and violence in particular, the formation of the capital differs in one central respect from other capitals. Most of the other capitals and their components are created, that is established anew: This may refer to initiating new social relations (social capital), or spurring engagement from a former state of passivity (political capital). Pro-social behaviour instead can require changing attitudes and behaviours that at the beginning are counter to what is needed to enable social productivity. Examples for counter-productive stances taken by individuals pertaining to the RA programme would be: the embracement of violence as a means to settle disputes, or picking on others to promote one's own position. An example in another area could be contempt for minority groups.

The need to revert attitudes suggests that minding the process involved in enhancing the cultural capital components is particularly important. This is the main reason why I have

chosen a qualitative research design, which is mainly based on semi-structured interviews (Patton, 2002; Smith, 1995; see Table 35 in the Appendix for the interview guide). These have been analysed systematically to arrive at reliable results. The qualitative investigation was paired with the quantitative analysis of a paper-based questionnaire to find out if and where the insights gained by qualitative and quantitative means differ. All pupils filled out the same questionnaire twice, once shortly before participating in the programme and a second time about one month after the programme. Teachers also completed a questionnaire, which was however narrower in scope and related to the whole class or groups of pupils rather than to each individual pupil for reasons of practicality. During the post-intervention visit both pupils and teachers have been interviewed (each interview was about 25-45 minutes in length).

In the survey participation by all pupils was sought to, which was partly difficult due to the high rates of absence among some pupils. Overall there were 46 pupils, who participated in the survey twice and in the intervention too. Five to seven pupils per class were selected to participate in the interview guided by the rationale of 'heterogeneity sampling' to arrive at a good representation of different types of pupils and both sexes. The ability to select was restricted by the fact that pupils had to participate on their own accord. 23 pupils have participated in the interviews in total. In addition one to three teachers per class have participated both in the survey and the interviews. The range of interviews has been determined by the point at which data saturation (Guest, Bunce, & Johnson, 2006) had been achieved, which was the case with a small to medium sized sample already, which based on insights from previous research in other fields (Charmaz, 2006, p. 114) could be expected for this thematically focussed study. The main attention in the investigation was clearly given to pupils, who have often been treated as "passive subjects of research" (Bassett, Beagan, Ristovski-Slijepcevic, & Chapman, 2008, p. 120).

The date of the post-intervention data gathering has been chosen based on two rationales: It ought to be long enough after the programme so as to allow time for effects to settle in and for youth to be able to provide a reflective account of those effects in retrospect. At the same time it ought to be close enough to the event as to minimize the influence of potential confounding variables from within and outside school. A period of about one month was chosen to satisfy to both aspects. Since the qualitative investigation allows for tracing processes of change retrospectively, I decided not to recruit a control group.

The quantitative data, once coded electronically, was analysed by using SPSS Statistics 23e. The qualitative analysis demands some more explanation. The investigation of the mechanisms at play in 'RA' shares some commonalities with an (adapted) grounded theory approach (Glaser & Strauss, 1967; Strauss & Corbin, 1998). The coding of interview material has been performed by Atlas.ti software. Due to the advanced state of knowledge on violent behaviour, in the data analysis I have used an 'integrated approach' to code structure, which "[…] applies the principles of inductive reasoning while also employing predetermined code types to guide data analysis and interpretation" (Bradley, Curry, & Devers, 2007, p. 1758). Thus, a naturalistic inquiry approach (Lincoln & Guba, 1985) was combined with a pre-defined 'thematic framework' (Ritchie & Spencer, 2002).

6.3.3 Effect measures

The analysis I performed on the effects of RA relates to nine measurement components across three different sorts of capital, spanning attitudes and actual behaviour: (1) endorsement of aggression factors and (2) aggressive behaviour, (3) cooperative behaviour and (4) consideration for others, (5) self-esteem (all five cultural capital); (6) interpersonal familiarity, (7) group cohesion (both social capital); (8) collective efficacy, (9) individual influence (both political capital).

Table 11 gives an overview of all components tested in the quantitative analysis. The questionnaire rests on established scales from existing research that have been reworked, adapted, combined or shortened to match the present setup effectively. The table gives an overview of the applied scales, including the number of questions, Cronbach's alpha as a measure of instrument reliability (Cronbach & Meehl, 1955) (except for individual influence, which is a single question and no scale), the applied scale intervals and the minimum and maximum scores. The analysis also included a confirmatory factor analysis (Brown, 2006). Only questions with a factor loading of around 0.5 and above on the construct to be measured by the individual scale were maintained.[37]

Table 11 Applied scales

#	Cap.	Category	Quest. (#)	α	Intervals	Labels	Min-Max
1	CC	Endorsement of aggression factors (physical and verbal)	9 (5&4)	0.80	1-4	„Strongly disagree“ „Disagree“ „Agree“ „Strongly agree“	9-36
2		Aggression (physical and verbal)	5 (3&2)	0.70	0-6+	„How many times have you done X in the last 7 days?“	0-30
3		Cooperative behaviour	6	0.73	0-6+	See 2	0-36
4		Consideration for others	7	0.81	1-5	See 1, plus „neither disagree nor agree“	7-35
5		Self-esteem	7	0.82	1-4	See 1	7-28
7/8	SC/PC	Group cohesion/ Collective efficacy	7	0.85	1-5	See 1	5-25
9	PC	Individual influence	1	-	1-5	See 4	1-5

Remarks: CC=Cul. Cap.; SC=Soc. Cap.; PC=Pol. Cap.

Sources for scales (revised before application): see footnote.[38]

[37] Please note that for variable 1 ("endorsement of aggression factors") 5 physical and 6 verbal questions have been tested separately for their factor loadings. All initial "physical questions" have been maintained, 2 verbal ones had to be dismissed for poor loadings (<0.4). The remaining 9 questions were combined to a single scale on aggression with an alpha of 0.8 as reported above. This complies with the application in van Schoiack-Edstrom, Frey, and Beland (2002), if with a reduced number of questions.

[38] 1: van Schoiack-Edstrom, et al. (2002)(derived from Slaby and Guerra (1988); Erdley and Asher (1998); Crick and Bigbee (1998); Crick and Grotpeter (1995)); 2: Dahlberg, Toal, Swahn, and Behrens (2005, p. 168) (in relation to Orpinas (1993, p. 74); Orpinas and Frankowski (2001, 64f.)); 3: Dahlberg, et al. (2005, p. 168) (in relation to Orpinas (1993, p. 74); Orpinas and Frankowski (2001, 64f.)); 4: Dahlberg, et al. (2005, 181f.) (in relation to Orpinas (1993); modified in Bosworth, Espelage, Du Bay, Dahlberg, and Daytner (1996, p. 73)); 5: Dahlberg, et al. (2005, 134ff.) (in relation to Weinberger and Schwartz (1990));

For testing the reliability of the scales all 109 completed questionnaires have been used (total number of individual questionnaires from the pre- and post-survey). For the analysis of effects the questionnaires of only 46 pupils have been used—the questionnaires of those pupils who participated in both the pre- and post-survey and the intervention itself. The scale intervals have been chosen based on conventions from previous research. Sources of the scales are indicated in the footnote to the table.

The composition of measurement components has been derived in two ways: First, from RA's goals and through discussions with staff of the organisation performing the intervention. Second, based on the insights I gained when I performed the qualitative research and during the analysis of the generated data. Both stages led to some revisions in the initial composition.

The qualitative interview guide and the questionnaire have been developed in parallel. The indicators applied are similar for both parts of the investigation, although the qualitative part is marked by greater depth and width of the issues covered. What is more, in the qualitative analysis the measures on the level of attitudes and those on the level of behaviour have been combined (for instance (1) and (2) as well as (3) and (4) from above). Some initial conception contained in the quantitative scales have been revised through the qualitative investigation as will be seen in the results and discussion sub-chapters.

7/8: Dahlberg, et al. (2005, p. 147) (in relation to Rosenberg (1989); Gray-Little, Williams, and Hancock (1997)); Dahlberg, et al. (2005, 324f., 46, 343) (in relation to Sampson, Raudenbush, and Earls (1997); Barnes, Sampson, Kindlon, and Reiss (1997) and Vessels (1998) and Perkins, Florin, Rich, Wandersman, and Chavis (1990)). #6 is missing on purpose in the list, since another SC variable emerged only in the qualitative investigation (see discussion). 7/8 group cohesion and collective efficacy were contained in one scale but emerged as separate variables in the qualitative research. See also Table 33-Table 34 in the Appendix for more details.

6.4 Results

I start with the presentation of quantitative results, which was mainly meant as a scoping exercise as to the trends in perceived effects. Quantitative questioning would need to be expanded to a wider range of participants to potentially increase the statistical significance of the results. Their practical significance is critically assessed in the presentation. The core of the analysis instead lies in the in-depth discussion of change processes evoked by RA and the link between the enhanced capitals and the programme components responsible for the shift.

6.4.1 Quantitative scoping

In order to arrive at effect sizes mean comparisons have been performed. These are illustrated in Table 12, which also gives the effectuated change in percent of the 'mean pre-intervention'. These are calculated by relating the mean difference detected to the mean before the intervention reported as the aggregated and averaged individual scores on the scales reported before.

All effects point into the direction that is intended by the programme. Endorsement of aggression factors and actual aggressive behaviour have both declined (-0.41; -1.00, the median is used in the latter case due to the distribution of the variable), while constructive behaviour, consideration for others, self-esteem, group cohesion/collective efficacy and individual influence have increased (+0.65; +0.43; +0.93; +0.43; +0.04).

Table 12 Quantitative effects

		N	Mean/ MD POST	Mean/ MD PRE	Mean/ MD difference	95% CI Min-Max	Diff. as % of Mean/ MD PRE
1	Endorsement of aggr. factors	46	18.37	18.78	-0,41 (0,43)	-1,29-0,46	-2%
2	Aggressive behaviour+	45	2.00	3.00	-1.00	-	-33%
3	Constructive behaviour	46	8.15	7.50	0,65 (1,01)	-1,38-2,68	9%
4	Consideration for others	46	24.57	24.13	0,43 (0,58)	-0,73-1,60	2%
5	Self-esteem	46	23.46	22.52	0,93* (0,45)	0,04-1,83	4%
7/8	Social cohesion/ Collective efficacy	46	19.52	19.09	0,43 (0,38)	-0,34-1,21	2%
9	Individ. influence	46	3.29	3.24	0,04 (0,13)	-0,22-0,31	1%

Remarks: Standard errors in brackets; CI=Confidence Intervall; +=Data not normally distributed, median used (=MD); *=p<0.05

However, all effects are small. If we look at the changes in relation to the states before the intervention, the largest change is -33% for actual aggression and the smallest one +1% for individual influence. However, while the effect on aggressive behaviour seems to be big, it really marks a decline by one "aggressive behaviour" (such as insulting someone, getting into a fight etc.) per week. What is more, only the increase in self-esteem by 4% is statistically significant at the 5% level. All other effects are not statistically significant, which also becomes apparent in the wide CIs at 95%.[39] All measures except for self-esteem and although generally tending more strongly into the expected direction, span the positive and negative part of the spectrum.[40]

I have tested whether separate analyses by socio-demographic characteristics, such as sex, age, cultural background or school class membership would lead to very different results. This was not the case, neither in terms of effect sizes, nor in terms of statistical signifi-

[39] Data on actual aggression are not normally distributed but positively skewed, which is why no CI has been indicated for the effect size on this scale.
[40] It is to be remarked that this is certainly partly due to the small sample size.

cance. Teachers also responded to a similar questionnaire, focussing on group rather than individual behaviour for reasons of practicality though. Only 9 teachers participated in that survey across the four classes, which is why I do not report results explicitly. The only thing that can be reasonably said is that teachers' accounts tended into the same direction as those of the pupils.

Positive, but rather small effect sizes and issues with statistical significance are not uncommon (Schick & Cierpka, 2013; Wilson & Lipsey, 2007). This is even the case with other studies that draw on a considerably higher number of participants—the low number of realised surveys certainly impedes the chance of results being significant. What does this suggest, in general but with regard to the intervention under study here in particular? One possibility is that there are no or only few effects that matter practically. Another possibility is that the intervention affects different types of individuals (personalities) in different ways and these effects might level each other out on the surface, leaving only the general, positive trace identified in the quantitative scoping.

Another issue is that a pre-drafted quantitative design might miss effects of relevance that had not been at the centre of attention from the start. Or it might start off with a repertoire of measures so broad that finding some positive effect is likely. Whether or not that effect is material, that is relevant for the target group though, or whether it is tied strongly to the very goals of the intervention can be less clear. It could also be difficult to say how that effect interacts with others. Most of these issues should be probed by the application of the initially sketched strategy. The insights generated thereby are presented in the following sections.

6.4.2 Qualitative analysis

The following analysis will show that the effects of the RA programme depend strongly on the personality traits of the participating pupils rather than clear-cut categories the pupils could be categorised by and are mainly found in areas that were not initially expected, which will be outlined more specifically in the discussion section. Cause-effect relations and the different mechanisms in building up the different kinds of capital for specific pupils could only be revealed in the qualitative analysis and categorised and contextualised effectively by the capitals framework underlying the analysis. Another value added of the analysis is the regard given to the context factors shaping individual behaviour, that is the individuals position within the wider social context.

6.4.2.1 Pupils' own accounts

To begin with I report pupils own accounts, then I add teachers perspectives to finally broaden up the perspective towards context factors that have emerged as important in the analysis and the mechanisms by which programme components spurred the detected impacts.

6.4.2.1.1 General assessment of the programme

All pupils, without exception, reported that they had fun during the programme. Many assessed it as a "completely new experience" or referred to the group work as a special opportunity (1, f).[41] More important for the question of impact though was that the latter two points made some pupils show degrees of motivation and engagement that they lacked otherwise. One female pupil said she was impressed with how well the boys, otherwise

[41] All direct quotes have been translated. To be able to differentiate statements, whether directly or indirectly quoted, each participating pupil has been allocated a number reported in brackets along with an indication of the person's sex (f=female, m=male). Pupils have not been numbered in the order of interviews. Thus, numbers are not indicative of group membership.

often raucous and disinterested, participated in the programme (20, f). One of the main benefits of the programme, according to both pupils and teachers, was that it was performed by persons from outside school, who could relate to the pupils' views and were impartial so that pupils could connect to them more strongly than to teachers. All except for two pupils would have appreciated more time for the project, since they would have liked to engage more deeply on both the boxing and the video work. 21 of all interviewed pupils said they would recommend the programme to others. However, the positive assessment and satisfaction with the programme shall not hide the fact that its impacts on pupils had been very varied.

6.4.2.1.2 Cultural capital

The first components when it comes to cultural capital is (non-)aggression. The effectiveness of the programme in altering aggression is one of the areas where the greatest variance could be observed. According to self-accounts as well as the assessment of others, three pupils who had usually been among those causing trouble in the class, had decreased their anti-social behaviour considerably (19, 23, and one who was not interviewed). A pupil, who himself was considered to seek trouble and had benefitted from the programme, said this about one of his peers:

> *XY was the worst of us all, but he has become a completely different person. He used to do bad things and now he has become very calm. He still is somewhat disrespectful towards teachers, but only when he's being bossed around all the time. (19, m; about another pupil)*

All three identified pupils independently reported that they avoided conflicts and were not actively seeking them out. Furthermore, they had become more interested in finding constructive solutions to conflicts. However, some confounding factors come in here. At least one of said pupils had been participating in an anti-aggression programme for about half a year, performed by an ex-convict at the school, in which he and others talked about their

daily problems, how they could avoid conflicts and how they could resolve disputes. Another of the three pupils had had to change schools twice already due to misconduct and was threatened by permanent expulsion from school in case of re-emerging problems. This was certainly a factor encouraging behavioural change independent of the intervention. With all three, RA seems to have had at least a supportive function. However, it was hard to make out the exact share of change caused by the programme. All three pupils had been 14-16 years old and from the same class.

In contrast to the above there had been four male pupils, whose negative inclinations had not been altered. One of them referred to peer-group pressure towards behaving anti-socially:

> *Well, respect for others, in the back of my mind I know that I should show it. […] But when the moment comes and everyone says: "You're a coward." When this moment comes, you just go ahead and jump [into trouble]. (18, m)*

When considering those that hadn't stood out by a particular degree of negative behaviour before the programme, about 5-6 boys and girls gave clear accounts of positive change. Both their physical and verbal aggression had decreased.

At least half of the pupils had previously experienced severe conflicts or other forms of violence, both physical and verbal. In particular the pupils who had once been the victims of violence found the contents of RA helpful. They had experienced an increase in feeling more self-assured and better able to resolve situations of conflict. Overall girls seemed to have benefitted more strongly than boys in this regard. Boys, even those who did not have trouble in maintaining a positive attitude towards others, sometimes felt peer-group pressure towards engaging in conflicts (13, m), confirming the quote above.

The second component in relation to cultural capital is ‘consideration for others’ and related ‘constructive behaviour’. About seven pupils reported how as a result of RA they tried to motivate their fellow pupils to engage more strongly with school, helped others resolve personal challenges or verbal and physical conflicts. One pupil with a learning disability (11, m) spoke about a situation in which he defended his friend in their free-time when other youths tried to mug the friend. Another pupil described how he and his friends were more inclined to address problem situations constructively, either through own initiative or by involving teachers (19, m). One female pupil reported how she had offered help to a fellow female pupil who was being bullied (7, f). The same girl described how she had succeeded in transferring what she had learned during the programme to her difficult family environment: in particular she said she had resorted to giving her mother the opportunity to regain her trust, although she had been betrayed by her mother repeatedly. Another female pupil spoke about positive effects on her conduct towards her siblings. In contrast to all the latter, another 13 pupils reported cases in which they had actively helped others, but pointed out that their inclination to do so had not been enhanced through the programme.

The third and final component within cultural capital is self-esteem. About 8 pupils, thereof 5 female and 3 male, told me that their self-esteem had increased. The majority of the effect was ascribed to the video work, during which the pupils had to overcome shyness and individual anxiety of acting in front of others. A higher awareness of their own achievement and abilities was the result. For three of them the boxing had been more important. One male pupil had a positive experience through acting as a co-trainer based on his previous boxing experience. One girl had increased her self-esteem through seeing how well she could keep up with the boys. Another girl seems to have benefitted from the boxing and the video work to the same extent. She talked about being self-assured and more relaxed, with follow-up effects for conflict-free behaviour and an increase in her feeling of

self-worth. This is what she responded to my question on whether she had felt more self-confident as a result of the programme:

> *Yes, and very much so. I have spoken up when someone insulted me. I was able to assert myself. I told myself that I was no longer the little girl from before, […] who never defended herself. I've become this new girl. Now I am able to defend myself. And when something unnerves me, I tell myself: „Calm down. Just leave. Let them talk. Don't give a shit about it." (10, f)*

Repeatedly I encountered reports about increased pro-activeness of shy and passive pupils. This seemed to hold for pupils independent of the reasons for their prior passivity, which lay for example in linguistic barriers, learning difficulties or simply the individuals' character. Thus, in terms of self-esteem the programme has helped promote a feeling of self-worth and overcoming passivity.

6.4.2.1.3 Social capital

The greatest effects of the programme were found in the area of group cohesion. They ranged from affecting the relation between few individuals only, to broader group level effects. It is remarkable that stronger cohesion spanned very different types of individuals. The programme especially brought together those that had little to do with each other before. At least 11 pupils reported positive effects in this regard. Here is an example, where the effects included the whole class:

> *We trusted each other more, we talked much more to each other than before. We all sat together and laughed. We had never done this before. (5, f)*

In some cases the closer contacts have been marked by a concomitant increase in empathy and mutual understanding, but this only applied to a minority of pupils (about 3-4). The univocal increase in social capital and the small variation across individuals and classes make a more detailed discussion redundant despite the centrality of the effect.

6.4.2.1.4 Political capital

In contrast to the effects on group cohesion, there was little increase in collective efficacy or problem solving ability at the class level. Neither was there a perceptible increase in the influence of individual pupils on collective decision making. Either collective-efficacy had been at a rather high level even before the programme, or pupils remained dependent on the moderating influence of teachers, as described in this pupil's account:

> *To be honest, it is sometimes a little hard for us to get started. […] It is primarily the teachers who,… how should I put it, have the overview of would we need to do, what we need to say, how we need to get started. The teachers really provide guidance in this regard. (17, m)*

Only rarely did the pupils mention that they dared more strongly bring in their own points of view into group discussions. The group of pupils, for whom this had been the case, had some overlap with those who had reported a higher level of self-esteem, that is pupils that had been markedly passive before. One specific case was remarkable, where increased self-esteem and self-initiative went hand-in-hand for a female pupil. She said after the programme she felt confident in pitching her own ideas to the class, while she had never dared do it before and always got goose-bumps whenever she had to say anything in front of others.

6.4.2.2 Teachers' assessments

Nine teachers have been interviewed in total, all female. Their assessments were quite diverse. About three teachers, each of them from a different class, reported very few to no changes in the pupils' behaviour. Six teachers instead explicitly reported positive effects, mainly in two dimensions: (1) increased group cohesion and a higher familiarity between pupils; (2) increased self-esteem, in particular among passive pupils. Teachers accounts are thus fairly congruent with the effects that emerged as central in the interviews with pupils.

Some further aspects were mentioned: In one specific class, two teachers said independently that verbal aggression had slightly decreased. Physical aggression was unaltered from the teachers' point of view. Greater collective efficacy, or individual engagement and assertiveness have neither been perceived by teachers. Beyond the latter dimensions eight teachers stressed that RA offered a sense of success and achievement, which the pupils otherwise rarely felt in school. Teachers highlighted the creative engagement demanded by the programme and that pupils could test out their limits. Only one teacher felt her pupils had already been too old for the programme. This was however mainly tied to the circumstance that her class had already participated in a variety of programmes promoting prosocial behaviour. Further positive changes according to that teacher were hard to achieve with her pupils, since they had experienced major learnings before—a view which cannot be entirely supported based on the pupils' own accounts.

The teachers' and pupils' assessment was identical in that the boxing and the video work were equally central to the project and that the project's duration should have been extended. In particular the teachers who had not perceived strong effects, tied this view to the hope that effects would be increased if the intervention were prolonged or repeated. Eight teachers said they would or already had recommended RA to others, or were already planning to participate with another class.

6.4.3 Mechanisms and moderating context factors

In order to arrive at a rounded picture of the causal process at work with regard to the RA programme, the identified effects have to be contextualised. Two levels need to be addressed when doing so: (1) the intervention and its components itself, more explicitly the question of how specific elements have produced the detected effects; and (2) the envi-

ronment in which the intervention has been performed and confounding factors the pupils are subject to. I will start with discussing the latter.

6.4.3.1 Confounding factors

My analysis confirmed the 'age decline' in conflicts in school referred to earlier. In particular pupils of the vocational schools reported a significant decrease of violent conflicts in their surrounding environment, once they had entered this higher level of schooling. But several pupils in the middle school class said similar things, referring to a higher level of conflict at younger school age. It has been mentioned that assessing changes in attitudes and behaviours can be complicated when a 'natural mechanism' into that direction is at work.

What comes in addition to this general effect is that the participating classes themselves were marked by a fairly low level of serious internal conflict and quite strong group cohesion even before participating in the programme. Although such a situation is generally desirable, this complicates the detection of improvements. In addition, in one class there had been a major confounding factor, namely a joint school trip that had taken place in between RA and the post-assessment, which came with some positive effects for group cohesion. In the interviews and their analysis I tried to bracket this factor out by explicitly identifying effects that had occurred through the trip. Anything like this has obviously been impossible to do within the quantitative analysis.

An even more important confounder than singular events was the pupils' family background, with both positive and negative influences on pupils' cultural capital. At least five pupils explicitly mentioned that their family had educated them to act in a pro-social way and that therefore RA could only back-up their pre-existing positive inclinations.

On the one hand, there were at least three male pupils, for whom negative family role models, in particular their fathers, played a major role in making them unreceptive for peaceful resolutions of conflict. One pupil talked about how he and his friends are constantly pushed towards violent behaviour by parents' expectations:

> *The goal [of RA] certainly is that you would never fight again or be involved in conflicts. […] But we don't accept that. For instance if someone insults us [several strong swear words as examples of such insults], we'd fight him! And it is always going to stay that way, no matter what anyone is trying to do about it. […] Because, if my father knew I was being insulted…; he always says: "You don't have to hit him right away." Well, alright, but if I didn't react angrily or with some aggression, I am sure my parents would think: "Why doesn't he do anything about it?" (4, m)*

He and another pupil mentioned charges in court against their fathers. It is unclear if and how single programmes can counteract such negative family influences. The respective pupils themselves provided some prompts of what would be needed to impress them: They mentioned the need for more 'authentic trainers' with own experience of violence. They said, they would be more likely to follow the advice of police men, youth workers, ex-convicts or experienced martial artists.

6.4.3.2 Effect mechanisms of the programme

As introduced initially the RA programme is based on three different pillars: (1) boxing; (2) video work; (3) trust exercises and discussions. Each of those has had specific effects on the pupils, which I describe below.

6.4.3.2.1 Boxing

The boxing exercises were assessed positively by over three quarters of the pupils. Meanwhile about one quarter didn't see any particular benefit coming out of boxing and would have been just as fine with practicing another sport, for instance football. Some pupils weren't quite sure about the meaning of the boxing within the programme beyond the physical exercise.

One girl in particular found that boxing was more appropriate for boys than for girls. In contrast to this I encountered explicit accounts of 6 girls and 5 boys, who were struck by the girls' performance in the boxing. Here is how one boy reacted when I asked why he was impressed with one girl he had mentioned when talking about his general assessment of the programme:

> *Because she had a good punch. I've seen it and found it quite interesting to learn that many of the girls were really up to it. (18, m).*

Some girls in reaction to the experience during the programme said they had decided to continue boxing. Their intention could however not be verified.

Almost univocally pupils saw boxing as a means for releasing pent-up emotions and physical energy and for reducing frustration, which they would "otherwise swallow and feel bad about" (8, m). Several pupils expressed the constant need for a way of reducing negative emotions and that regular boxing exercises could serve this purpose. One boy reported feeling "more relaxed and peaceful" (19, m) after the experience. Unexpectedly about 6 pupils, male and female, had had previous boxing experience and some of them were happy to share their skills with their fellow pupils. In particular two males (15 and 23) gave some impressive accounts of what they believed boxing could achieve. Here is one of them:

> *When you've just started boxing, you think: „Wow, I am so cool, I can box." You are always after trouble because you feel strong. But when you've been doing it for a while, you calm down. You don't want to get into any trouble, not with the police nor anyone. (23, m)*

According to this pupil, after an initial phase of potentially increasing aggression, boxing may weaken violent inclinations significantly.

6.4.3.2.2 Video work

The video work was rated positively by all pupils without exception, although about 5-6 did so for the mere fun of it rather than for any actual effects perceived. Some of the statements in the positions pointing in the latter direction remained ambivalent though. One pupil had first reported that the video work had not increased respectful behaviour. At the same time he said:

> *Of course it made sense. You can't solve everything by fighting. There are so many conflict situations per day. If we started a fight for each of them, we'd be no better than cavemen. (4, m)*

This underlines that it is hard to extract the more subtle effects triggered by the programme, that is those which enable, stabilise or confirm a positive inclination but might not be 'grand' enough to detected as a major change. However, several other pupils reported such major change explicitly. Some described how important it was for them to feel what it is like to take on different roles in a situation of bullying, in order to learn how to deal with it (8, m). Others reported that the video work had provided learnings on how to support others in daily life (17, m). The need to come out of their comfort zone and stand in front of the camera was also mentioned as a positive aspect. The same goes for the earnestness that was demanded from pupils by the situation and that they managed to uphold, while they lacked this skill in many other situations in school.

6.4.3.2.3 Trust exercises and discussions

The trust exercises were assessed as childish by a minority of 2-3 male pupils, while many others rated them as positive for building a closer relationship to or a greater understanding for others, especially to those they had not been in direct contact with previously. One female pupil said she had gained trust in two male pupils that she would not have trusted

before (7, f). Another, very memorable, report refers to the revelation of personal fears. It was given by a girl who was very impressed by the report of one of her fellow pupils:

> *XY, the tall one, […] he was afraid of going to school when it's still dark, since he now has to walk through a tunnel, and although he's so tall, he said he was afraid of doing it. […] I've never heard XY say: "I feel anxious on my way to school." What I mean is you got to know that person better. (22, f)*

She went on to say that this experience had strengthened her self-image, because she had seen that others have moments of uncertainty at times. The fact that the boy was among the strongest pupils in the class and among those most often involved in trouble, made this insight the more impressive. Unfortunately not all complementary elements of the programme could be analysed in a systematic fashion, since some of them had only been applied in some of the classes. This is also why trust is not explicitly mentioned as a category in the analysis.

6.5 Discussion

Overall it is clear that the elements that RA is trying to promote were highly relevant to the pupils, independent of their age, school type or location. But it has also become evident that the programme's impacts vary strongly across different personalities and that the main impacts were found in other areas than initially expected. For instance social capital has been enhanced significantly while cultural capital was only weakly influenced as regards aggression and cooperative behaviour, while there was a considerable effect on self-esteem. Political capital, that is individual engagement or group efficacy, were also only (very) slightly enhanced. Although the variation of effects makes it difficult to provide a condensed overview, the following paragraphs give an aggregated account of the effects achieved. In contrast to my other two empirical cases the display in a radar chart marks the evolvement of themes in the discussion and an indication of effects rather than an explicit testing, which is why its composition is posited within the discussion section.

6.5.1 Main effects produced and missed

The results have shown the use of differentiating between different capital components to denote the range of relevant effects. The application of the capital based approach in combination with a qualitative research design has helped extend and systemise the initial quantitative conception in several regards along the insights gained in the interviews. 'Interpersonal familiarity' for instance has only emerged in the qualitative analysis and its importance to the intervention had not been anticipated. Based on prior research, group cohesion and collective efficacy questions were contained in a composed measure within the quantitative research. However, the qualitative analysis together with the capital based reasoning suggested that these were two different effects. They helped untangle the two by highlighting that one referred merely to the existence of social relations and their strength (group cohesion) and the other to joint solving of problems and initiation of activities (collective efficacy). For its engagement character the latter was subsumed under political capital. The same goes for 'individual influence', which originally was seen as being part of collective efficacy, but then spun-off from it for poor conceptual fit. It now appears as a single question in the quantitative testing and as an element in the qualitative analysis.

Beyond these benefits the capital based reasoning has introduced an important perspective in assessing effects of value with regard to the RA programme. Although almost all pupils reported they were satisfied with the programme and would have recommended it to others, specific individuals benefitted from the programme much more than others, with important implications for the adaption of intervention mechanisms of the programme or a recalibration of its goals. The programme has certainly enhanced some pupils' capabilities, for instance in relation to making them better able to deal with conflicts. However, some enabling effects were likely too subtle to translate into categories that would be directly relevant to improving the pupils' capacity to act. Self-esteem for instance is a prerequisite

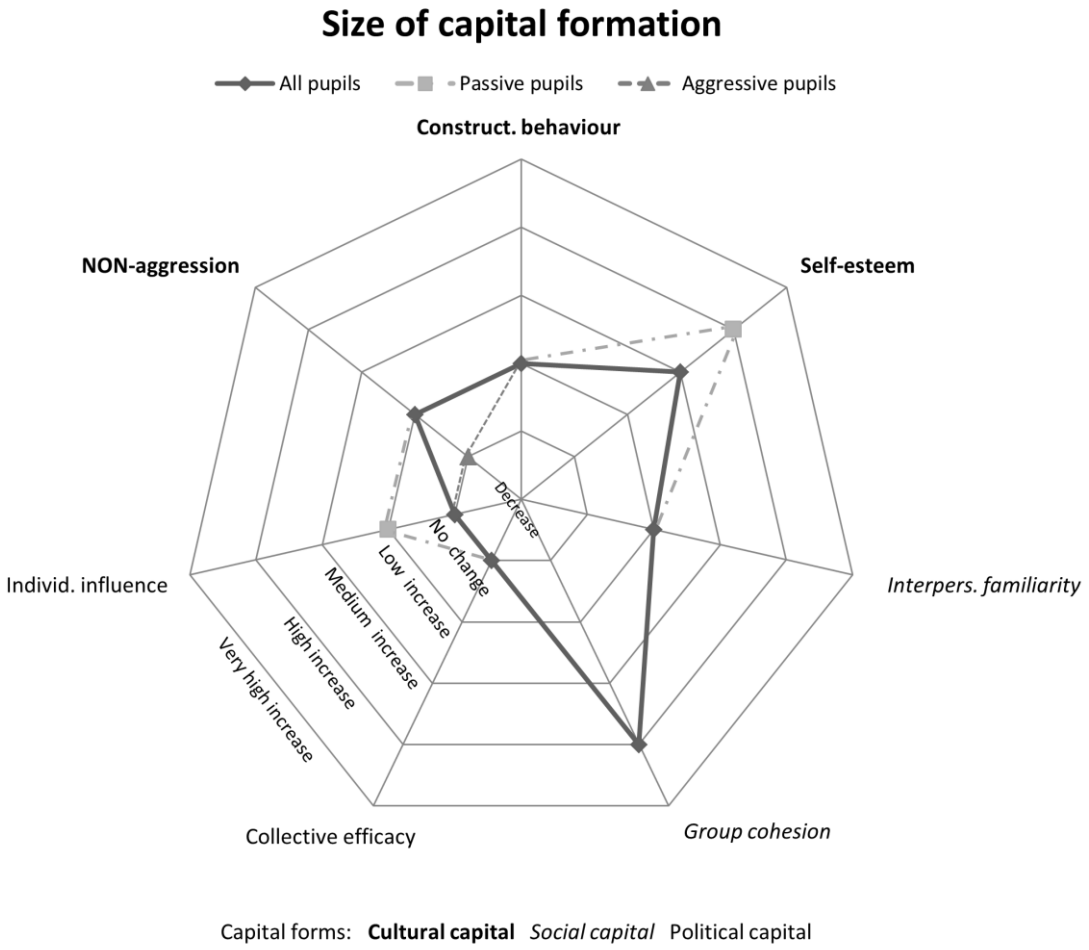
for doing so, but some shifts therein were expressed by pupils with some uncertainty and thus hard to grasp. It is difficult to assess whether they will be meaningful for the future achievement, but they have certainly mattered in the context of the specific class community. Some of the variables in fact (cohesion or interpersonal familiarity) were only relevant in relation to the group context and less so for the individual if looked at in a more isolated fashion, for instance in relation to her utility arising from getting to know some of her peers better.

Figure 5 below illustrates the enhancement of capital forms by RA. It relates to the nine dimensions introduced at the beginning, but aggregates measures of attitudes and behaviour where appropriate ((1)&(2) and (3)&(4)). It also extends the conception in the ways just mentioned. The radar chart marks the capital forms the different measures belong to and distinguishes change in capitals by five regular intervals ranging from 'very high increase' to 'no change' (equal to the level of capitals before the intervention). The midpoint of the graph instead marks a potential decrease, which is however irrelevant in the present case. I differentiate between types of pupils, marked separately, where these have emerged as relevant in the analysis.

The analysis has shown that an increase in self-esteem ('medium increase') was one of the two most strongly perceived effects. Passive and shy pupils have benefitted the most from this effect ('high increase'). An increase in group cohesion was the other main effect affecting the majority of pupils in similar ways ('high increase'), which did not always go along with a greater interpersonal familiarity among the pupils ('low increase'). The higher cohesion can probably unfold even more strongly at the beginning of a school year, when the pupils still know each other the least. Constructive behaviour and caring for others has been quite pronounced by several pupils, but no specific type of pupils could be made out

who were more receptive for it, and not been perceived at all by others (therefore 'low increase' overall). Effects on collective efficacy and the influence of individuals on group decision making have been very weak in general (therefore 'no change'), whereby passive pupils increased their individual influence more than others ('low increase').

Figure 5 Capital formation by RA



Effects regarding aggression have been varied most strongly and need to be explained more explicitly. Pupils with violent inclinations, in particular if male and at higher age levels, almost haven't changed in aggression at all ('no change'). About two to three younger pupils with similar inclinations to the latter, have experienced a reduction in violent behaviour. Due to confounding factors, such as a change of schools or participation in parallel interventions promoting pro-social behaviour, a statement about the exact effects of RA in

this regard is difficult. When asked about it, a slight majority of pupils assessed the potential impact of the programme towards violent pupils as limited. Some rather moderately tempered pupils have however also benefitted from the programme in terms of reducing aggression (therefore 'low increase' on *non*-aggression).

In any case a longer duration, a more drastic experience and more 'authentic' instructors were named as ways to impress pupils with a high degree of behavioural difficulties. The programme would have to be developed further in this regard or partner with other organisations. It would also have to more strongly include family backgrounds, which represent the most potent influence on young people's behaviour. In some cases the intervention has shown that it can spur the transfer of learnings into family contexts.

Both the boxing and the video work were found to be indispensable, although or because they have had different effects for different individuals. Most pupils have described boxing as a means of releasing physical energy and reducing frustration. Boxing also caused a slight increase in physical self-esteem and self-confidence in conflict situations. Besides, it was very effective in strengthening the self-image of girls just as the one that boys had of their abilities. Although these had not been in focus during this investigation, the programme seemed to unfold effects on gender roles.

The video work helped pupils to get to know themselves and others in new ways. In particular passive pupils have gained a higher level of emotional self-esteem (self-worthiness) and social self-esteem (interacting with others). Provided the films dealt with conflicts (bullying, conflict resolution, respectful behaviour), most pupils have said they had been sensitized more strongly to these subjects than before. However, those pupils that chose only a remotely related subject (smoking, a portrait of their city quarter etc.) have experi-

enced these effects to a lesser extent. It thus seems recommendable to restrict the choice of subjects. It also was not clear to all pupils what boxing had to do with the other issues. These connections would have to be stressed more strongly.

The teachers' assessment was largely in line with that of pupils, although the assessment was not given in the same breadth and depth. One aspect that only came up in the teacher interviews was the value ascribed to the new mode of learning as well as the feeling of individual and group achievement the programme offered. According to the teachers these elements are usually lacking in everyday school practice.

6.5.2 Limitations

While this study has laid the foundation for future research on the effects of this specific intervention and others, it has also been marked by some limitations. For instance, it has been partly difficult to capture changes, since the general 'class climate' has been largely positive to begin with. However, in particular one of the classes was affected by serious conflicts at the school level between classes and pupils of different ages. Which effects RA could unfold towards the wider environment could not be assessed by the present design, not least because it would likely have needed more classes to participate in the programme. RA seems to have stronger effects on the aggression of pupils aged about 14-16 years than on those who are older, but it is not clear whether and how the effects revealed can be extended towards even younger pupils, since primary schools were not in the focus of this analysis. An extension of the impact assessment in that direction is however promising with regard to the illustrated trend.

6.6 Conclusion

The analysis I performed shows that in fact RA's impact as to the creation of cultural capital is varied. This applies in particular to one of its presumed main component, the promotion of behaviour that is free of violence. In turn it fosters another dimension related to the youths' equipment with cultural capital: the programme is effective in increasing self-esteem, in particular among shy and passive pupils. A positive cultural side-effect is that it dissolves gender prejudices (another cultural capital factor). Furthermore it strongly promotes group cohesion as a component of social capital , while its impact is limited as regards the promotion of collective efficacy (for instance joint problem solving) and individuals' influence on group decisions, both being elements of political capital. In addition to its systemising of impacts, the capital based approach together with the flexible research design has helped derive recommendations on how to improve organisational performance by establishing links between capital components and the core pillars of the programme.

7 Social impact of a political web-platform

This chapter reports the results of the study that focussed on the political capital created by abgeordnetenwatch.de (AW).

AW is a web-platform that aims to promote a dialogue between citizens and politicians, increase political transparency, and limit the influence of lobbyists on politics. It does so by a spectrum of activities ranging from offering a public forum for political exchange, to documenting politician's voting behaviour, to initiating legal action against misconduct. 745 citizens and 255 politicians participated in an online survey to express their views on the effects of AW. Some comparisons have been drawn to data from the European Social Survey (ESS), both as regards traits of AW users against the general population and in assessing AW's effects.

My research finds that AW users have higher levels of literacy and engagement than ESS respondents who are representative of the general population, and only part of the differences are due to different inclinations of the two groups. AW users are also politically more conscious (cultural capital) than the population. AW users express much more positive views than politicians about the platform's effects in terms of information, awareness raising, exposure of politicians to the public and sensitisation for lobbyism. Given the positive assessment of impacts by citizens it is striking that politicians are rather indifferent towards the platform's effects. This applies even more so as activity on AW seems a replacement for rather than a way into political engagement within formal organisations.

7.1 Background

Studying political capital is a key to understanding actor participation and influence within political systems, both of which are central to democratic functioning and thus wider social productivity. Existing research has pushed the boundaries in understanding the moderating factors of political capital as people being politically engaged within the democratic system. On the level of nations research has for instance focussed on the effects of employment programmes on the target groups' political behaviour (Blattman, Emeriau, & Fiala, 2017). On the field or group level is has looked at the formation of networks or actor coalitions promoting political issues (Sørensen & Torfing, 2003; Weible, 2006), and on the level of individuals on practices and skills acquired in professional and associational life (Brady, Verba, & Schlozman, 1995).[42]

Ways of engaging politically online have received increasing attention lately, supposedly due to their wider spread, lower threshold for engagement, and higher accessibility than offline activities. Most studies focus on how the internet is used as a means of political mobilization (Di Gennaro, 2006; Stanley & Weare, 2016), as a forum for political debate (Albrecht, 2006), or as a new mode of access to politically relevant information (Gil De Zúñiga, Puig-I-Abril, & Rojas, 2009; Mossberger, Tolbert, & McNeal, 2010). Often they have a focus on how it affects young people (Vissers & Stolle, 2014). However, we lack insights on whether and how targeted online initiatives that aim at promoting citizens' political capital have an actual impact. The activities of AW provide an excellent space for testing whether and how the capital based approach to social impact can be fruitfully applied, with a particular focus on political capital.

[42] See for instance Dacin, Goodstein, and Scott (2002) for a distinction of the different levels, however, in relation to institutional research.

For the concept of political capital I apply here, I relate to its discussion in the theory chapter. Throughout my work it refers to the "engagement and activity of individuals within social settings" and comprises variables that relate to activism, involvement, or influence. In the context of AW, and in contrast to the other studies, it is furthermore applied within the explicit realm of politics. Therefore it addresses the "discursive exchange between a large variety of constituents", including ordinary citizens and politicians alike. It comprises aspects such as the "engagement in political debate, or participation in policy formation" and related components (compare to the theory chapter). The effects of AW on these variables have been tested in a survey for citizens and a separate one for politicians. The effects have been assessed in two ways: (1) by means of comparison between the levels of capitals held by respondents to the survey on the one side and those held by the German population as derived from representative data of the European Social Survey (ESS) on the other side, and (2) by asking citizens and politicians explicitly whether they saw or experienced a range of effects. In the first approach the effects are measured indirectly, in the second one they are measured directly.

Predispositions in cultural capital (relating to norms and values such as trust, solidarity or tolerance) and satisfaction measures (for instance with how democracy works) have been assessed to account for the specific characteristics of CS respondents that might have had a moderating influence of the measured effects. Two further cultural capital items (being interested in current issues; caring for others) and one social capital item (forming communities of interest) have been used as measures of effects complementary to the political capital ones.

7.2 Specifics of the web-platform

The web-platform “abgeordnetenwatch.de” is an independent, nonpartisan internet space on which citizens can get in touch with members of parliament (MPs) of the German National Parliament (Bundestag), the European Parliament, and the parliaments on the German federal state level. AW was founded in 2004. At the time it was only the citizens of Hamburg who could contact their MPs publicly online. Citizens’ questions and politicians’ answers are documented on the website. In late 2006 AW was extended to address the level of the Bundestag and 2008 that of the European Parliament. In 2009 over 90 percent of the politicians that stood for election into the German or the European parliament engaged in some political dialogue with citizens through the platform. AW is the biggest political discussion platform in Germany with about 125.000 users and 275.000 page clicks per year as of August 2017. AW is considered a social enterprise and the founder has been elected an Ashoka fellow.

In addition to the moderation of political dialogue AW performs several targeted activities surrounding political transparency, lobbyism and corruption. Prominent activities of AW include: (1) the documentation of MP’s voting behaviour in parliamentary decisions on politicians’ web profiles on the website; (2) a legal appeal to the Bundestag to reveal information on permanent access permits for lobbyists; (3) an initiative for greater transparency of MPs’ auxiliary income, that is for work performed alongside their political mandate; (4) a petition and subsequently passed legislation for punishing corruption among MPs; (5) an initiative for lowering the threshold for party donations that require revelation of the donors’ identity. The main means for exerting influence on these matters are public communication through online, television and print media, petitions and legal appeals.

7.3 Methods

7.3.1 Participants

Participation in the citizen survey (CS) was not restricted by any mode of preselection, apart from the one that participants had to be users of AW, and a high number of completed questionnaires was sought to increase practical significance of the data. 745 citizens participated in the online survey in total. It was open for ten weeks beginning in late February 2016. Calls for participation in the survey had been issued by AW on their website, in their newsletter and on social media. Access through AW had to be sought since there was no other way of attracting a critical mass of participants otherwise. The calls for participation stressed the independence of the research and that the researcher was in no way affiliated with AW. No direct contact to potential participants was made apart from the modes of access mentioned above. Participants for the politician survey were recruited through contacting MPs on the national, federal state and EU level via their official email accounts, including three reminders. The politician survey (PS) was initiated about a week after the citizen survey but closed at the same time as the latter. Over 2500 MPs (1800 on the federal state level, 630 on the national level, 87 on the EU level) were contacted in total and 255 MPs chose to participate, yielding a response rate of roughly 10 percent.

7.3.2 Analytic strategy

In addition to standard demographic variables, such as gender, region of residence, income and education, some more specific variables have been assessed to describe the target group. In the absence of a proper control group, the characteristics of the group of respondents have been compared to those of respondents to ESS conducted in 2014.[43] In addition to comparisons of socio-economic variables further differences and commonali-

[43] Data of ESS Round 8 conducted in 2016 had not been released yet at the time of the present analysis.

ties between CS and ESS have been explored. Comparisons were drawn on people's level of trust and satisfaction for example, but also on that of their political orientation. Instead of higher level variables pointing to 'fundamental beliefs', such as a preference for security or economic well-being (Sabatier & Jenkins-Smith, 1993, p. 252), participants were asked about their political orientation more directly and offered categories such as 'conservative' or 'liberal' (some call these "policy core beliefs"; Weible et al., 2009).

ESS data have not only been used for comparing general characteristics, but also to reckon how much the users of AW differed in terms of their interest in political issues, their assessment of possibilities for participation in the political system or their estimation of the influence individual citizens can have on political decisions. Citizens' responses were reflected in yet another way, namely through comparing their answers to the views held by politicians as expressed in the separate PS.

7.3.3 Socio-demographics of the groups of respondents

In the following I describe commonalities and differences between the participants in the citizen survey (CS), and thus supposedly the general user structure of AW, and characteristics found in ESS data. The latter are representative of the general population as regards gender, age, education, and region (federal state) when "post-stratification weights" are applied (European Social Survey, 2014d). A further comparison is drawn to the characteristics of the group of politicians that participated in the separate politician survey (PS).

Table 13 provides a summary of these characteristics. While the share of women in ESS is around 50 percent, only about 25 percent of CS respondents were women, and around 30

percent of respondents to PS.[44] In comparison with ESS, people aged 55 years and over are overrepresented in the citizen survey, while those aged between 15 and 34 are underrepresented.[45] The age shares are 16-30-54 percent in CS (lowest-medium-highest stratum), 28-36-36 percent in ESS, and 10-56-34 in PS. There are also fewer people who were not born in Germany or were not a German citizen in CS (about 4 and 1 percent) and PS (about 3.5 percent)[46] than in ESS (about 12 and 6 percent). People living in (1) cities and (2) suburbs are overrepresented against people living in (3) smaller towns and (4) villages in CS and PS as compared to ESS. The shares are: 29-16-33-21 percent in CS, 33-15-33-19 percent in PS and 16-12-37-35 percent in ESS.

The partnership status of respondents differs in several regards. As compared to ESS, PS contains a larger share of married people (73 percent) and a lower share of celibates (16 percent). The last trait is shared by respondents to CS (16 percent of celibates), whereas the percentage of divorced people is higher in CS, especially when compared to ESS (28 percent in CS versus 8 percent in ESS). This is also reflected in the larger share of single person households in CS: 27 percent versus 18 percent in ESS, and 11 in PS.

[44] All figures are rounded to full percentage points in the text.
[45] These and another category between the two are the usual age strata applied in the calculation of ESS weights (European Social Survey (2014a)).
[46] All MPs need to have German citizenship.

Table 13 Socio-demographics CS, PS and ESS

#	Variable	Values	Citizens		ESS		Politicians		N Total		
			N	%	N	%	N	%	Cit	ESS	Pol
1	Gender	female	176	24.5%	1570	51.6%	69	28.8%	717	3045	240
2	Age STRATA	15-34	115	16.2%	847	27.9%	23	10.1%	711	3031	227
		35-54	213	30.0%	1083	35.7%	126	55.5%			
		55+	383	53.9%	1101	36.3%	78	34.4%			
3	Born in Germany	No	27	3.8%	363	11.9%	8	3.4%	711	3045	237
4	German citizen	No	9	1.3%	191	6.3%	-		716	3041	-
5	Locality	City	211	29.0%	487	16.0%	79	32.8%	727	3043	241
		Suburb	118	16.2%	381	12.5%	36	14.9%			
		Town	241	33.1%	1123	36.9%	80	33.2%			
		Village	157	21.6%	1053	34.6%	46	19.1%			
6	Partnership	Married	355	51.1%	1589	52.4%	160	73.1%	695	3031	219
		Divorced	195	28.1%	250	8.3%	23	10.5%			
		Widowed	20	2.9%	213	7.0%	1	0.5%			
		Celibataire	125	18.0%	978	32.3%	35	16.0%			
7	Household #	1	189	27.2%	538	17.7%	23	10.6%	695	3037	216
		2	309	44.5%	1158	38.1%	71	32.9%			
		3	91	13.1%	537	17.7%	53	24.5%			
		4+	106	15.3%	804	26.5%	69	31.9%			
8	Income	0 -1.380 (20%ile)	120	19.0%	451	17.0%	-		630	2648	-
		1.381-2040 (40%ile)	86	13.7%	584	22.0%	-				
		2.041-3.280 (70%ile)	185	29.4%	779	29.4%	-				
		3.281-4.960 (90%ile)	165	26.2%	534	20.2%	-				
		4.961+ (91%ile)	74	11.7%	300	11.3%	-				
9	Occupation	Paid work	364	51.5%	1458	48.8%	-		707	2988	-
		Education	65	9.2%	385	12.9%	-				
		Unemployed	56	7.9%	193	6.5%	-				
		Retired	211	29.8%	710	23.8%	-				
		Household; child	11	1.6%	242	8.1%	-				
10	Education STRATA	Lower (GS, HS)	36	5.0%	1038	34.2%	6	2.6%	722	3031	230
		Medium (RS, GYM)	258	35.7%	1576	52.0%	47	20.4%			
		Higher (Uni)	428	59.3%	417	13.7%	177	77.0%			

As regards income, ESS respondents comply fairly accurately with the decile-based German income distribution, which has been collapsed into five categories in the table. CS respondents also do so for the largest part. However in CS, households earning between 1,381-2,040 € account for only 14 percent whereas those earning 3,281-4,960 € are at 26 percent, while both shares should be at 20 percent to be representative of the general popu-

lation.[47] As regards their occupation respondents to ESS and CS are pretty similar with only some differences, for instance a higher share of retired people (30 percent in CS versus 24 percent in ESS) and a lower share of people doing housework and child caring (2 percent in CS versus 8 percent in ESS).

CS data are marked by a considerable negative skew in education, that is a large share of people in the highest educational stratum:[48] Only 5 percent are at the lower educational level (Hauptschule or no secondary school degree), while almost 60 percent are at the highest educational level (went to university). Another 36 percent are at the medium level (Realschule and Gymnasium). This skew is even stronger in PS, with 77 percent of respondents having been educated at a university. For a comparison only 14 percent of ESS respondents possess higher level degrees, while 34 percent are located at the lowest level.

Table 14 Federal states CS, PS and ESS

#	Variable	Values	Citizens		ESS		Politicians		N Total		
			N	%	N	%	N	%	Cit	ESS	Pol
11	Location	Baden-Württemberg	99	13.7%	386	12.7%	50	19.6%	722	3045	255
		Bayern	112	15.5%	449	14.8%	28	11.0%			
		Berlin	53	7.3%	128	4.2%	16	6.3%			
		Brandenburg	14	1.9%	98	3.2%	4	1.6%			
		Bremen	6	0.8%	25	0.8%	0	0.0%			
		Hamburg	29	4.0%	65	2.1%	13	5.1%			
		Hessen	77	10.7%	223	7.3%	17	6.7%			
		Mecklenburg-Vorpommern	8	1.1%	68	2.2%	0	0.0%			
		Niedersachsen	60	8.3%	290	9.5%	14	5.5%			
		Nordrhein-Westfalen	150	20.8%	662	21.7%	31	12.2%			
		Rheinland-Pfalz	37	5.1%	149	4.9%	26	10.2%			
		Saarland	7	1.0%	40	1.3%	1	0.4%			
		Sachsen	21	2.9%	168	5.5%	17	6.7%			
		Sachsen-Anhalt	8	1.1%	99	3.2%	1	0.4%			
		Schleswig-Holstein	33	4.6%	102	3.4%	13	5.1%			
		Thüringen	8	1.1%	93	3.1%	1	0.4%			

[47] MPs have not been asked about their income, since due to their standard salaries, they should all be in the highest decile.
[48] Strata applied correspond to those used in ESS; see European Social Survey (2014a).

Finally, Table 14 gives a summary of the location of respondents in the 16 German federal states. Generally there is a tendency for people from Eastern Germany (for instance from Brandenburg, Sachsen-Anhalt, Thüringen) to be underrepresented in CS and PS as compared to ESS. Overall, however, CS and PS respondents comply for the most part the with regional provenience pattern of the German population.

7.3.4 Further characteristics: capitals baselines and satisfaction

All three groups have furthermore been compared as to their general level of: (1) political capital (voting; political orientation; obtaining political information on TV); (2) cultural capital (tolerance; solidarity; sense of community; trust); and (3) satisfaction (with the national government; how democracy works; life in general). The measures were chosen as available in ESS. None of them are supposed to be fundamentally affected by the use of AW.[49] The purpose of this comparison was to get a better sense of the character of the different groups.[50] Table 15 contains two political capital dimensions (voting and political orientation), "interest in political information" instead is displayed in Table 16 due to its different format.

About 80 percent of both CS and ESS respondents indicated that they had voted in the last national elections (item 1). Taking ESS as a benchmark, CS data contain much fewer people that describe their political orientation (item 2) as "conservative" (about 9 versus 38 percent) and "nationalist" (about 2 versus 5 percent). At the same time the participation of

[49] Voting in CS referred to several previous elections and has been assessed by a scale ranging from "never" to "seldom" to "mostly" to "always". The latter category has been transformed to "yes" and the rest to "no" for comparison with ESS data. Since the median length of respondents' use of AW was only at about 3 years, the voting variable has been treated as an inclination unaltered by AW.

[50] Initially some further social capital items had been considered, such as "how often do you meet socially with your friends, relatives or colleagues", or "how many people are there, with whom you can discuss intimate or personal matters." These were dismissed due to their relatively lower importance for the present investigation as compared to those listed above. They were also disregarded due to practical restraints to the overall length of the questionnaire and the time needed for completion by respondents.

"communitarian", "green", "liberal" and "left-liberal"[51] people was higher in CS—shares in CS were mostly twice as high as in ESS, in the case of "liberal" thrice as high. Participation in PS was more in line with ESS, yet with a markedly lower participation from "conservatives" (about 24 percent in PS) and higher participation from "liberal" and "left-liberal" politicians (about 16 and 4 percent in PS). While conservative politicians in PS occupied a higher share than conservative citizens in CS, the shares of liberal and left-liberal individuals were almost at an equal level in both surveys.

Table 15 Voting and political orientation CS, PS and ESS

#	Cap.	Category	Item	Citizens		ESS		Politicians		N Total		
				N	%	N	%	N	%	Cit	ESS	Pol
1	PC	Voting	yes	604	82.0%	2149	79.5%	-		737	2703	-
2		Pol. orient.	Conservative	58	9.1%	718	38.3%	59	24.4%	640	1876	242
			Social-democratic	169	26.4%	572	30.5%	76	31.4%			
			Communitarian	101	15.8%	144	7.7%	17	7.0%			
			Green	186	29.1%	236	12.6%	40	16.5%			
			Liberal	92	14.4%	83	4.4%	38	15.7%			
			Nationalist	10	1.6%	88	4.7%	2	0.8%			
			Left-liberal	24	3.8%	34	1.8%	10	4.1%			

Table 16 illustrates the results on the differences in cultural capital and satisfaction (and gathering political information by watching TV as another political capital item). These differences have been assessed by comparing the mean and standard deviation as well as the median in the groups, due to the non-normal distribution of most of the data. They have been furthermore assessed graphically and by means of the Mann-Whitney U test for two independent samples (confidence interval at 95 percent, CI=95%). Z-scores of the test have

[51] Left-liberal was not initially a category, but emerged through frequent mentions in the category "other." In conjunction with PS data, it turned out that mostly people who identified with the "Piratenpartei" used this term to describe their political orientation.

been translated into effect sizes as measured by Cohen's d (1992).[52] A further comparison to politicians has only been drawn for illustrative purposes.

The comparisons show that people responding to CS and ESS have a number of traits in common, which are discussed first before I turn to detected differences.[53] Respondents to both surveys are similarly tolerant, both with regards to accepting sexual diversity (see item 4 Table 16; corresponding question: "gay men and lesbians should be able to live their lives as they wish; medians between "agree" and "fully agree") and cultural diversity (item 5; medians at "allow *some* people from a different cultural background to come and live in Germany"). Politicians are almost at an equal level.

Table 16 Political capital, cultural capital and satisfaction CS, PS and ESS

#	Cap.	Category	Item	Citizens			ESS			Politicians			N Total			Diff. CS-ESS		
				M	SD	MD	M	SD	MD	M	SD	MD	Cit	ESS	Pol	z*	d	x
3	PC	Interest	Polit. info.	2.45	1.616	2	1.80	1.154	2	-	-	-	732	2879	-	-11.66	0.19	x
4	CC	Tolerance	Sexual diversity	1.46	0.729	1	1.79	0.940	2	1.38	0.645	1	724	3022	237	-9.43	0.15	
5			Cultural diversity	2.35	0.870	2	2.09	0.781	2	2.14	0.823	2	731	3024	235	-7.00	0.11	
6		Solidarity	Income inequal.	1.55	0.846	1	2.20	0.971	2	2.07	1.187	2	698	3001	207	-18.57	0.31	x
7		Unity	EU unif.	2.23	1.327	2	2.87	1.199	3	1.69	0.980	1	683	2964	237	-12.87	0.21	
8		Trust	People	6.33	1.810	7	4.95	2.207	5	6.74	1.663	7	734	3044	231	-16.32	0.27	x
9			Legal system	4.97	2.544	5	5.73	2.474	6	7.45	1.968	8	738	3025	240	-7.35	0.12	
10			Political parties	2.49	2.032	2	3.91	2.133	4	5.93	2.186	6	734	3005	231	-15.69	0.26	x
11	Sat.		Govern-ment	2.97	2.319	3	5.00	2.178	5	4.68	2.723	5	733	2929	242	-19.53	0.32	x
12			Democr.	3.72	2.685	4	5.87	2.383	6	6.59	2.650	7	739	3015	242	-18.59	0.30	x
13			Life	7.39	2.078	8	7.38	2.107	8	8.27	1.505	8	727	3038	233	-0.20	0.00	

Remarks:*= z (and d) all p=0.000, except item 13 (not sig.); x= marked difference (highlighted in grey: displayed separately in density plot); M=mean; SD=standard deviation; MD=median; PC=Pol. Cap.; CC=Cul. Cap.; SC=Soc. cap.

[52] A Cohen's d of 0.1 indicates a small effect size, 0.3 a medium size and 0.5 a large size.

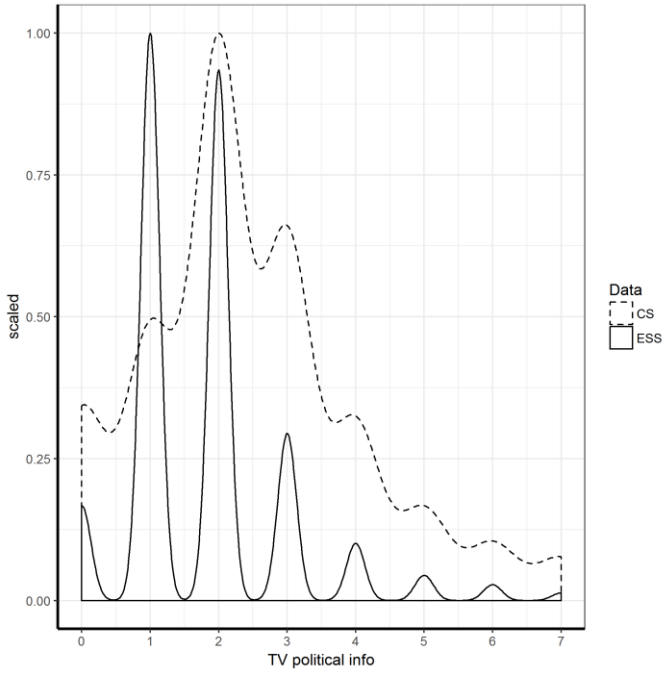
[53] In order to simplify responses for the participants of CS and PS, all questions # 4-7 were recorded on a 1-5 scale ranging from "agree strongly"-"disagree strongly", and all questions # 8-14 on a 0-10 scale ranging from "not at all" to "totally." Table 42 in the Appendix illustrates the original scales from ESS, how they were applied in CS and PS and how the two versions were unified to allow for direct comparisons.

Similarities have also been detected as regards EU unification (item 7), with some higher variation than in the two items before. CS participants tend to agree that "unification should go further", while ESS respondents were located rather at "neither…nor" on that question. Politicians are more pro-unification than both, CS and ESS. Overall, politicians are also more satisfied with their life (item 13; median 8 and mean 8.3) than CS and ESS respondents, who are almost on a par (median 8 and mean 7.4). Respondents to CS and ESS also have a similar level of trust in the legal system (item 9; median 5 for CS and 6 for ESS), while politicians' trust in the legal system is clearly higher (median 8).

In contrast to the above variables, there are several areas in which CS and ESS differ (marked with "x" in the table). Some of the items that exhibit differences have been selected to outline discrepancies in more depth through the illustration of (scaled) density plots (highlighted additionally in grey).[54] To begin with, CS and ESS respondents differ in the time they spend watching news and political information on TV (item 3; see Figure 6).[55] While the majority in ESS consume such information for up to half an hour per day (1 on the scale), the majority in CS do so for half an hour to one hour (2 on the scale), and a substantially higher share of people do so for even greater amounts of time.

[54] Scaled density plots are transformed so that the spot of highest density for each group lies at 1.
[55] Question not contained in PS survey. Scale starting with "no time at all" with subsequent increases by 30 minute intervals up to 7="more than 3 hours."

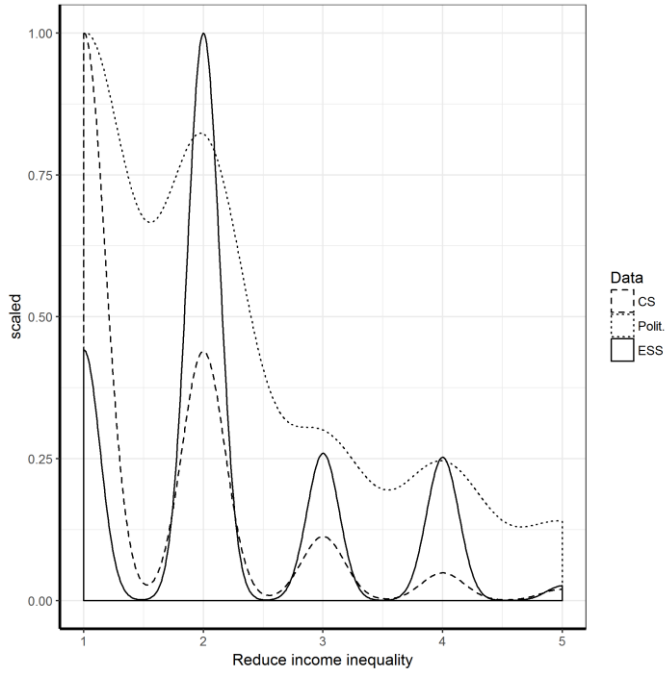
Figure 6 Watching news and political information on TV



In reference to item 6 called "solidarity" and as displayed in Figure 7, CS and PS are similar in that a majority of respondents "agree strongly"[56] that the "government should take measures to reduce income inequalities" (peaks in both groups at 1). However, the relative difference between "full support" and lower levels is clearly more pronounced for CS, who are thus more supportive of the claim. ESS respondents, while remaining on the supportive side, have their stress on "agree" but not on "agree strongly."

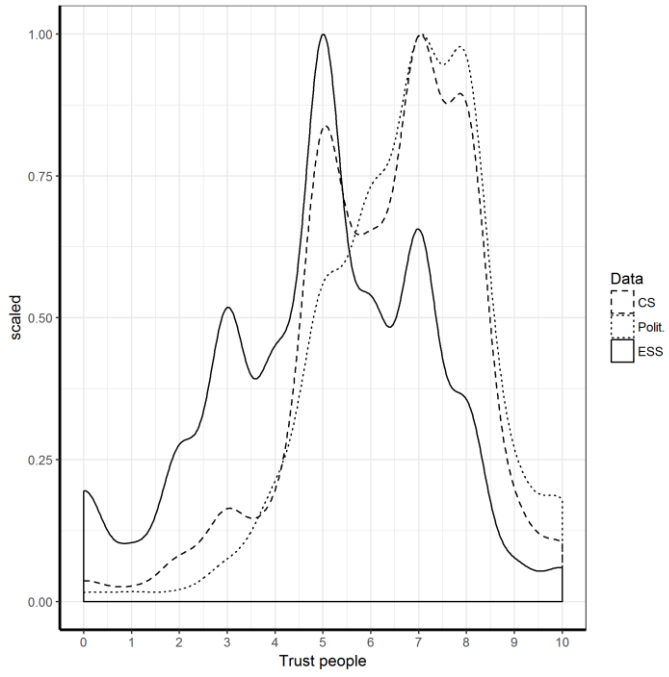
[56] Solidarity scale: 1="agree strongly"–5="disagree strongly."

Figure 7 Reduction of income inequalities



When it comes to trusting people (item 8) the picture looks different, as shown in Figure 8. The plots for CS and PS are almost identical and inclined towards trusting people, while the ESS distribution is centred at a medium level (scale: 0="don't trust"–10="fully trust").

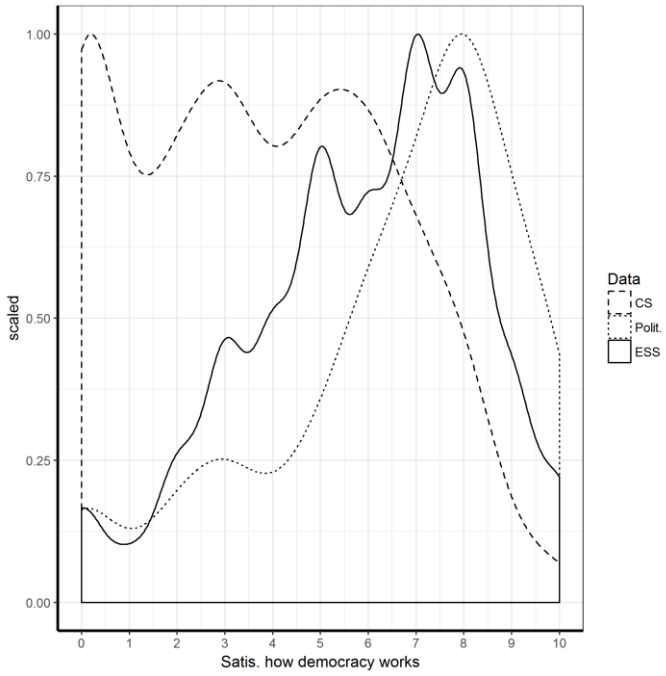
Figure 8 Trust in people



In contrast to trust in people, the participants in CS are more mistrusting of political parties than those in ESS (item 10; median 2 in CS versus 4 in ESS; both levels expressing scepticism though). CS respondents are also more dissatisfied with the work of the government (item 11; median 3 in CS versus 5 in ESS). It needs to be remarked that the 'refugee crisis' which peaked in 2015/2016 might have caused shifts in the level of satisfaction with the government in the population, which cannot be detected in the 2014 ESS data.

A further difference relating to the scepticism towards the political system is the respondents' different take of how well democracy works in Germany (item 12). As shown in Figure 9 the share of people dissatisfied with "how democracy works" is prominent in CS (dashed line), while the distribution in ESS is left-skewed leaning towards a positive assessment (solid line). Politicians instead are even more positive (dotted line). In particular as compared to CS, politicians are also much more trusting of political parties (item 10; median 6 for ESS versus 2 for CS) and satisfied with the national government (item 11; median 5 for ESS versus 3 for CS).

Figure 9 Satisfaction with how democracy works



Based on the above observations we can say that respondents to CS in comparison to the general population are: more interested politically, more solidary and more trusting of people in general, but also more sceptical of the established political system (more mistrusting of parties, and less satisfied with the government and how democracy works).

7.3.5 Effect measures

The questions posed in the surveys referred to two main populations that AW potentially affects: politicians and citizens. Both were incorporated in each of the surveys. So I can mostly present the 'self' and the 'other' view of effects on each target population. Results to be presented in the following are, however, not structured by the target groups, but by the mechanisms by which AW can produce impacts: (a) effects of the website and its use; (b) AW's organisational activities (information on MPs' voting behaviour, lobbyism etc.).

There are furthermore two types of effects the website can have: immediate and mediated effects (see Table 17). The immediate effects do not need much time to evolve and can best be assessed by asking directly whether an effect has been experienced or not. Mediated effects in contrast to immediate ones need to grow and manifest over time. They include potentially unintended spill-overs of using AW and do not focus on intentional change only. They are best assessed by estimating how capital levels in CS differ from the level of capitals held in other groups (here: the general population). But they can also be assessed by asking for effects directly.

Table 17 Effect measures

#	Level	Mechanism	Capital*	Category
1	Effects of platform	Immediate		Communication
2				Advocacy
3			Cultural	Inclusion (of formerly excluded)
4			Social	Community (of interest)
5		Mediated	Cultural	Consciousness
6				Literacy
7				Engagement
8				Influence
9	Effects of AW activities	Immediate		Information
10				Politicians' effort
11				Politicians' exposure
12				Restriction of third party influence

Remark: *=All except for those marked=Pol. Cap.

The table only outlines the general categories used to differentiate the capitals further. It will be seen further which specific items these are composed of. For a lack of previous studies on the effects of specific interventions and the unique character of AW's activities, the immediate impacts have been assessed by a number of generic questions. The categories assessed comprise: (1) communication (in relation to revelation of political positions, but also political dialogue); (2) advocacy (promotion of thematic issues); (3) inclusion (in the sense of the creating political interest in people, who were previously disinterested); (4) community (in the sense of enabling the formation of like-minded interest groups).

Category 4 called "community" contains some reference to coordinated behaviour and coalition building (Sabatier, 1998). Effectively it is more about building a community among users though and its relational character makes it a social capital variable. Category 3 referred to as "inclusion" also deviates from the activity-based character of political capital. It refers to an initial sensitization for political issues of people formerly disinterested. The normative character of showing interest in politics makes it a cultural capital variable. All other variables belong to political capital.

In contrast to the immediate ones, the questions on the mediated effects on citizens were mostly derived from existing literature along the four broad categories: (5) consciousness (knowing and caring about issues); (6) literacy (being able to participate in political processes); (7) engagement (passive and active); (8) perceived influence (access to and being able to shape policy making).

Most of the questions applied have been used in assessing the state of contemporary democracy in the 'Citizenship, Involvement, Democracy' (CID) survey, which was performed between 2000-2004 in several European countries and in the USA (Andersen et al., 2007; Howard, Gibson, & Stolle, 2006; van Deth, Montero, & Westholm, 2008). Many of these questions have consecutively found their way into the European Social Survey and been discussed in-depth by (Dalton, 2008). I therefore refer to his work quite intensely, though relabel most of the concepts to relate more strongly to the capitals terminology, which makes it easier to distil the essence of the different components. The same applies to some other previous work I refer to (for instance Sørensen & Torfing, 2003).

Political consciousness relates to what Dalton calls 'social citizenship' (2008), that is perceived ethical and moral responsibility towards others. The latter comprises for instance awareness for societal issues that matter and the inclination of caring for people with weak social positions in society. It thus shares some similarities with the 'spurring interest in disinterested citizens' variable from above. Since both these aspects comprise norms or are motivated by them, such as the idea that it is important to know what is going on in society, and although they relate to politics, political consciousness has no or only a weak activity-based character. For this very reason, political consciousness, in contrast to all categories to follow, is treated as cultural capital.

Political literacy in turn refers to some of the items that Dalton calls 'autonomy', namely forming a political attitude independent of others and the competence to communicate positions in a political dialogue.

Political engagement includes passive forms of action, for instance gathering information, following discussions etc., just as "direct and individualised forms of action" (Dalton, 2008, p. 92). The active forms of engagement have been specified further in a follow-up question. Citizens were asked to indicate how often they engaged in activities ranging from sharing political information, to protesting, to petitions, to founding new political groups or getting engaged in existing ones, to donating to political parties, to "political consumerism" (Norris, 2013), that is boycotting products out of political reasons. An in-between item in terms of activity and passivity is the aspect of "encouraging others to vote" (Kalaycioglu & Turan, 1981), and has thus been listed separately.

Political influence is occupied with the question of whether people have access to politicians, whether they can influence politicians' agendas and decisions, and whether they overall have the feeling that they can make a difference within the political system in place (see for instance Sørensen & Torfing, 2003).

Finally, the potential effects of AW's organisational activities are always immediate and span: (9) offering structured political information; (10) pushing politicians' efforts; (11) leverage on politicians' public exposure; (12) the restriction of third party influence on policy making.

7.4 Results

The presentation of results happens against a general description of patterns of use of the platform, which I provide as a grounding reference here: The average length of using AW was about 3 years for citizens and 5.5 years for politicians. The majority of citizens used the platform one or several times per month, the majority of politicians less than once per month. The largest share of respondents in both groups spent less than half an hour per week on AW. While the main purpose for politicians was to communicate with citizens, cititzens themselves indicated that the main use of AW for them was for gathering information. For citizens communication with politicians on AW seems to have 'crowded-out' or replaced other electronic modes communication, for instance via e-mail or other websites. No such change could be observed for politicians. Using AW for them is rather in line with a general increase in the use of electronic communication in relation to other modes, such as personal contacts. The main observation in terms of patterns of use and effects triggered thereby is that the share of citizens, who had "never been in touch" with politicians in any way—electronic or personal—decreased from about 39 to 17 percent after citizens had started using the platform.

Further details on how AW is employed by its users are reported in the Appendix. The information provided also refers to issues such as which parties citizens had previously been in touch with and which of them they viewed as most accessible. The results presented here instead directly relate to the measures of AW's effects introduced in the methods section.

7.4.1 Immediate effects on capitals

All variables have been tested by asking citizens and politicians directly about their opinion if AW had a particular effect in a number of dimensions.[57] As outlined in Table 18 the immediate effects refer to five political capital variables. The first three relate to information and communication, another two to advocacy (items 4 and 5). The strongest effect as regards political capital is the informational one (item 1). Both, citizens and politicians agree that AW "makes politicians reveal their positions", that is increased transparency. However, citizens assess the effect to be considerably higher than politicians (median in CS=8, in PS=6).

Table 18 Capital formation: immediate effects

#		Cap.	Category	Item	Citizens			Politicians			N Total		Diff. CS-PS		
					M	SD	MD	M	SD	MD	Cit.	Pol.	MD diff.	z*	d
1	1+	PC	Communication	Make politicians reveal positions	7.70	2.352	8	5.44	3.083	6	714	227	2	-10.085	0.33
2	2			Stimulate dialogue between citizens and politicians	5.99	2.684	6	3.81	2.943	3	686	226	3	-9.394	0.31
3	3			Stimulate dialogue between politicians across	5.57	2.899	6	2.40	2.591	2	650	225	4	-13.044	0.44
4	I		Advocacy	Raise awareness for neglected	7.85	2.196	8	4.79	2.881	5	710	228	3	-13.870	0.45
5	II			Promote specific political issues	7.24	2.483	8	3.74	2.805	3	685	223	5	-14.342	0.48
6		CC	Inclusion	Spur interest in disinterested	5.97	2.819	6	2.36	2.327	1.5	668	216	4.5	-14.579	0.49
7		SC	Community	Interest group formation	6.99	2.632	7	2.88	2.593	2	659	198	5	-15.343	0.52

Remarks: *= z (and d) all p=0.000; M=mean; SD=standard deviation; MD=median; PC=Pol. Cap.; CC=Cul. Cap.; SC=Soc. cap.
+= Listing used in corresponding Figure for clustering along capital components.

As a follow-up to this informational aspect two further questions were implemented in the survey. One question referred to whether AW had "increased the coherence of politicians' positions" through documenting their statements on certain issues over time. Another ques-

[57] All items have been assessed on an 11-point scale ranging from 0-10 ("not at all"-"totally"). Due to the non-normality of the data median comparisons are more telling and non-parametric tests, namely the Mann-Whitney U test (CI=95%), have been used to examine the size of differences between the judgment of citizens and that of politicians. Cohen's d has been calculated to account for the size of the differences.

tion asked respondents whether AW "made politicians' decisions and line of argumentation more traceable." The coherence aspect was neutrally assessed by both groups (medians at 5). The same goes for the traceability aspect when it comes to politicians, while there was slightly more support for it among citizens (median at 6). For the apparent unimportance of the items they have not been reported in the table.

A contrast between citizens' and politicians' views becomes apparent in the 'dialogue items' (items 2 and 3). Citizens support the claim that AW "stimulates a dialogue between citizens and politicians" and "among politicians across party borders" (both medians at 6), whereas politicians deny the claim (medians at 3 and 2). The difference between citizens' and politicians' views is "medium" when judged by Cohen's d for item 1 and 2 (d=0.33 and 0.31) and tends towards a "strong" effect for item 3 (d=0.44).

Among citizens, the level of support for the two advocacy items "AW helps raise awareness among politicians for neglected issues" (item 4) and "AW helps citizens promote specific issues" (item 5) is close to the support for item 1 (medians for both items at 8). The contrast to politicians' views on these two items, however, is even sharper than in the previous case. Politicians' support for the claim is at a median of 5 for item 4, and at a median of only 3 for item 5. Consequently, the difference between the two groups is high according to Cohen's d (=0.45 and 0.48 respectively). Citizens had the chance to point out which neglected or specific issues AW helped promote in a text box. Among the ones most frequently mentioned were: TTIP, lobbyism, refugees, welfare/social policy (issues surrounding inequality), citizen engagement, and ecology.

Two further capitals have been assessed as direct effects, each by a single item, namely cultural capital (item 6) and social capital (item 7). The cultural capital measure was a

question on whether AW "was able to spur interest in politics in groups or individuals previously disinterested with politics." While politicians strongly disagreed (median=2), there was slight agreement among citizens (median=6). The social capital question referred to AW's "capacity to bring together likeminded citizens and to form communities of interest." Citizens agreed with the statement (median=7), while politicians clearly disagreed (median=2). The stark contrast in the groups' view on the matter resulted in large effect differences (d=0.49 for item 6 and 0.52 for item 7).

Figure 10 Immediate effects: differences between citizens and politicians

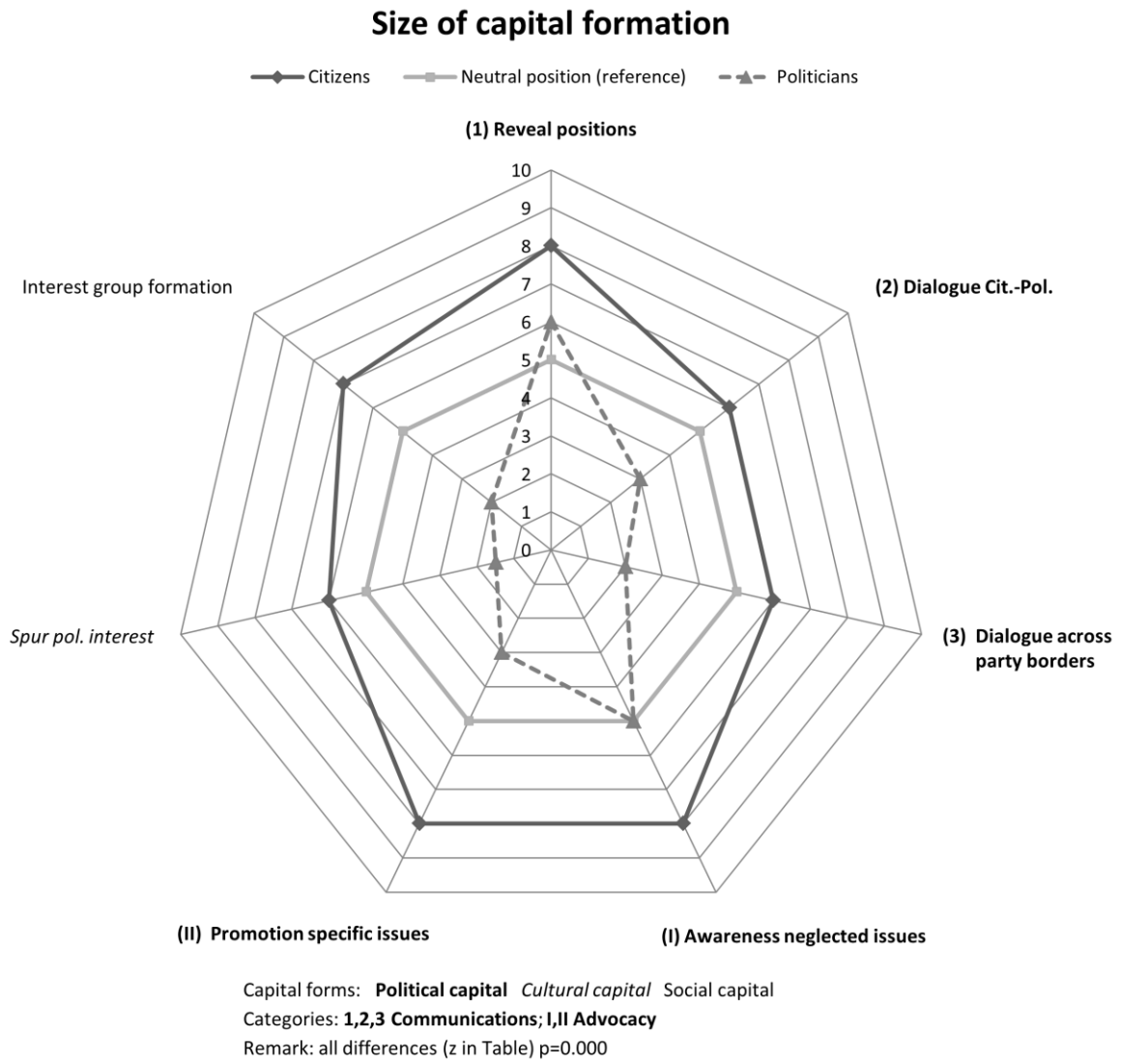


Figure 10 summarises the views of citizens and politicians in a radar chart, in which 0 marks total dismissal and 10 full support for a claimed effect. The level of 5 is marked as the "neutral position" in the graph, where respondents are undecided whether there is an effect or not. Several observations are striking: Politicians only agree with a single item, the one on the "revelation of political positions" by the website (item 1), and are neutral towards another, the one on "raising awareness for neglected issues" (item 4). They clearly dismiss all others, that is think there is no effect on these dimensions. Citizens in contrast tend to agree with all the assessed items. Yet, there is a lower level of support for three of them: Two on the promotion of a political dialogue, either "between politicians and citizens" (item 2), or "between politicians across party borders" (item 3). The third one relates to the aspect of inclusion, more specifically the cultural "sensitisation for political issues of citizens previously disengaged with politics" (item 6). In addition to the "revelation of political positions", citizens are very positive about the variables "raising awareness for neglected issues" (item 4) and "the targeted promotion of specific issues" (item 5) as well as about "the community building function" of AW (item 7).

7.4.2 Mediated effects on capitals

Since some standard assessment of the investigated variables is available in ESS—if partly in a different format—a two-step analytic procedure has been chosen to account for mediated effects on citizens:

First, users of AW were asked about the size of their own political capital (political literacy, engagement, influence) and cultural capital (political awareness).[58] Differences in the

[58] In parallel to previous questions, answers were recorded on an 11-point scale ranging from "not at all" to "totally."

level of capitals between CS and ESS are a first indication of impacts of AW, whereby ESS respondents serve as a quasi-control group.

Second, the research needed to find out whether differences were mainly due to the fact that AW users and the general population had different inclinations to begin with, as partly suggested by the initial comparison of the base levels of political capital, cultural capital and satisfaction in CS and ESS data discussed in the methods section. For this reason, respondents were asked directly to indicate which influence AW had had on the level of their capitals. In order to reveal differences between the effects perceived by citizens and politicians' estimation of AW's effects, politicians were asked to make the same assessment. For politicians, there was obviously no initial question about the level of citizens' capitals, but a direct question about AW's suggested influences.

In order to probe the area of active engagement and complementary to the above there had been some follow up questions on the connection between using AW and other forms of engagement.

7.4.2.1 Comparison of capital levels

Table 19 outlines citizens' levels of several capital (components) as found in CS and ESS. The ESS questions that correspond to those asked in CS can be found in the footnote to the table. While differences on all other questions were tested based on an 11-point scale, the comparators for items 1 and 2 necessitated some data conversions on both sides.[59]

[59] For item 1 in order to meet ESS data, which was recorded on a 1-4 scale (with 1 marking the highest point and 4 the lowest one), CS data was reverse-coded and condensed to match the latter scale. Recoding of CS: 0, 1 & 2 → 4; 3, 4, & half of 5→ 3; other half of 5, 6 & 7 → 2; 8, 9 & 10 → 1; (5 split by half and allocated to 2 and 3).
For item 2, both CS and ESS data were condensed to a 1-5 scale (with 1 marking the highest point and 5 the lowest; as for item 1 CS also had to be reverse-coded). Recoding of CS: 0 & 1 → 5; 2, 3 & 4 → 4; 5 →3; 6, 7 & 8 → 2; 9 & 10 → 1. Recoding ESS: 1 & 2 unchanged; 4 &5 → 3; 5 → 4; 6 → 5.

In CS, the level of "awareness for important political issues" and of "awareness for the needs of fellow citizens" (cultural capital) measured by items 1 and 2 is high in both cases (median=9 or 1 on the converted scale for item 1; median=8 or 2 for item 2). But there is a huge discrepancy with regard to the difference between CS and ESS as measured by Cohen's d. While the difference between CS and ESS in relation to item 1 is very large (d=0.78), the difference between the two in relation to item 2 is very small (d=0.07). The German population on average thus cares for their fellow citizens as much as the users of AW, but not nearly as much about political issues as AW users.

Table 19 Mediated effects: levels of capitals CS and ESS

#		Cap.	Category	Item	Citizens			ESS			N Total		Diff. Cit.-ESS		
					M	SD	MD	M	SD	MD	Cit.	Pol.	MD diff.	z*	d
1	1+	CC	Awareness	Care about important political issues[1]	8.87	1.359	9								
				1 converted to 1-4[x]	1.35	0.527	1	3.44	0.831	3	713	3043	2	-47.619	0.86
2	2			Want to help worse-off[2]	7.93	2.154	8								
				2 converted to 1-5[x]	2.08	1.424	2	1.91	0.742	2	697	3001	0	-4.135	0.08
3	I	PC	Literacy	Form own opinion	8.25	1.784	8				701	-			
4	II			Able to communicate political positions[4]	7.85	1.870	8	4.83	2.628	5	702	3010	3	-26.946	0.44
5	A		Engagement	Motivate others to vote	7.51	2.870	8				686	-			
6	B			Engage passively	8.05	2.562	9				683	-			
7	C			Engage actively[7]	7.12	2.772	8	4.04	2.843	4	694	3023	4	-23.424	0.38
8	D			Want to be party/group member	3.47	3.623	2				660	-			
9	i		Influence	Have access to politicians	5.72	3.261	6				671	-			
10	ii			Can influence political decisions[10]	5.19	3.315	5	3.77	2.487	4	669	3007	1	-10.124	0.17
11	iii			Think what they want/ do matters politically[11]	5.57	3.245	6	3.56	2.412	3	674	3014	3	-14.769	0.24

Remarks: [x] for comparison to ESS; *= z (and d) all p=0.000; M=mean; SD=standard deviation; MD=median; PC=Pol. cap.; CC=Cul. cap.; [+]= Listing used in corresponding Figure for clustering along capital components.

[1, 2, 4, 7. 10. 11]= ESS reference questions per item: 1: "How interested are you in politics?"; 2: "It is important to help the people around you and care for their well-being." (originally 6-point scale, collapsed to 5-point for comparison); 4: "How confident are you in your own ability to participate in poltitics?"; 7: "How able do you think you are to take an active role in a political group?"; 10: "How much would you say the political system allows people like you to have an influence on politics?"; 11: "How much would you say the political system allows people like you to have a say in what the government does?"

For CS respondents, the two political literacy items, namely "forming one's own opinion independent of others" (item 3) and "feeling able to communicate political positions" (item 4) are at the same level (medians=8) as the awareness items discussed before. Only one of the two could be compared to ESS data and the difference turned out to be close to large

when judged by Cohen's d (0.44), but not as large as the one on "caring about political issues."

The engagement variables reveal an interesting distinction. Items 5-7 ("motivating others to vote", "engaging passively", engaging actively") are again at a high level in CS with medians at a level of 8 or 9 on the scale. When item 7, namely "engaging actively", is compared to ESS data, we see that the difference is quite pronounced, its size being between a medium and large effect (d=0.38). Item 8 "wanting to be a party member or engage in another political group" in contrast is at a very low level in CS (median=2). This suggests that engagement that is not formalized in an organization is much more relevant to AW users than the latter.

Although marked by a small positive tendency, the three political influence items are clearly at lower levels than the other items. The medians for "having access to politicians" (item 9), "being able to influence political positions" (item 10) and "what citizens do and want matters politically" (item 11) are at 6 and 5 on the scale respectively. And when items 10 and 11 are compared to ESS data (medians at 3 and 4), the size of the difference is only in between "small" and "medium" (d=0.17 and 0.24).

Figure 11 Mediated effects: differences CS and ESS

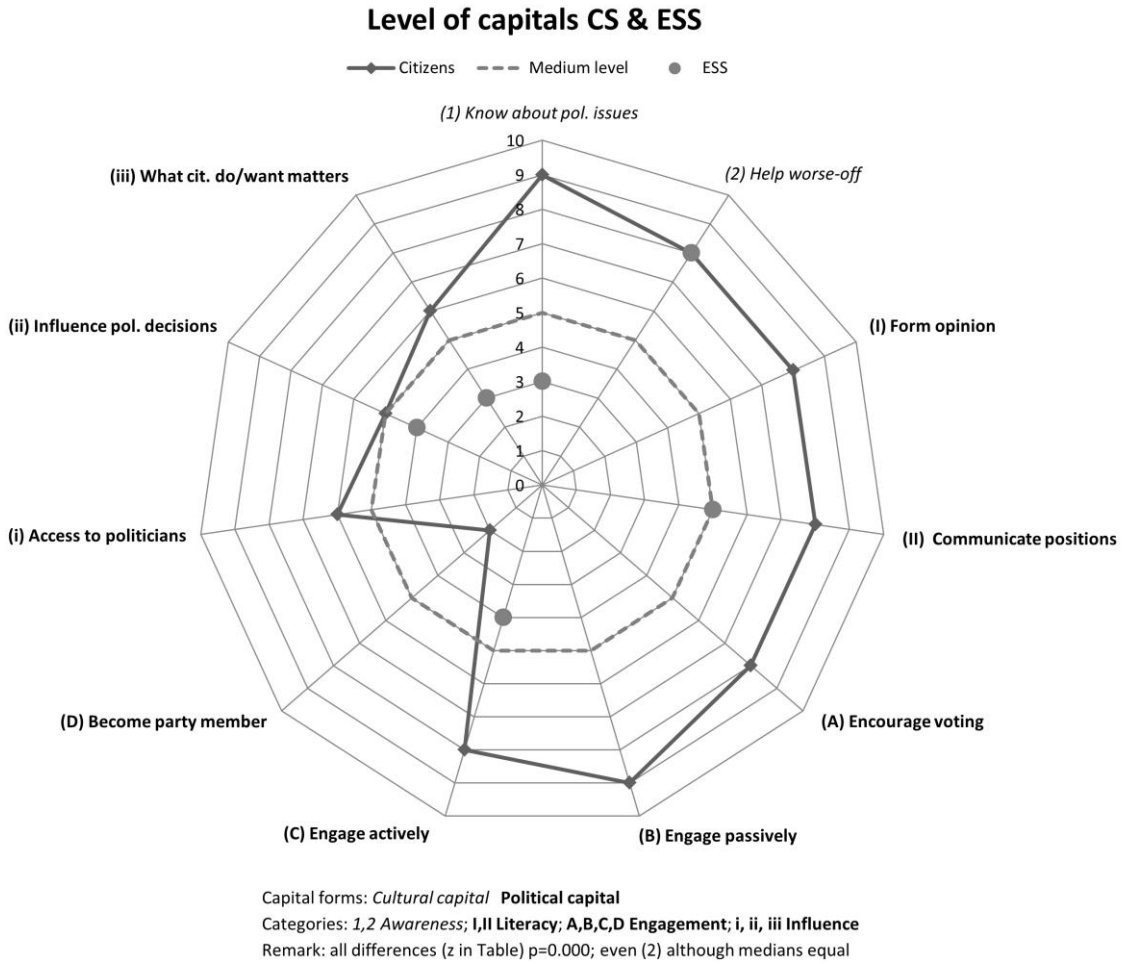


Figure 11 illustrates the differences between the levels of the capitals found in CS and ESS. Items 1 and 2 have been rescaled to a 0-10 scheme.[60] The dark solid line marks the levels measured in CS and the dots mark ESS reference positions, for those items where comparisons were possible. The dashed line marks the "medium or neutral level" of 5 and serves only for more easily identifying high or low levels of the capitals. It is becoming clear that there is a considerable difference between CS and ESS throughout all four categories, except for item 2 in the awareness category. It is also evident that capital levels are higher on 'awareness', 'literacy', and 'engagement' than on 'influence'. The outlier in the

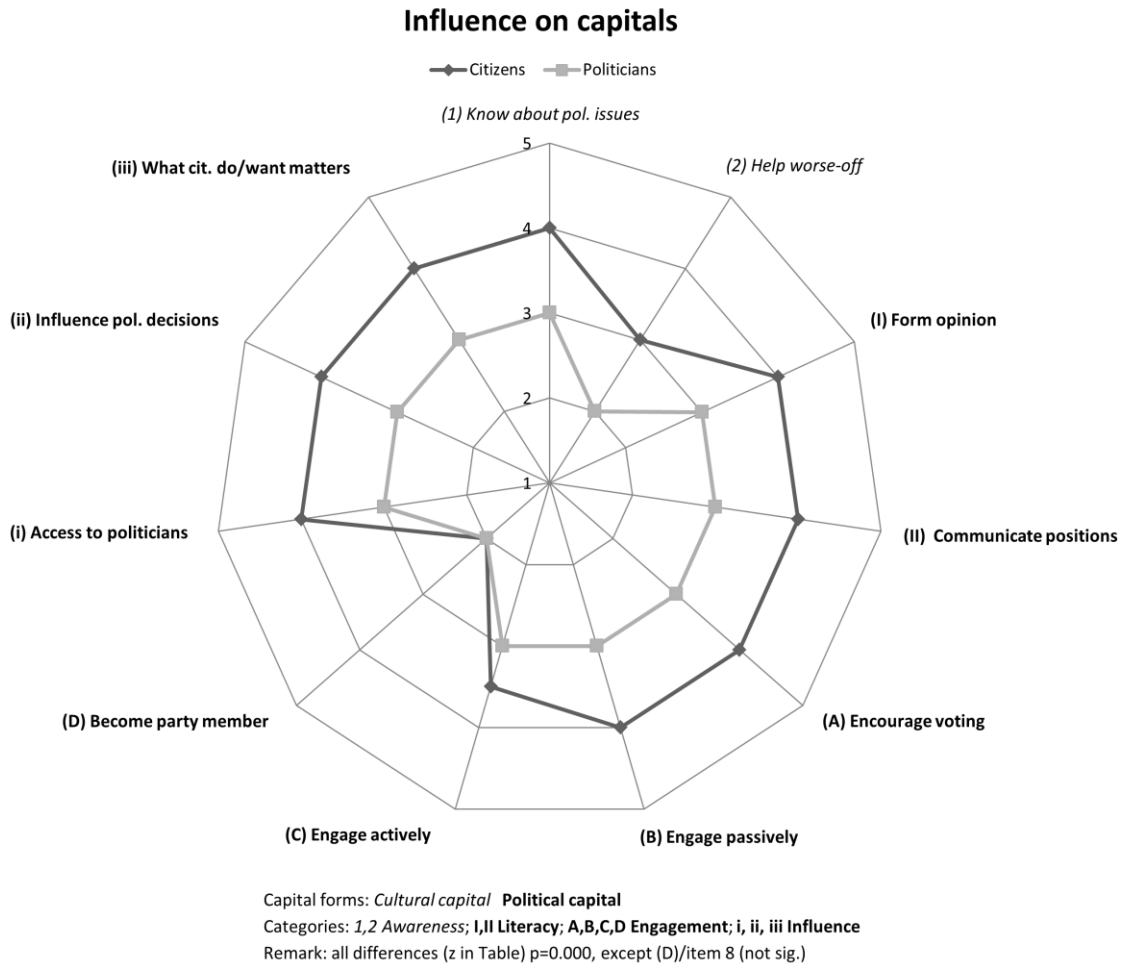
[60] The rescaling of the items was based on the percent difference between the medians in the two groups in relation to the respective applied scale. While this difference for item 1 is at 2/3 (yielding an ESS position of 3 on the 0-10 scale, given an original median of 9 for CS), it is close to 0 for item 2 (which is why ESS is at the same level as CS on the 0-10 scale).

engagement category when it comes to formalized engagement (item 8; D in the graph), reconciles with the scepticism towards political parties and the government found in the initially assessed base levels of trust and satisfaction of CS respondents.

7.4.2.2 Citizens' and politicians' assessment of mediated effects

As referred to in the above description of the two steps in the analysis, citizens were also asked to indicate explicitly how much AW influenced the formation of capitals.[61] Citizens' statements were compared to politicians' estimation of effects. Both are summarised and displayed in Figure 12.

Figure 12 Influence of AW on capital formation: differences CS and PS



Capital forms: *Cultural capital* **Political capital**
Categories: *1,2 Awareness*; **I,II Literacy**; **A,B,C,D Engagement**; **i, ii, iii Influence**
Remark: all differences (z in Table) p=0.000, except (D)/item 8 (not sig.)

[61] Assessment on a 5-point scale ranging from "none at all", to "very low influence", to "some influence", to high influence", to "very high influence."

The figure shows that citizens assess AW's influence as "high" on all tested items, except for "wanting to help others" (item 2, median at "some influence") and "becoming a party member" (item 8, median at "low influence"). The pattern of politicians' assessment is almost a mirror of that of citizens, yet transferred to a lower level each, except for item 9 "becoming a party member", where the medians in CS and PS are equal. Item 2 "helping worse-off" is thus at "low influence" and all other items are at "some influence." Thus, although politicians seem to see an influence by the platform, its magnitude is smaller than when assessed by citizens themselves.

Table 20 Influence of AW on capital formation: levels CS and PS

#		Cap.	Category	Item	Citizens			Politicians			N Total		Diff. Cit.-Pol.		
					M	SD	MD	M	SD	MD	Cit.	Pol.	MD diff.	z*	d
1	1+	CC	Awareness	Care about important political issues	4.09	0.830	4	2.97	0.880	3	674	233	1	-14.751	0.49
2	2			Want to help worse-off	3.29	0.992	3	2.63	0.744	2	624	227	1	-11.037	0.38
3	I	PC	Literacy	Form own opinion	4.03	0.830	4	2.90	0.864	3	634	231	1	-14.761	0.50
4	II			Able to communicate political positions	3.85	0.850	4	3.04	0.932	3	626	232	1	-9.562	0.33
5	A		Engagement	Motivate others to vote	3.48	1.071	4	2.69	0.769	3	617	230	1	-10.220	0.35
6	B			Engage passively	3.96	0.952	4	3.07	0.925	3	636	230	1	-11.392	0.39
7	C			Engage actively	3.47	1.027	4	2.96	0.919	3	640	228	1	-6.859	0.23
8	D			Want to be party/group member*	2.63	1.058	2	2.48	0.654	2	606	227	0	-1.33	0.05
9	i		Influence	Have access to politiciancs	3.67	0.978	4	3.25	0.920	3	633	229	1	-5.738	0.20
10	ii			Can influence political decisions	3.57	0.987	4	2.78	0.764	3	635	230	1	-10.769	0.37
11	iii			Think what they want/ do matters politically	3.50	1.013	4	2.92	0.852	3	602	229	1	-7.980	0.28

Remarks: *= z (and d) all p=0.000, except item 8 (not sig.); M=mean; SD=standard deviation; MD=median; PC=Pol. cap.; CC=Cul. cap.; += Listing used in corresponding Figure for clustering along capital components.

A closer look at the distribution of responses and not only the medians (as previously, assessed by the Mann-Whitney U test and translated to Cohen's d), gives deeper insights into the differences between citizens' and politicians' views. The results are displayed in Table 20. Apart from the item on party membership (item 8; d=0.05), the size of the differences is least on item 9 ("access to politicians"; d=0.2) and item 7 ("engage actively"; d=0.23). It is biggest in relation to item 3 ("form own opinion"; d=0.5) and item 1 ("know about polit-

ical issues"; d=0.49). Although there is no general pattern, deviations between politicians' and citizens' standpoints are smaller on those capital components where the difference between AW and ESS respondents is also small—in particular the "political influence" items.

7.4.2.3 Probing active engagement

In parallel to the strategy above and as a complementary analysis, I have taken a more detailed look at forms of active engagement AW could have an influence on. Participants in CS were given a list of items and asked how often they performed the particular activities: (1) participating in a political event, (2) initiating or signing a petition, (3) sharing information electronically, (4) boycotting products, (5) engaging in a political group, (6) engaging in a political party, (7) being involved in party campaigning, and (8) donating to a political party. Citizens were asked to indicate how often they performed the listed activities on a scale ranging from "never", to "once per year", to "several times per year", to "several times per month", to "weekly". Citizen distributions were compared to ESS data, available for all activities except for items 3, 7 and 8, and available only in a binary format.

Table 21 Active political engagement: differences CS and ESS

		Citizens						ESS		Infl. AW
#	Activity	Never	1 x year	>1 x year	>=1 x month	weekly	N	Never	N	Share (N=723)
1	Political event	29.8%	29.5%	32.9%	6.3%	1.6%	699	91.3%	3043	1.7%
2	Petition	1.5%	6.4%	36.8%	27.9%	27.4%	720	67.3%	3038	52.0%
3	Sharing info. electronically	9.5%	5.2%	31.0%	27.0%	27.3%	714			20.6%
4	Boycotting products	16.3%	6.6%	33.8%	15.8%	27.5%	669	68.2%	3041	7.2%
5	Engaging in a political group	53.3%	6.5%	17.3%	11.1%	11.9%	649	73.7%	3042	1.4%
6	Engaging in a political party	82.3%	4.0%	5.1%	4.4%	4.2%	643	96.4%	3042	0.7%
7	Party campaigning	63.6%	8.4%	16.3%	6.4%	5.3%	656			1.2%
8	Party donations	56.2%	16.0%	19.6%	7.1%	1.2%	689			2.6%
9	(None)									12.6%

Remarks: Medians marked in grey.

CS respondents were additionally asked about *the* main activity AW had an influence on. Politicians did not have to answer these more detailed questions, since they really only are of value when reported as actual activities by citizens and not assessed as hypothetical ones by politicians. Table 21 illustrates the CS–ESS comparison.

Items 5-8 serve to explore in more detail item 8 from above ("becoming a member of a political party or another political group"). This is done by differentiating between parties and other political groups and also by listing other forms of active support than becoming a member (donating and campaigning). With the median lying at "never" for all items (all medians marked in grey in the table), they outline that the AW users tend to be more averse to organizational engagement in politics, be it membership, campaigning, or donations, than to individualised forms of engagement. There are some nuances to be considered though. First, membership/engagement in some other group occurs more often in CS than membership in an established political party. Also, the share of people, who are a member in either is considerably higher in CS than in ESS. In turn the share of people, who have campaigned for a party or donated to one, is higher than of those directly engaged in a party. The overall degree of formal political engagement is thus not low at the absolute level, but less important in relative terms. What is more, the spill-overs from AW to formal engagement are particularly small as will be seen in the following.

Items 1-4 are much more prevalent than items 5-8, with item 2 and 3 ("petitions" and "sharing information") as the most prominent ones. The median in both cases is at "several times per month." There is a huge difference between CS and ESS when it comes to petitions. While about 67 percent of the population have never signed a petition, it is not even 2 percent of AW users. Although at different levels, the contrast is similarly strong with regard to participation in a political event: Only 30 percent of AW users have never done

it, but over 90 percent of ESS respondents. For CS respondents, the median lies at "once a year" when it comes to participation in political events. Finally, the contrast in "boycotting products" (median at "several times per year" in CS) is somewhat smaller but still huge, with 68 percent in ESS and only about 16 percent in CS, who have never done it.

Items 2, 3 and 4 (petitions, sharing information, boycotting) are also the most prominent ones with regard to activities on which AW had an influence (see last column in Table 21): 52 percent indicated "petitions" as the one activity that AW had influenced most, about 21 percent picked "sharing information electronically" and about 7 percent "boycotting products." All other activities were of minor significance compared to the latter. It is striking to see that only about 13 percent of AW users reported that AW had no influence on any of the listed activities (line number 9 in Table 21).

7.4.3 Effects of organisational activities on political capital

With regard to AW's additional activities, both citizens and politicians had first been asked whether they knew about them at all. Activities included: (1) politicians' profiles, on which their voting behaviour is documented alongside their communication with citizens; (2) a legal appeal to reveal permanent access permits for lobbyists to the Bundestag; (3) an initiative to reveal MPs' auxiliary income; (4) a petition and subsequent legislation against corruption among MPs; (5) an initiative for lowering the requirements for mandatory public reporting of donations to political parties.

In a second step, respondents were asked to choose from a list of potential effects that these activities might have, ranging from the structuring of available political information to increasing the public exposure of politicians. Finally, respondents were asked to assess whether they found the effect they had selected positive or negative on a 0-10 scale. Table

22 outlines whether respondents knew the listed activities. It shows that citizens know AW's activities better than politicians, except for politicians' profiles where the publicity is nearly the same in both groups (at 73-74 percent). It also shows that items 2-4, relating to lobbyism, auxiliary income and party donations are more well-known (at or above 88 percent in CS and between 55 and 65 percent in PS) than the initiative against corruption (65 percent in CS, about 41 percent in PS).

Table 22 Knowledge of AW's activities CS and PS

		Citizens			Politicians		
#	Activity	Yes	%	N	Yes	%	N
1	Politicians' profiles & voting statistics	523	74.4	703	177	73.1	242
2	Legal appeal on revealing permanent access permits for lobbyists	690	96.5	715	133	55.0	242
3	Initiative for greater transparency on MPs' auxiliary income	679	94.8	716	156	65.3	239
4	Initiative for lowering threshold for mandatory reporting of party donations	626	87.8	713	133	55.6	239
5	Initiation of legislation against corruption among MPs	459	65.0	706	98	41.4	237

Table 23 highlights which effects citizens and politicians see as most prominent in relation to the above mentioned activities. The ranking performed in the table is a heuristic one oriented at the combined recognition of an effect by citizens and politicians.

Table 23 Effects of AW's activities

		Is effect?*					Assessment of effect							
		Cit.		Polit.		Rank	Cit.				Polit.			
#		Yes	%	Yes	%		Mean	SD	MD	N	Mean	SD	MD	N
1	Expose politicians to the public	622	86.9%	151	62.4%	1	8.96	1.552	10	592	7.01	2.718	7	145
2	Inform citizens	599	83.7%	104	43.0%	2	9.20	1.322	10	578	8.09	2.182	9	96
3	Reveal party influencers	595	83.1%	49	20.2%	3	9.27	1.275	10	579	8.30	1.910	9	47
4	Spur discussions about corruption	522	72.9%	71	29.3%	3	9.14	1.447	10	509	7.75	2.378	8	68
5	Force politicians to focus on politics	381	53.2%	91	37.6%	4	8.55	1.910	10	357	7.56	2.189	8	86
6	Structure information	288	40.2%	56	23.1%	5	8.91	1.609	10	276	7.54	2.252	7	50
7	Lower influence of lobbyists	333	46.5%	16	6.6%	5	8.79	1.895	10	321	8.20	1.859	8	15
8	Support politicians in following their own opinion	212	29.6%	36	14.9%									
9	Make politicians work more	130	18.2%	18	7.4%									
10	Force politicians to stick with position	65	9.1%	21	8.7%									

Remarks: * % of "Is effect" calculated in relation to N max (716 cit., 242 polit.) of "knowing activities" in previous table; M=mean; SD=standard deviation; MD=median.

The most prominent effect is the "increased exposure of politicians to the public" (considered an important effect by about 87 percent in CS and about 62 percent in PS). In second place comes that the platform "provides important information" to its users (about 84 percent in CS, 43 percent in PS). Two further effects are in third place, since the priority allocated to them differs between the groups: the "revelation of party influencers" is supported by about 83 percent of citizens and about 20 percent of politicians, while about 73 percent of citizens and about 29 percent of politicians support the claim that AW's activities "spur discussions about corruption." Item 5 "forcing politicians to focus on politics" is backed by about 53 percent of citizens and about 38 percent of politicians. The two items ranked in fifth place ("structure information" and "lower influence of lobbyists") don't have a majority of support in either of the groups, but they are still more prominent than the rest (about 40 percent in CS and 23 percent in PS for item 6; and about 47 and 7 percent for item 7). Items 8-10, namely "supporting politicians in following their individual opinion", "making

politicians work more", and "forcing politicians to stick with positions" only find negligible levels of support.

When it comes to the assessment of effects (right part of the table), those who have recognized the existence of one of the effects are generally positive in its assessment.[62] This pattern is apparent in both groups. Yet, there are some differences. First, the median is at 10 for all items in CS, whereas it is at a lower level in PS ranging from 7 to 9. The aspect seen as least positive by politicians is the "exposure of politicians to the public." Since there is not much difference between CS and PS in their assessment of the effects, it is not displayed separately in a radar chart as per AW's other impacts.

7.5 Discussion

The following sections highlight the main effects identified by and with regard to citizens. Thereafter I discuss politicians' perspectives and then demonstrate insights generated from a sensitivity analysis, before outlining limitations.

7.5.1 Main effects produced from citizens' perspective

In contrast to the other sections, citizens' perspectives are illustrated on three different levels: (1) immediate effect and (2) mediated effects of the platform as well as (3) immediate effects of AW's additional activities.

7.5.1.1 Immediate effects

Citizens judge the communication and advocacy (political capital), and the community effects (social capital) of the platform to be most prominent. At a closer look, the promo-

[62] Due to the negligibility of the effects, items 8-10 are not listed in the assessment of effects, but the general pattern from the other variables applies in a similar way, although with some lower levels of positive assessments by about 1-2 points on the scale for both, citizens and politicians.

tion of advocacy is clearly the strongest effect, since both its components, namely "raising of awareness for neglected subjects" as well as the active "promotion of specific issues" by citizens are rated as "high." The mentioned main subjects, ranging from TTIP, to social policy, to ecology are very relevant for contemporary society, which makes this impact the more important.

In relation to communication and information, it is really only the "revelation of positions held by politicians and parties" that finds citizens' strong support. AW does not seem effective (enough) in promoting intense "political dialogue." Since there were no follow-up questions on this issue, it is hard to judge whether this is due to reluctance on the side of citizens, on that of politicians, or caused by some other factor, for instance the structure of the website. In contrast to this, it has been surprising to see that users thought AW was somewhat more effective in forming "communities of interest among citizens" (social capital). An outside observer would probably have expected to see the opposite, since the stress of AW seems to lie on getting in touch with politicians. More needs to be understood about the dynamics of citizen-politician dialogues and user-user communities to enable AW to promote one further without neglecting the other.

Another issue in relation to the informational aspects of the platform comes up through the observations on follow-up effects arising from available information, or rather the lack of follow-up effects. The results suggest that a higher amount of available information on AW does not necessarily make politicians stick with their positions (coherence). Nor does it make the overall line of argumentation or the evolvement of politicians' positions automatically more traceable. While the first aspect is mainly in the politicians' hands, AW could try to promote devices that better map the positioning and communication of politicians over time.

7.5.1.2 Mediated effects

At the level of mediated effects citizens highlight consciousness (cultural capital), literacy and engagement (political capital) as most important. When it comes to the levels of capitals held by CS respondents against ESS respondents and to citizens' own judgement of whether AW had an effect, political literacy stands out as unambiguously positive. All three, the level, the difference to ESS and the perceived effect are high on both, "forming an own opinion" and "communicating political positions."

The level of political consciousness (cultural capital) is also high, but the line of argumentation just outlined only holds for "knowing about political issues." The item "caring for others" instead is neither very different from the level found in ESS nor promoted strongly by AW. This could suggest that the two do not effectively belong into one category, or that improvements in the latter variable are hard to achieve or even unnecessary, since it is generally at a very high level in the population. In any case, it needs to be remarked that cultural capital aspects were not as central as political capital ones in this investigation, but their perceived existence suggests some interesting dynamics between different capitals that would be worthwhile exploring further.

The stimulation of political engagement through AW is also strong, but with very specific limits. Users of AW seem to prefer acting individually or in online communities rather than in formal organisations. While this is bad news for political parties, it is good news for engagement in general, since it seems to be effectively spurred by using AW. And this engagement, though maybe not formalised, certainly has 'real world' consequences. Petitions will lead to legislation in the best of cases and boycotting products can have significant effects on producers. The range of connected citizen actions shows that AW produces

some important spill-overs in particular to individualised forms of engagement, whether intended or not.

With regard to levering political influence, AW's ability seems to be restricted. Although citizens say that the influence of AW was high, the level and difference to ESS was not very pronounced. It is not entirely clear whether the effect would need more time to unfold, whether representative democracies have a 'natural cap', that is a threshold for citizen influence, or whether the perceived level of political influence of AW users may have been particularly low before they started using AW, which would magnify the meaning of an above average level of influence. There is some indication that the improvement of a previously low level could in fact have been the case, since there was some support among citizens for the claim that AW could spur interest in politics among those previously disinterested (immediate effects).However, whether they said so in relation to others or themselves cannot be sais based on the data, and the perceived effect was not very strong. Another hint in this direction would be that the share of CS respondents, who had "never been in touch" with politicians dropped from 39 to 17 percent after they had started using AW. Thus, although neither a real dialogue, nor considerably more influence for citizens has emerged, there is definitely more interaction between citizens and politicians. Longitudinal observations and thus more precise tracking over time could help untangle some of these issues.

7.5.1.3 Effects of organisational activities

In relation to AW's activities, citizens have highlighted five effects that can really be merged into three: The first and most prominent one is moving politicians into the limelight and exposing them to public scrutiny (items 1 "exposure" and 5 "make politicians focus on politics"). The second aspect is the provision of information (item 2), yet not only

with regard to political positions (as tested within the category of immediate effects of the platform), but also on lobbyism, misconduct etc. The third one directly relates to the latter in that the effect of AW does not stop at providing information, but furthermore "promotes public discussions" about these issues (items 3 and 4).

If read in conjunction with the effects that have found only little support, some further interpretations are possible: In the users' view AW's activities are seen as useful in "providing information" to citizens (item 2), but improvements might be needed in the "structuration of information", since this effect has not found major support in CS (item 6). AW's activities furthermore "reveal party influencers" (item 3) and "stimulate discussions about corruption" (item 4), but they are not effective (enough) in "lowering lobbyists' influence" (item 7). Finally, AW "exposes politicians to the public" (item 1) and this might "push politicians to focus on politics" rather than additional activities (item 5, but only weak support), but the exposure does not translate into increased effort of politicians (item 9, "work more"), "help politicians follow their own opinion" (item 8), or force them to be coherent in their position by "sticking to their claims" (item 10).

To close the loop, missing support for all the latter is consistent with citizens' position that communication between citizens and politicians through AW helps reveal politicians' positions, but does not increase traceability of political decision making or the coherence of politicians' positions (see immediate effects of the platform).

7.5.2 Main effects produced from politicians' perspective

In what I say below, it is not to be forgotten that some further effects than those discussed here have been backed by a fair proportion of politicians that participated in the survey. The discussion, however, relates only to those variables that were supported by the majori-

ty of politicians and/or could clearly be identified as a positive aspect (items above 5 on the 0-10 scale for instance).

When these criteria are applied, there are really only two items for which politicians recognise a positive effect by AW. In the immediate effects, it is only the informational aspect, more specifically the "revelation of political positions." All other items, with the exception of "raising awareness for neglected issues", which is at the neutral level, are nowhere near a positive effect from the perspective of politicians. When it comes to the mediated effects, politicians have a somewhat more positive take and ascribe "some influence" on several items to AW. However, no single influence is assessed as "high", which is why we can't effectively say that politicians see any profound impacts by AW in this regard. We find the second positive assessment in relation to AW's activities when politicians back the claim of "increases in the exposure of politicians" and assess its function as generally positive (median at 7). None of the other effects of AW's own activities is supported by a majority of politicians.

Although I had learnt about some similarly sceptical positions from politicians that chose not to participate in the survey but got into touch to communicate their scepticism, such a huge gap between citizens' and politicians' view in response to the surveys was very surprising to find. Some of those who did not participate said they had made that decision, because they found the platform unimportant, or because AW was tendentious in giving more room to the big and established parties.[63] Some others expressed that in their view AW was mainly used by a particular kind of people, with a specific political orientation and a narrow set of prioritised issues.

[63] Please note that all these comments are anecdotal evidence based on some emails I received and not a proper analysis. They might still be useful for embedding the analysis into the wider context.

The validity of some of these presumptions is relativized by the sensitivity analysis to follow.

7.5.3 Sensitivity of results

To test for the sensitivity of the obtained results CS data have been weighted to match ESS data in gender, education and age distribution. Adjustments for regional provenience of respondents have not been implemented, since that distribution was rather representative to begin with (and only a little modified after the weighting just referred to). Thus, after weighting CS data complied with ESS data on all variables typically used for assessing whether data are representative of the German population (European Social Survey, 2014a). Political orientation has been included as an additional variable into the weighting, since it stands to reason that results might be strongly dependent on the fact that conservative voters were strongly underrepresented in the CS data. A perfect match between CS and ESS on all five levels could not be achieved. One reason was the strong original deviation of CS in terms of gender and political orientation in particular, another the share of missing values, in particular for political orientation.[64] Overall, I have arrived at a reasonably high correspondence between the two data sets. For illustration the distribution of political orientations after weighting was about (excluding missing values): 36 percent conservative, 34 percent social-democratic, 9 percent communitarian, 12 percent green, 4 percent liberal, 3 percent nationalist, 2 percent left-liberal (and thus close to the distribution in ESS).

The balancing out of the data on the levels of age, gender and education has led to some other shifts in the socio-demographic characteristics of CS, of which, however, only one is

[64] While the share of missing data was about 3-4 percent for age, gender and education, it lay at about 14 percent for political orientation.

very prominent: the weighting has increased the share of people in the 20th income percentile (0-1,380 €; now 28 and previously 19 percent) against those in between the 40th and the 70th (2,041-3,280 €; now 23 and previously 29 percent) as well as those above the 91st percentile (above 4,960 €; now 7 and previously 11 percent). This has been accompanied by a concurrent shift towards single person households (now 37 and previously 27 percent) as well as an increase in the share of unemployed people (now 14 and previously 8 percent). There were only less prominent shifts in other socio-demographic variables.

When it comes to base levels of political and cultural capital and satisfaction, some further differences after weighting become apparent. For instance, the number of people who voted in the last election declined to about 72 percent (previously 82 percent). Visual inspection together with distributions test, such as the Mann-Whitney U test used before, highlighted other shifts. After weighting people in CS are: somewhat less in favour of cultural diversity and EU unification, and somewhat more mistrusting of the legal system as well as more dissatisfied with the national government and the way democracy works in Germany.[65] All other variables have remained largely unaltered (for example levels of solidarity or trust in people).

Effect tests have been re-run with the weighted data, which is then not only closer to ESS in terms of socio-demographic variables and thus to the general population, but also in terms of shares of political orientations. It is striking to note that all effects, the immediate and mediated ones of the platform (including forms of citizens' active engagement) as well as the activities of AW, remain remarkably stable. Of course some shifts have occurred, but most of them are negligible. If anything, effect sizes (with ESS or PS data as comparators) were slightly higher than before. The impacts of AW identified in this study thus do

[65] The general tendency remains unaltered; it is only the level that changes to some degree.

not seem to depend strongly on the particular structure of the CS respondent group (supposedly AW's user structure).

7.5.4 Limitations

The comparison between CS and ESS mainly served a heuristic rather than an analytic function for two main reasons: First, and although the likelihood is small, there might be some proportion of AW users in the sample of respondents to the ESS, distorting the separation between the two groups. Second, the user structure of AW is not representative of the German population in a number of regards. The difference between the two groups in variables supposed to be affected by AW could result because of different inclinations in related areas. Most relevant in this regard was the higher degree of political interest of AW users. It needs to be remarked, however, that this relates mainly to the difference in following the news or political contents when watching TV. The difference was not that pronounced though. The size of the detected difference was between 'small' and 'medium'. The difference in people who indicated they had "always voted" in national elections was even less stark: 82 percent of respondents in CS and 79.5 percent in ESS said they had done so. Any differences detected at all, have been uncovered through controlling for socio-economic characteristics and base levels of capitals, enabling readers to put results into perspective.

What is more, the sensitivity analysis of effects just outlined and the weighting procedure applied therein balanced out many of the differences. Nevertheless, perfect control is not possible in the present setup. There is always the risk of omitted variables that might have an important influence, but could not be controlled for. The direct responses of participants on the effective influence AW had on their capitals is generally more indicative of effects than the comparison between the levels of capitals in CS and ESS. The analysis performed

is thus somewhat more useful in detecting the difference AW has made for citizens, rather than for a comparison of the absolute levels of the tested variables against the general population.

Further limitations relate to some of the aspects mentioned further above: the social and cultural capital aspects could for instance have been pursued further, had there been more room within a single survey or a separate investigation. Most of the pursued issues, including the assessment of the levels of capitals, would also benefit from a longitudinal tracking to follow up on the cross-sectional observations presented here. It also should be remarked that the assessment of capitals has been performed in reference to several individual items (questions) rather than multi-item aggregative scales. Finally, a qualitative component in the form of interviews or focus groups would be a useful addition to get a better sense of who and how the users of AW really are. It would also help to explore why they perceive certain impacts and how they think they could be levered further. A reflective discussion with politicians would be useful to find out what they think about the huge divide between their own assessment and that of citizens.

7.6 Conclusion

The research has outlined how impacts of the promotion of political activity and democratic functioning can be studied and compared to existing data. Based on the analysis it can be said that political capital has grown through AW. The study has specified the main impacts of AW as perceived by citizens, which cover a wide range of areas. It has also served to highlight a large discrepancy between the impacts felt by citizens and the assessment of impacts provided by politicians. It is undeniable that AW users, at least those who chose to participate in the survey, exhibit some traits that differentiate them from the general popu-

lation and that some groups of people, in particular ‘conservatives’, seem to be underrepresented in AW’s user structure. AW should strive to become more inclusive in this regard.

All characteristics taken together though (socio-demographics and base levels of political and cultural capital as well as satisfaction) suggest that AW users are a group of people well inside the spectrum of potential voters for most parties. Together with the positive view of citizens of the platform this suggests a catalysing function of organisations like AW for democratic functioning in times of political polarisation and disenfranchisement between citizens and political parties. The stark contrast between the positive view of citizens and the more indifferent standpoint held by politicians outlines that such platforms might need to be taken more seriously and that we need to give greater attention to new forms of participation and their effective impact on society. I have shown how the capital based approach can be applied to capture and differentiate this impact.

8 Discussion

The specifics and implications of the individual empirical chapters have been discussed in great detail in the corresponding chapters and shall therefore not be repeated here. Instead, the discussion chapter does two things: First, it aims to provide some reflections on how the interplay between theory and empirical work has advanced the capital based approach. In this regard I will restrict myself to discussing the evolvement and interplay of different capital forms as well as the relation of capitals to measures from other approaches within the area of practical application. I will leave to the conclusion chapter a critical appraisal of whether the goals postulated in the introduction have been achieved (identifying how social impact could best be studied), and how the conceptual claims formulated in the theory chapter were confirmed or relativized (advantages of the capital based approach over other approaches). Second, the discussion will outline how well the work undertaken empirically reconciles with the principles and guidelines put forth in the methodology chapter.

The two goals of the discussion when combined point to its main function. It lies in generating an understanding of social impact in the style of the 'sociology of knowledge', that is by tracing the learnings that emerged in the process of deeply examining the social phenomenon. [66]

8.1 The 'multi-sites multi-capitals' setup and its uses

None of the three empirical studies has been completed as initially drafted, but all were extended and revised along the way, in particular as regards the emergence and accommodation of unexpected or complementary effects. Following the conceptual drafting per-

[66] See Keller (2011) for a succinct overview on the evolvement of the sociology of knowledge.

formed in my theory and methods chapters, I applied a multi-capitals lens to each of the organisations, yet one which started off with a very distinct focus on one capital. The study on community oriented housing (CO) was mainly drafted to examine social capital, that of RespAct (RA) to measure effects on cultural capital and that of Abgeordnetenwatch (AW) to assess political capital. As promoted initially, the focus was derived from the goals of the organisations as found in their mission.

Focussing on one type of capital per study proved useful in two regards. It had a confirmatory and a relativizing function: In CO and AW the focus was clearly maintained and other factors were mainly given attention to complement the analysis (confirmatory). In the investigation of RA in contrast there was a marked shift when it comes to the capitals enhanced the most through the intervention. I found that social capital components were at least as important as cultural capital components in judging the effects of the intervention (relativizing). At the same time, however, the formation of social capital by RA was rather clear and observable almost across all pupils alike. In contrast to this, the cultural capital components, specifically those relating to violence, were more multi-facetted and played out in different ways for different individuals. This is one reason why focussing on them was valuable despite the observed shifts in focus.

Alongside maintaining flexibility in the analysis, the empirical application of the capital based reasoning confirmed my proposed meanings of the different capitals. Most of the suppositions were confirmed in the empirical application: for instance that it would prove useful to narrow down the concept of social capital to relational aspects or to extend cultural capital towards including more basic modifiers of cultural composure and beliefs. This happened when aspects initially considered as belonging to the same category turned out not to work well together in the empirical testing and thus had to be allocated to anoth-

er one (for instance group cohesion as an aspect of social capital, and problem solving or individual influence as a political capital component). The discovery of incompatibilities in the empirical work served as a reinforcer of what I had proposed conceptually. There was a back and forth between the theory building and the empirical application. The theory as it stands now could not have been drafted in the same way without the empirical testing. The research was both explorative and 'generative' in the sense that it was based on an iterative process rather than following a linear one. There also were significant spill-overs between the three empirical cases resulting in a joint framework of capitals, which will be outlined below. This framework will need to be extended, since the three organisations I have studied provide only a snapshot of the activities of social purpose organisations. However, the initial composition of such a framework does not only confirm the use of the multi-site setup of my empirical work, but can also serve as a point of departure for future research.

Finally, it needs to be repeated that my proposition of the capital based approach does not negate the merits of the other conceptual approaches, nor is it supposed to replace them entirely. It is rather apt to complement them in calibrating the analysis. The approaches can be combined. I related to aspects of capability, cost-benefit rationales and satisfaction measures throughout the research and believe the combination of the perspectives, albeit superficial since not at the centre of the investigation, have outlined some significant synergies. Their co-occurrence has also offered deeper insights into where the individual approaches have their weak points empirically.

In the following paragraphs I sketch out how the capital based approach enabled the evolvement of complementary perspectives and enhanced the adaptability of the analysis to the gradual emergence of capitals and further variables of importance. I do this in rela-

tion to each of the empirical studies separately but with constant cross-references to the other cases. The individual perspectives are finally merged.

8.1.1 Social capital in the community context

As regards CO and social capital I started with a focus on social relations and networks, supposing that given the goals of the concept people in CO would have more and higher quality neighbour contacts. Following the dominant conceptions of social capital in the literature, these contacts would serve a 'capacity oriented' function by increasing mutual support in various regards, and an 'interaction function' by stimulating exchange between individuals or by promoting participation in joint events. Two other categories initially were part of the social capital focus, since I had encountered them in previous research: trust (Rothstein & Stolle, 2008)[67] and self-efficacy (Lillbacka, 2006).

However throughout the analysis and against the critique of an over-inclusive understanding of social capital (Fine, 2010; Flap & Völker, 2004) I came to question their allocation. Trust defined in the way encountered in most surveys, namely that of generalised trust (Bjørnskov, 2006a), as the very nomenclature suggests has little to do with individual relations held. It is a general attitude or norm, which is more similar to a pro-social attitude as examined in the study of RA than to the 'capacity building' or 'interaction' aspects of social capital. That it is undoubtedly important when looking at social capital, however, is clear from the close correlation between the two variables (Putnam, 2004b). I therefore maintained trust in the analysis but as a norm-based cultural capital component (see van Lange, 2015 on links between trust and culture).

[67] The link between trust and social capital is nearly omnipresent but particularly supported in for instance in the cited article.

Self-efficacy in the empirical testing of both the pre-disposition variables as well as in the attempt of combining several measures into an index of social capital defied all these attempts and remained clearly separate from the latter. The insights on self-efficacy were met by the concurrent observation that a survey question on joining forces to tackle problems coming up in the facility did not fit well with aspects of group cohesion, such as the 'sense of belonging'. I tried and failed to fortify the link between the aspects of problem solving and group cohesion by adding a question on the individual influence of the inhabitants within the facility in the second CO survey. Instead, my analysis came to suggest to regard joint problem solving and individual influence as two related yet distinct sides of agency and thus as a component of political capital.

Although the context clearly differs, the two variables are not at a far distance from other sorts of engagement as comprised by my concept of political capital and as encountered in the AW study. The similarity applies although within the latter setting engagement is more obviously political. 'Influence' is recognised as a separate political capital variable in the investigation of AW too. Both, joint problem solving as an example of 'collective efficacy' and individual influence on group processes have also found recognition in my study of RA. Collective efficacy as the ability to promote change and handle difficult situations is not only discussed with regard to informal group processes, but also within organisations (see Patras & Klest, 2012 for an application within social service organisations). The conceptual idea was thus also applied to CO, yet not in the form of a fine-grained scale but, due to its stepwise emergence, as a single question. Self-efficacy not only by wording is a reflection of collective efficacy at the individual level. Alongside individual influence the two concepts were thus regarded as belonging to political capital.

Thus, not only did my social capital components become clearer and more aligned through this stepwise process. The two complementary perspectives of cultural and political capital emerged from it too.

Right from the beginning measures to test satisfaction were implemented as an integral part into the research on CO. The interest was to see whether effects detected with regard to the identified capitals would be mirrored in concurrent increases in satisfaction. It was applied in a version directly tied to the intervention by asking inhabitants how satisfied they were with their own living situation. At a more general level participants were asked about their life satisfaction, which has been established as a standard measure of well-being (Diener et al., 2009). A more specific application referred to satisfaction or 'quality-of-life' in relation to health, following the existing suggestions that greater interaction could improve health status (Hanson et al., 2013). The value of the first two satisfaction measures was impeded by the latent variable of previous life experiences that made this particular target group very satisfied in general by way of comparison to the historical ardours endured during younger ages. A health effect in turn could either not be detected because it does not exist, or since a follow-up period of only one year was too short to make out significant differences, even if any were caused as follow-up effects of the higher level of capitals in CO.

In the chapter on CO I mentioned that the costs of operating CO compared to AL models have been assessed to add to the analysis. This had become an issue by way of interest of the organisations funding the research. The providers were interested whether CO models could be operated at lower costs than AL through the fact that service provision by staff is replaced in part by the self-initiative of the inhabitants. The results based on the average costs of several exemplary facilities were mixed. It was hard to establish a profound diag-

nosis, since there were considerable differences in costs between the two separate providers promoting CO, while costs in AL were at similar levels. What can be said with confidence, however, is that generally CO will not require higher costs of operation than AL. Judging the alternatives based on this account alone would not have been very useful. When applying a cost-benefit rationale, one would of course also have assessed the benefits produced. However, the links to this would have been restricted in so far as effects that could effectively be monetised, such as health gains, were not detected in the given timeframe. In conjunction with the higher levels of capitals measured, however, the lower or equal level of costs signals a profound advantage in the effectiveness of CO against AL.

8.1.2 Pro-social behaviour as cultural capital

The initial focus in the study of RA could rather clearly be derived from the programme's main goals. Several aspects of pro-sociality were classified as elements of cultural capital following the idea that cultural values are shaped by processes of "socialisation" (Goldthorpe, 2007), and that these values at a very basic level enable individuals to be productive within the wider social system. Productivity is not only embodied by educational performance and achieving social status, which is what most of the cultural capital literature has focussed on (Lareau & Weininger, 2005)—or the downside of the process, namely the production of inequalities (Small et al., 2010). Instead, there are more fundamental cultural codes and determinants of individual composure that determine whether an individual can strive in society. Many of these have been debated for long in psychology, for instance the embracement of violence or its reverse, namely non-aggression (see for instance Frączek & Zumkley, 1992 on socialisation in this context), self-esteem and other factors that shape personalities (Vignoles, Regalia, Manzi, Golledge, & Scabini, 2006).

At the beginning of my study, group level aspects were only included as an add-on, and group cohesion and collective efficacy had been conflated in one category. As the research progressed, and in particular through the interviews, it became clear that the latter two needed to be separated. This resulted in the formation of two separate categories: social capital for the component of group cohesion and political capital for collective efficacy, in line with the above discussion of CO. It also became clear that the research needed to be complemented by further components that initially hadn't been conceived at all, for instance interpersonal familiarity (social capital) as well as individual influence and engagement (political capital; see again the discussion of CO).

Each of them emerged as an individual level variable complementing the aforementioned group level variables. In some cases self-esteem, which was a variable considered from the beginning as a part of cultural capital, seemed to translate into a higher level of self-initiative by the pupils. The aspect of interpersonal familiarity in turn came out as important through unexpected statements by pupils describing how they had established bonds to other pupils they had not been in contact with before, or how they had started to see some of their peers in a different light. The latter aspects share commonalities with the 'importance' and 'intimacy' components in assessing the quality of people's social contacts within the CO study. It was not possible to implement all of these aspects into the quantitative analysis, but in the qualitative study they have been carved out. Despite their relevance, at least the 'influence' component played a role only for specific pupils, in particular those who had formerly been passive.

The last aspect, namely differences of effects depending on the individuals' personality, is another area in which knowledge was produced in the course of the research. The extension and systemisation of measurement variables enhanced by the capitals framework, to-

gether with the in-depth insights enabled through the qualitative analysis, were vital in getting aware of important effects. The discrepancy between goals postulated by the organisation but not achieved to the full, for instance promoting non-violence, and those effectively reached, for instance stronger bonds between pupils, suggests that an adaption of the mission, or a fortification of the intervention in certain respects would be necessary.

Satisfaction measures were more indicative of impact or effects in the context of RA than in that of CO. The fact that pupils were highly satisfied with the intervention and most of them would have recommended it to others is a positive sign in its own right. This is particularly so when combined with the statements of teachers who stressed the 'experience character' of RA and a sense of achievement that pupils felt when tackling and mastering challenges that were very different from what they usually encountered in school. Teachers also highlighted that pupils had learned practical skills, through the video work in particular, which might be of future use to them. The programme has thus, even if unintended, enhanced the formation of human capital and promoted capabilities in some respects. Another unexpected effect was revealed when pupils reported that the intervention, especially the boxing, had affected their view on gender roles. Tolerance and inclusiveness are thus further cultural capital components of potential relevance to RA.

The detected breadth of effects of the programme could either be interpreted as missing focus in the intervention or as an advantage. It certainly outlines that the intervention is effective in initiating change on several levels, in particular when some positive inclinations are present, but would likely not be effective in reverting deficiencies in a context of serious conflict.

The trait of broad but partly shallow effects would make it difficult to focus on extensions in pupils' individual capabilities as a measure of RA's impact. RA has only levered the capabilities of some of the pupils who took part in the project. However, it did produce effects for instance at the group level which are of relevance independent of changes at the individual level. Furthermore, even regarding the individual level it is not entirely clear how the detected shifts in capitals would translate into capabilities that enable future gains. The capability approach would be more easily applicable in a context where effects can more directly be translated into achievements expected to arise from them.

8.1.3 Political capital and democratic functioning

The contribution of AW to democratic functioning by levering political capital was inarguably the most focussed of the three studies. This was not only because engagement and influence in this case were literally embedded in the context of politics—generating a closer proximity to the terminology of political capital—but also since the instrument of an online survey offered the least room for probing issues or for the analysis to be comprehensive. The breadth of opportunity provided by the platform and the effects by AW's additional activities already demanded a multi-pronged analytic strategy, which is why the composition of measures was almost exclusively focussed on political capital.

The relevance of social and cultural capital was really identified only in the data analysis, where aspects of community and of awareness had been detected, which although they related to politics did not reconcile with the activity-based foundation of political capital. The attitude of "being aware of important issues", "wanting to care for other citizens" as well as the "sensitisation of disinterested citizens for political issues" were all interpreted as cultural capital due to their norm and value-based character. There are in fact literal parallels between these three aspects in terms of the standard scales applied for measuring

them.[68] The formation of "communities of interest among users", although placed in an online environment, refers to social capital formation.

Although not central to the analysis itself, other cultural capital measures together with satisfaction measures had been critically important in getting a sense of who exactly the user group of AW was and whether this would qualify, in the sense of restrain, the interpretation of detected effects. The user group in fact exhibited traits that made higher detected effects more likely, but a sensitivity analysis showed that results did not depend heavily on the specific group characteristics.

The identified parallels between measures across the different cases have underpinned the use of the capital based approach as an integrated framework, which is specified in more detail below.

8.1.4 Capitals combined

The applied capitals can be systemised based on the commonalities detected above. It will be seen that several capital components reoccurred in the empirical work, although with some variation. This holds in particular for social and political capital but less so cultural capital. Table 24 gives an overview of the variety of components assessed. The order in which individual capital components appear in relation to the interventions does not give an indication of the importance of that component to the intervention. Instead it has been chosen to highlight patterns across the three cases through aligning shared focuses. Two components stand out in that they appear in all three empirical settings: (1) variables relating to the ties between groups of individuals, in the form of cohesion or community build-

[68] Compare for instance questions on caring, solidarity, or tolerance in ESS (2014b) and those on "consideration for others" in personality studies (Weinberger and Schwartz (1990) in Dahlberg, et al. (2005)).

ing (social capital), (2) influence, either in relation to a more narrow (facility, class) or a wider social setting (political system). Three further aspects were of importance in two of the empirical studies: (1) interpersonal familiarity, as a dimension on its own or as a part of the quality in individuals' contact network (relevant in CO and RA); (2) collective-efficacy, as capacity for joint problem solving (relevant in CO and RA); and (3) engagement, not in the sense of passive participation, but of distinct actions and contributions to a cause or event (relevant in CO and AW).

Table 24 Multi-capitals framework

	Community oriented housing	**RespAct programme**	**Abgeordnetenwatch**
Social capital	Group cohesion Contacts Mutual support Participation	Group cohesion Interp. familiarity	Community formation
Cultural capital	Trust	Non-aggression Cooperation Self-esteem	Inclusion Consciousness
Political capital	Self-efficacy Collective efficacy Influence Engagement	Collective efficacy Influence	Influence Engagement Dialogue/debate Advocacy Literacy Information

What follows is a description of components exclusively occurring in one of the studied contexts. Participation as a measure of general social exchange was of importance in the CO context only. What is more, social capital in CO was differentiated further to not only account for aspects of community and interpersonal relations, but also for how ties would translate into actual support. Political capital in the context of AW in turn was spelled out

into further variables, two of which point at 'engagement capacity' (literacy and the availability of information) and two at actual 'discourse' (dialogue and debate between constituents[69], or advocacy for certain issues).

When looking at the spread of components across the table, it is evident that social capital and political capital solidify more as a concept than cultural capital. The reasons for this with regard to social capital are twofold. First, the literature and measurement of social capital are very advanced, thus offering a lot of guidance. Second, my narrow conception of social capital has helped establish focus in the investigation. In contrast to this, the study of political capital in the version I am focussing on, namely that referring to democratic functioning within the realm of politics or processes of engagement and influence more generally, is far less advanced. Yet, the just mentioned definitional anchors have enabled a set of measures to emerge around a number of shared components. Some further condensation and future development would be desirable and seems realistic as the classification of aspects around 'engagement capacity' and 'discourse' show.

The system of cultural capital in contrast is still quite scattered, and systemisation at least in view of the measures applied in the present three cases is hard to achieve without overstretching analogies. Two categories emerge nonetheless. Some of the variables surround 'individual composure' (self-esteem, consciousness of pressing issues), others the 'stance towards others' (trust, non-aggression, inclusion, cooperation). This is not to say that self-esteem for instance does not contain an element of comparison to others. What I suggest is that it is not 'held towards' others in the same sense as for example trust. Some further

[69] With regard to 'dialogue and debate' two aspects are merged, which have been held separate in the analysis of AW, since they were caused by different mechanisms: the communication between citizens and politicians (promoted through the platform) and the stimulation of public debate about lobbyism and corruption (promoted through AW's own activities).

interpretation might be possible: In the second category we can observe a transgression from the more passive stance of the attitudes of trust and non-aggression to more action-based concepts of inclusive or cooperative behaviour. Despite the more inchoate state of cultural capital, my 'extended' version that refers to a range of determinants of shared cultural conduct has proved useful as a frame of reference. This is becoming visible for example in the delineation of the above components from the formation of human capital or capabilities through new learning formats, the acquisition of new skill sets and the experience of achievement, which had been promoted by RA as a side-effect.

Its use is also underlined through the application of several cultural capital variables in the scoping for personality traits of the user group of AW. Commonalities are found between these variables and those considered in the study of RA, such as trust, solidarity, tolerance or unity/unification. This overlap in cultural capital can be observed although the studies are located at very different levels (group level versus national level) and measures have therefore been drawn from different disciplines: the ones for RA mainly from psychology and the ones for AW from macro level social science.

Table 25 summarises the components just mentioned, along with others that were considered complementary to the investigations or used as control variables. Satisfaction measures mattered in both ways, and some of the variables reoccurred across settings, such as life satisfaction (CO and AW) or satisfaction with the experienced intervention (RA and CO). Satisfaction could furthermore refer to one's individual position, as seen in the application of health satisfaction in CO, or to the wider social setting, as seen in satisfaction with the political system in place (in AW).

Table 25 Secondary effect variables and control variables

	Community oriented housing	RespAct programme	Abgeordnetenwatch*
Human capital/	New skills		
Capabilities	Achievement		
Satisfaction	Programme	Living environment	
		Life	Life
		Health	
			Political system
Costs-benefits		Resource intensity	
Cultural capital			Trust
			Solidarity
			Tolerance
			Unity

*=variables assessed as control variables, not as effect variables

A cost-benefit rationale especially mattered in CO, since two discrete alternatives were compared. Accounting for differences in resource intensity seems more advisable here than when trying to arrive at an absolute rather than a relative monetary value of a 'non-market good' (Dolan, Fujiwara, & Metcalfe, 2011). Resource intensity when seen in combination with the levels of capitals produced can provide a rationale for judging the effectiveness of two alternatives. It however only really adds value where the option that levers the capitals more strongly is at a significant cost disadvantage. In the present case, where CO required an equal or smaller amount of resource investment while producing higher levels of capitals almost throughout the board, it is of limited additional value.

The interaction of capitals discussed in relation to all three cases suggests that the capitals are reinforcing each other rather than requiring significant trade-offs. This is congruent with the conceptual proposition that all capitals, if in different ways, contribute to social productivity and that their assessment can serve as an effective unit for measuring the social impact created by social purpose organisations. Before I move on to challenges and

open questions with regard to the capital based approach, I provide some reflections on the link between the approach and the empirical research designs applied.

8.2 The 'Capital Based Approach' and research designs

The empirical research encompassed almost all methodological options in terms of research design that were outlined in the methods chapter, except for the experimental design. Table 26 outlines the main dimensions along which the executed studies need to be compared when it comes to methodology.

Table 26 Specifics of the empirical research designs

	Community oriented housing	**RespAct programme**	**Abgeordnetenwatch**
Design	Cohort (longitudinal & controlled)	Observational qualitative	Cross-sectional
Level	Group level (larger range, almost 'community' character)	Group level (small range)	National level
Participants	Beneficiaries: inhabitants	Beneficiaries & observers: pupils & teachers	Beneficiaries (2 types): citizens & politicians
Controls	Baseline accounts (retrospective); matching	Pre-post survey; tracing of change processes	Direct questions about effects; comparison to general population
Methods	Quantitative/mixed: Person-assisted questionnaire (& interviewer notes)	Qualitative/mixed: Interviews (& questionnaire)	Quantitative: Web survey
Duration	Long (> 1 h; performed twice)	Medium-long (40 min.: 30 min. qual. + 10 min. quant.)	Short (15-20 min.)

The table comprises six aspects: (1) the research design applied; (2) the level at which the intervention and the respective investigation was placed (group, community, nation); (3)

the participants of the research, with a particular emphasis on types of beneficiaries addressed in the measurement; (4) means taken to deal with and filter out confounding to ensure the reliability of results; (5) the methods applied and the types of data generated; (6) the duration of the data collection per participating individual. The listed aspects will be referred to throughout a discussion of how the capital based approaches and research designs have been linked depending on the specific contexts and which benefits or challenges this produced.

The variety of combinations of methods I arrived at was not motivated by a desire to illustrate the many outfits a social impact measurement design can take. Instead, the studies were drafted so as to best suit the particular setting, taking into account the way the interventions were provided and their specific goals. All this happened against three different versions of how the interventions shape their beneficiaries' lives: CO is a permanent offer and shapes the inhabitants lives throughout, AW is also a permanent offer but can be used by varying degrees of intensity and for different purposes, and RA is a one-off intervention that needs to unfold its impact at the single point in time when it is performed.

The scope of effects that can be assessed depends not only on the specific traits of the participant group, but also on the corresponding instrument necessary to gather the data. Since the CO questionnaire targeted elderly people, it was important for the survey to be person-assisted. In combination with the range of issues covered, the specific needs of the target group in the survey interview resulted in a long duration for completion, more precisely more than one hour on average. AW's target groups in contrast were accessed through an online survey, which had practical limits to how much time the survey could take. It was held within a range of about 15-20 minutes. The questionnaire and interviews on RA lay in between the two poles and at a medium to long level with around 40 minutes.

The most important consideration guiding the choice of the research design and its specifics concerned the mechanism of intended change the interventions pursued, and which comparator would be best to turn out effects. CO is one among potentially many living models for the elderly. At the same time it is a discrete, location bound offer and when inhabitants move in, they usually do not have a way of judging how CO compares to other living models, since they mostly have not experienced other forms of living, except living independently in a flat or a house. Thus, rather than through asking inhabitants directly which effects they thought CO had, the effect was arrived at by means of a comparison between the level of capitals reported by inhabitants in CO and the levels in a separate and more 'standard' living model (here: AL as a control group).

A repeated, longitudinal data gathering was performed to verify results given the impracticality of establishing a proper baseline on the capitals held by individuals right at the point when they moved into the facility. Retrospective accounts were used instead. A second assessment was also realised to solidify and extend insights by probing issues that emerge as particularly interesting or controversial in the preliminary analysis after first data collection. Finally, it has also been done to explore some variables that could provide links for future research, for instance the question whether greater social capital would extend from the 'community of inhabitants' towards the wider neighbourhood.

AW in contrast to CO is a unique intervention and there is no comparator that could be reasonably used other than outlining differences between citizens using AW and the general population. The fact that the latter and the user structure of AW differed considerably in some regards made the comparisons less robust than if a good comparator had been at hand and a precise matching could have been achieved. The ideal situation as regards matching is that it is not needed at all, which is the case when the control and intervention

group are almost identical. This in fact was largely the case in the CO study. Matching against the general population, or in other words making the structure of individuals in the intervention group 'representative' of the general population, is difficult since no clear rationale exists on how to control for confounders effectively. Even if the user structure were representative of the general population according to the standard socio-demographic variables (typically age, gender, education, and location), this would not rule out the influence of one or several latent variables that could still differ between the groups.

Considering political orientation in the sensitivity analysis of results in the AW study was one attempt of assessing an aspect that appeared to be one of the most relevant confounders. None of the assessed control variables had a profound influence on results, but this does not indicate that there are no other relevant but omitted control variables. Asking directly for the effects perceived was a good option in this case, with the caveat that these effects hold only for the specific target group and may not be generalizable. What is more, exact effect size is hard to come by when participants need to estimate the 'amount of change' with neither a robust baseline—which would again have required surveying and tracking people who had only just started using the platform—nor an effective comparator available.

In neither of the two cases of CO and AW is it clear how a randomised procedure would help solve the challenges. A random allocation of living environments is hardly conceivable, neither ethically nor practically. Nor does a forced use of an offer that is fundamentally based on intent and proactivity—one needs to want to know or learn about something to use AW, and one cannot simply be 'participating' in a passive sense—seem capable of providing more reliable insights than addressing an existent user group.

The distinctive feature of RA as regards demands towards the research design is that the intervention does not only seek to promote a property, composure or action but also needed to revert deficiencies in the levels of capital held. For instance, it is of course of value that RA was effective in fortifying and further promoting positive inclinations towards cooperative and peaceful behaviour. However, with an intervention such as RA it is the individuals that think and act violently, that is in an anti-social manner, or those who lack self-esteem where the real social impact potential lies. For all those exhibiting no lack one could always argue that they would have been just as well-off without the intervention and it is difficult to disprove this claim. Any attempt of tracking follow-up effects over time trying to spot positive turns in a life trajectory would open doors to so many confounders that no reliable attribution would be possible despite the effort.

What could be done effectively, however, was perform a careful tracing of shifts in thinking and composure, and whether and how these shifts were connected to the intervention. The analysis was based on the interviewees' own accounts, but also those of others (teachers and fellow pupils). Deciding on whether a shift exists and whether it represents a genuine change rather than a reproduction of 'lessons learnt' through an intervention, can best be judged when the researcher is able to probe issues and reflect them from several sides. This works best when qualitative methods are applied. These do not only help reveal change processes and the profoundness of the detected changes, but also help tie the latter to the programme components responsible for them. Besides, they enable the revelation of effect modifiers (Miettinen, 1974), that is stratification of effects among subgroups. Effect modification in RA was observable as regards pupils with behavioural difficulties, for whom an enhancement of non-violent action was more restricted than for other pupils. It also became apparent in relation to passive pupils, who benefitted more than others in terms of increased self-esteem.

From the observations above it follows that cultural capital is not only challenging through its potential breadth of components, but also through the circumstance that it may comprise positive and negative pre-dispositions. Negative pre-dispositions are hardly imaginable with regard to social or political capital. People may have no social contacts or be excluded from communities, but they would normally not actively seek to destroy their social links or undermine interactions within communities. Along the same lines: some people may be disengaged or uninvolved politically, or within another realm where aspects of activism and engagement matter. Only few people would however try to behave in a way that would crowd-out their own and others' engagement. Political extremism might be one of the few cases in point. But in that case the root cause and thus the target of any intervention seeking positive impact would be the cultural attitude that motivates extremism. The engagement of an individual holding extremist views seeking to fight the engagement of others would only be one of the symptoms of that attitude.

The last example outlines that although capitals should be considered in parallel, the mechanisms by which they contribute to social productivity, or deteriorate it, should be carefully distinguished. Scholars have for instance argued that social capital can have negative effects when norms adverse to commonly shared standards are imposed on group members within networks marked by strong ties. Youth gangs or the Mafia are given as examples for the negative consequences social capital may have (see Ledogar & Fleming, 2008). The very point in this, however, is that norms and thus cultural capital play a separate role herein. The networks in which individuals and the cultural capital they hold are embedded will of course play a role for how the members of that group are socialised. But in the given example it is the defective norms, and thus a lack of cultural capital, that cause the "public 'bads'" (Portes, 1998, p. 18). The social capital instead is only responsible for spreading and imprinting it onto group members.

This reasoning does not only fortify the distinction between capitals within the capital based approach. It also demonstrates some potential links to the general literature on the different capitals, which the observations made within the capital based investigation of social impact could contribute to. Further links will be outlined in the conclusions. Before I close with these, I present some unresolved issues in measurement and design that would need to be addressed in and through future applications of the capital based approach.

8.3 The 'Capital Based Approach' and organisational learning

Due to the restrictions in time and capacity that come with executing this research as part of a doctoral thesis, I was not able to initiate and accompany a process of organisational learning and change that the insights derived from the analysis would have allowed for. I have provided the organisations with the results of the analysis, partly in a one-on-one format, but I have no detailed insights on how the results have been used by them. Therefore, the following considerations outline possibilities for organisational learning as a consequence of the executed research rather than a factual documentation of occurred lessons.

In future research capital based impact analyses should be tied more closely to the organisational learning and change they enable in pursuit of increased organisational performance. For being capable of doing that the model of social productivity would need to be extended vertically, that is in relation to a more fine-grained understanding of the exact links between organisational models, strategies and activities and produced capitals. In thus far I grant that in view of organisation theory my research in many respects focusses on the question of *what* works and—as research on evidence-based management more generally (see Langley, Smallman, Tsoukas, & van de Ven, 2013, 4f. for this argument)—is limited in providing definitive answers to the question of *how* organisational change and its desired positive effects on performance are to be achieved.

Despite the above restrictions it remains that by tying organisational activities and the formation of capitals closely together the capital based approach to social impact is likely more apt to support organisational learning than the alternative conceptions. Thereby it might help close a perceived divide between 'understanding' and 'measurement' sometimes articulated in process theories of organisation (see Hernes, 2009, 143f.). Impact measurement along capitals may enable organisational learning in the traditional sense: It could lead participating organisations to provide a *different* response under the condition of the *same* stimulus (no shift in mission or target group) (compare to Weick, 1991). Alternatively organisations may adapt their organisational mission in view of a seeming impracticality of changes necessary to achieve initially postulated goals.

For judging the likelihood of such changes in social purpose organisations, we need to take the central importance of mission into account. The strong focus on mission could have two conflicting effects: On the one side it might increase the organisations' willingness and flexibility to adapt their inputs and activities in order to increase performance in view of its goals. On the other side, a strong organisational imprinting can lead to persistence and reproduction (Schreyögg & Sydow, 2011), in particular if change is not appreciated or hard to communicate towards external stakeholders, such as investors or regulators. In the following I will abstract from internal and external sources of resistance to change and assume organisational embracement of change when it helps aligning mission and produced capitals, and/or levering the latter.

In particular the research on the school intervention, due to its qualitative design, offers opportunities for organisational change that could be beneficial to the social enterprise executing the programme. It is the case where the link between performed activities and achieved effects is outlined most explicitly. While the boxing and the video work had dif-

ferent effects on the participating pupils, the video work unexpectedly had made a stronger impression on them, which is why it seems beneficial to extend it. There have also been positive effects of the reflective discussions and the trust games which were not consistently applied in each of the investigated applications of the programme. The organisation would need to maintain more consistency in this regard if it wants to fully harvest a breadth of effects on cultural capital. One of the strongest deficits uncovered in the analysis was the lack of shifts the intervention made on the cultural capital of male pupils with a high degree of behavioural problems. These deficits could be coped with in two ways. First, the organisation might consider employing trainers that as stressed by the pupils themselves would be able to make a stronger impression on them (martial artists, youth workers or repentant ex-offending youths). Second, the organisation could partner with other organisations that specialise on at risk male youths who exhibit a high degree of aggression, such as "(ex-)convicts helping youth" (Gefangene helfen Jugendlichen (GHJ), 2017).

A change in mission or in the communication of the intervention's goals might be an alternative to the latter changes. I have detected pronounced effects with regard to the self-esteem or pro-activity of shy pupils, which do not play a major role in the postulated goals of the intervention as of now. The same goes for the enabling effects on the inner-class social capital or the detected, yet not fully explored, aspect of challenging gender prejudices. All three are undoubtedly positive but point into different directions than what is currently in the organisation's focus. The limited ability to affect older male 'trouble makers', in combination with the insight that there were it least some, yet younger pupils with behavioural difficulties who benefitted from the programme, might point to the necessity of sticking with the organisation's original focus on younger secondary or elementary school pupils.

Though more superficial, since quantitative investigations are less capable of displaying causal lines of effects, the providers of community oriented housing have two major sources of potential change in organisational strategy based on the performed research. The first one relates to the realisation that community oriented models do not help promote intimate friend-like or family-like structures visible in the non-superiority of such models as compared to assisted living when it comes to emotional support in particular. The providers could consider introducing sponsorships between existing and new inhabitants that might enhance interaction on an individual level. The support of other formats, such as conversation circles might be another vehicle to promote such relations. However, it would be important for the organisations to not be prescriptive and avoid crowding-out self-initiative that is at the heart of the community oriented concept.

A second prompt lies in the observation that, despite the fact that they are superior in this regard to assisted living, there is still room for improving the inclusiveness of the CO models. An important question to answer before trying to take any measures on this account would be why some inhabitants choose to be rather isolated despite choosing to live in a model that is community oriented. Of course there will be people that prefer being on their own and it would certainly not be reasonable to try to draw these into community life. Yet, in particular when regarded on a regional or national scale, there are many old people who live in forced isolation, who could benefit more from the community oriented models than those who do not want to be part of the community to begin with. And then there might also be cases of individual isolation that result from a loss of contacts over time due to casualties among their contacts. Ways of maintaining connections to the larger community might need to be sought to counteract such developments. An alternative to making organisational adaptions is of course the recognition that community oriented models are restricted in these regards.

A third aspect is relevant only for the facilities with a focus on multi-generational housing. Counter-intuitively intergenerational contacts or support did not seem to be a major factor contributing to the higher social and political capital in the facilities. It would need to be explored further if the effects of younger people in the facilities are more indirect. They might bring 'more life' into the facilities merely by their presence, or be engaged more in communal activities than in individual one-on-one support or exchange. However, in deeper investigations it might also turn out that they are not centrally important for enhancing capitals at all, which could obviously lead to major shifts in the organisational model away from focussing on a multi-generational mix of inhabitants.

The prompts for organisational change and learning are most superficial for the political platform. The reason is the cross-sectional design and the absence of any reasonable control group in the study. Yet, there are still some potential learnings to be made for the organisation based on the results of the levered capitals. Most striking is the potential for a fortification of the communication strategy towards MPs. While arguments for the use of the platform currently seem mainly based on appealing to MPs' functions as representatives of the population and the corresponding duty of answering their questions, the positive effects on citizens' political capital could serve as a different motivating force. The same goes for the realisation that the platform along with other forms of online participation seem to replace formal organised engagement, in particular in established parties.

The contrast between citizens' and politicians' views has been most pronounced in relation to the own activities of the platform, most of which focus on exposing and counteracting misbehaviour or dubious practice among politicians. There is certainly a tension between involving more politicians in increasing political transparency by using the platform and taking a critical stance towards MPs' actions. Obviously the organisation will not want to

trade off the one against the other, in particular since citizen users have a uniformly positive image of both fields of activity. Yet, the present tension might hinder effective scaling of the reach of the platform and a clearer separation of activities might be advisable.

As regards the 'exchange platform function' of AW there is at least a positive tendency among users in judging its capacity in evoking interest in political issues among those typically disengaged. As of now this has been no explicit priority of the organisation and thorough recommendations of how it can be advanced would require more research. A potential lever could be found in harnessing potential spill-overs from online activities on petition platforms or social media. The latter are closer to citizens' every-day realities and probably more broadly used, but do not offer the broad set of coordinated offers and activities the platform provides. The organisations could try to more actively attract first time or infrequent users of such 'low threshold' online forms of political engagement.

The points discussed show that when organisational change and learning is sought, the assessment of impact must not be a one-off exercise, but needs to be repeated—if in an adapted fashion and with a specific focus on the new or different responses given by the organisation—in order to enhance performance in an evolutionary process. As also seen above some interventions, due to their nature and the modes of assessment they allow for, are more suitable for producing insights on necessary organisational adaptions. This has been and will continue to be a source of shortcomings and challenges of impact studies, which I want to discuss in greater detail in the following.

8.4 Shortcomings, challenges and open questions

Before discussing some of the things that did not work well in the execution of the research I want to forestall that the three studied interventions have been chosen not although they

were challenging to assess, but exactly because of this property. My ambition was to test whether and how the capital based approach would be capable of demonstrating impact, not only in each of the three areas most neglected in the current practice of impact measurement, but also with regard to interventions that would typically be treated as very hard to assess for their very grounding in these areas (social, cultural and political capital).

The absence of a control group can be highly problematic in assessing social impact and has affected the analysis of AW and the RespAct programme in particular. The arising problems are independent of whether control comparisons are impossible to realise since no alternative intervention exists (AW) or the recruitment of a control group is very difficult, since there are few incentives for such a group to participate (RespAct). Purposefully sampled populations will be prone to selection effects, leading to sampling bias which needs to be critically acknowledged. Without a control group there is little that we can do to account for such a bias. While a clear source of fallacy in the medical testing of drugs, psychology (Cronbach, 1990) or macro sociology (Berk, 1983) for instance, selection effects will be of lower if any severity in the assessment of organisational social impact. Take AW as an example: It has become evident that a certain group of people, such with some inclination towards political issues, will be more likely to use the platform than others. In how far does that restrict the assessment of the impact of the platform? Only in a restricted sense: It of course needs to be acknowledged that findings cannot be extrapolated to the general population, or as mentioned in the section on organisational learning to particular subgroups such as people previously disengaged with politics. What is more, the analysis needs to include direct questions for perceived effects among users ("Has AW influenced X, and if so by how much?") rather than only indirect ones ("What is your level of X now that you are using AW?"). As long as these two points are adhered to in the col-

lection of data and the interpretation of results, the claim of having assessed impact can be maintained.

A different strategy has been applied in the case of RespAct. The primary strategy of asking for impact has been an indirect one using questions such as "Have you behaved differently in regard X recently?" Only in a second instance have I ensured by means of direct questions whether any perceived change could be ascribed to the pupils' participation in the programme. This might still invoke social desirability bias, meaning that in the interview situation respondents could have been prone to tell the interviewer what they thought would please him. Desirability bias might be bigger in interview situation than in responding to a questionnaire (see for example Crowne & Marlowe, 1960 on how to measure this very bias). But the discussion between the interrogator and the respondent also helps understand other potential sources of behavioural change, such as influences by single events, family or friends, better than any pre-post pair of questions could, including a comparison with a control group.

Two shortcomings remain. First, pointed questions for changes in certain dimensions may themselves lead to other response biases. Second, the assessed shifts in capitals are relative and due to the lack of a comparator cannot be given in absolute terms. This would hardly ever change in the case of the political web platform, since in the absence of similar interventions it would always need to be assessed against the non-existence of any intervention. As long as the intervention does not address a population nearly congruent with the general population major restrictions on generalisability will remain. These can be controlled for by statistical procedures, as done in the elderly living facilities case. Yet corrections of bias are easier when two fairly similar groups can be compared, as was the case when comparing CO and AL, than if representativeness to the general population is to be achieved (see

for instance Stolzenberg & Relles, 1990 for a critique of modes of correction for nonexperimental data). In the school intervention case the 'relativity' of measures might change if a number of similar studies were executed using my framework, method and assessment tools. In other words the issue might decrease through a process of benchmarking and an evolution of assessment standards for the capital based approach.

In view of a lack of established standards I have tried to make use of established scales and designs throughout the research, but also performed a large amount of reinterpretation, condensation and calibration of measures in order to match the needs of the interventions more accurately. I have also done so for making them practically relevant to academic and non-academic analysts of social impact and for enabling the implementation of measurement results into continuous organisational reporting on social impact achieved. The latter is demanded by both, the imperative on organisations to "manage for impact" (Cronin & Dearing, 2017) and the necessity for evaluating intervention "success" (Edmiston & Nicholls, 2017)[70] within new public contracting practices such as social impact bonds. As already suggested in the methodology chapter and confirmed in the empirical application, many available repertoires of measures are either too specific and deep (psychology), or too broad (social science surveys) to be directly applicable within the measurement of social impact. What is more, the available measures often lack a conceptual grounding that would ease classification, such as the framework presented here, and application across fields of activity. On the other side there is practitioner driven indicator development that encompasses a whole range of activities (Social Value UK, 2017), but this often is not linked to the wide range of scientific measurement scales available—supposedly for the impracticality of the latter. I suggest the capital based approach is capable of providing a balance between the two, which I hope the cases themselves and this comparative discus-

[70] See for the U.S. use of „Pay for Success Schemes" for referring to social impact bonds.

sion have shown. It would require a far more extensive application to prove its worth in this regard.

Despite the guidance taken from existing research and an initial idea of the multi-capitals framework sketched out in this chapter, my research has been marked by a number of redundancies and dead-ends in the collection of data. Some dimensions did not solidify and had to be dropped, or had to be split or rearranged into different constructs. This thinned out some of the assessed dimensions so much that they had to be treated as single questions rather than scales measuring an underlying construct. Single questions are obviously weaker in capturing an achieved leverage in capitals than scales. The elderly living facilities case could benefit from two waves of data collection, but these issues still prevailed. While the second wave helped clarify some open questions from the first round, it also produced new questions or incomplete perspectives that could not be fully addressed with the available data. The inchoate state of the data applied in particular to the investigation of the political web platform, where a more efficient structuration and potential aggregation of dimensions could have been achieved in a second data gathering, which however was practically impossible. The same can be said about the minor quantitative part of the research on the school intervention, which in contrast to the qualitative interview guide could not be extended or adapted, at least to some degree, as the research progressed.

A further problem of the research was that fertilisation and the evolvement of insights across the cases was restricted in the execution of the research and could only take place during the data analysis. This applied since the empirical data collection in all three cases has been performed largely at the same time. Although data collection has not taken place literally in parallel, the analysis in none of the cases had been finished before the conceptualisation and data collection of the other cases had started. This led to a reproduction of

some challenges and restrictions in the explanatory character of the data rather than a stepwise resolution of such issues. If I had to re-execute the research many of the applied questions would be leaner and better concerted. There would also be more purposeful coherence in the capital components occurring across the cases, which in the actual research process only emerged at the very end. Such uncertainties and the necessary solidification of knowledge was however an unavoidable property of the exploratory, multi-site investigation and could therefore hardly be avoided. The parallel rather than sequential execution of the research was owed to the practicability of the execution of the research and restrictions in the time that could be invested into it.

As regards open questions it remains for me to outline further fields in which the application of the capital based approach could contribute to a standardised canon for its application, including an extension of the multi-capitals framework. I have previously named relevant fields of application for the capital based approach, which I want to repeat here, arranged by the respective main capital. A prime place for assessing social impact by the formation of social capital would be community development initiatives of many sorts, be they performed within a specific geographical setting such as a city quarter or online. Establishing social ties and thus studying social capital would also be instrumental in assessing interventions that try to promote the integration of refugees, or such that tackle the reduction of individual isolation of many sorts, for instance in old age or when people get unemployed. Political capital is relevant in any intervention fostering literal political engagement of citizens, but it may also be applied to the enabling of consumers within the range of consumer protection, or with regard to cooperative movements in which people come together to promote a joint cause. A focus on cultural capital formation can in turn be applied to activities of different kinds that influence identity formation, spanning from a

dignified self-image of women to encouraging youths from working class backgrounds to seek higher education.

An important criterion in any research seeking to explore impacts by means of capital formation will be to find a balance between focus and breadth in the analysis, as proposed theoretically and demonstrated empirically. A close match between organisational mission and the capital components in focus should not crowd out attention for complementary measures.

While the discussion chapter has focussed on a condensation of knowledge produced by the empirical investigations against the proposed methodological framework, the conclusions to follow will critically appraise the performed research in relation to theory building, and the advancement of future research. They will also provide managerial and policy implications of the research beyond designs and measures.

9 Conclusions

This thesis meant to address two connected research questions. The first was concerned with how to define social impact in a way that allows us to understand and meet the challenges it produces, most notably that it is often considered unfathomable. The second demanded the exploration of how, based on the theorising, the empirical measurement of social impact could be enhanced.

Essentially this work was thus about transferring Henry David Thoreau's imperative "be good for something" from the individual to the organisational level, and about finding an effective way to address the question: "*What* are organisations good for?"

I outline here how my research has performed in this regard. In this I will focus more on the issue of theory. The empirical part follows from the latter quite naturally. I then go on to discuss limitations and implications for policy and practice, and give prompts for future research.

9.1 Contributions

My research has shown that it is beneficial to place social impact within a theory of 'social productivity' on processes of production that happen within society and contribute to positive social outcomes. Social and political stability, equity, or inclusiveness are examples of such outcomes. To arrive at them, actors employ a variety of capital factors within a process of production, similar in many respects to economic production. Several capitals play a central role in the process. These include natural capital (the environmental ecosystem), human capital (skills and knowledge), social capital (relations), cultural capital (norms and values), and political capital (engagement and influence). Many institutions and organisa-

tions contribute to the creation or preservation of these capitals, those pursuing a ‘social purpose’ in particular.

While sharing similarities with economic production, the chain of social production also differs in important ways from it and produces challenges in the assessment of contributions to social productivity. First, it typically involves a multitude of capitals at the same time, which produce a range of outputs that in different ways lever social outcomes. This happens against the background of complex social realities, which are prone to a variety of confounding factors. This makes attribution between individual actions and final social outcomes difficult. Another complication arises in that in contrast to measures of economic productivity, which rely on market prices and are therefore considered to reflect the value of welfare efficiently, social outcomes are fuzzy and hard to define—not to speak of break down into measureable components. Mainly as a consequence of the latter aggregation of single measures is hardly possible. It may not even be desirable, since the different dimensions may represent different sorts of value.

Instead of trying to grasp social impact through encompassing accounts along several lines and with regard to highly aggregated concepts, the social impact of organisations can be understood and assessed by accounting for organisations’ contributions to the formation of the above capitals. Especially social purpose organisations can be understood as capital producers that lever social productivity. Such an approach has several advantages over existing approaches to social impact which are mainly oriented at monetary costs and benefits, satisfaction, or capabilities.

The existing approaches, despite their great merits, often do not compensate the challenges social impact poses and partly exacerbate them due to their conceptual foundation. Capa-

bilities for instance in many cases have remained abstract concepts and are therefore hard to measure. Costs and benefits as measures of impact face the same challenge, but for the opposite reason. They focus on measures that can be monetised and often disregard aspects for which this is hard to do. Both examples show that they are restricted in addressing the measurement problem. Life satisfaction in turn is a measure composed of many different and subjective feelings. This adds to the problems of attribution and aggregation. All three come with the additional challenge of being oriented at the individual and her preferences, utilities or functionings. This specific foundation of the approaches may result in neglect for trade-offs that are located at the group or community level.

The capital based approach is meant to mitigate these issues. First, through being grounded in a variety of social sciences, it can relate to a rich body of theory and empirical research—including measures—ranging from social psychology to comparative macro sociology. This helps address the measurement problem. Second and based on the latter, it can distinguish several capitals and split those into more fine-grained components. This helps address the aggregation problem. Third, the capitals are located at a closer proximity to the organisational activities performed and provide focus in the execution of an impact analysis. This helps address the attribution problem. Fourth and finally, the capital based approach is derived from the macro level perspective of social productivity and the employed capitals often do not relate to the individual in an isolated fashion. Instead they point for instance to her relations (social capital) and positioning towards others (cultural capital), or to her activity and activism within her social setting. By its collective grounding the capital based approach links up to the societal legitimisation process of what is regarded as socially productive. This helps address the foundation problem.

Based on this theoretical proposition, the research has contributed to spelling out how the capital based approach can be applied in the empirical measurement of social impact. I suggest that impact assessments apply a 'multi-capitals setup' that is capable of providing both, focus and breadth in the analysis. Focus can be achieved by marking one primary capital within the analysis. The most important capital shall be derived from the goals found in the mission statement of an organisation. A match between core goals and a particular form of capital should be possible for most organisations at the level of the specific intervention to be investigated. Breadth comes in through considering several other capitals alongside the main one. Although they might not be of the same importance as the latter, effects on a variety of capitals are likely to coincide in most interventions.

As alluded to in relation to theory, the operationalisation of measures for the capitals should be guided by existing research. It is unlikely, however, to find measures that can be applied within social impact measurement right away. Considerable effort needs to be put into revising and tailoring measures to match the nature of the intervention and that of its target group. Measures that are relevant for the direct beneficiaries should receive priority, since social purpose organisations foremost exist to serve them and only in the second instance other stakeholders.

The multi-capitals setup of the research can effectively be combined with research designs of all kinds, ranging from experimental studies to systematic qualitative analysis. A review of current practice has shown that their application is far from common in the measurement of social impact. Combinations of the approach and research designs should be drafted so as to best suit the intervention under study. Particularities of the target group or the supposed mechanisms by which the intervention creates impact are important criteria to take into account. As has been shown the conceptual approach in itself is capable of addressing

the existing challenges in the measurement of social impact. It is particularly useful with regard to decreasing the measurement, aggregation and foundation problems. Rigorous research designs in turn are important to further lever its capacity in mitigating the attribution problem.

My research has demonstrated how an empirical application can look like with a specific focus on social, cultural and political capital, for those are the most challenging capitals to assess and also the ones most disregarded in the current practice of social impact measurement. The 'capitals framework' derived across the cases has provided a guiding reference in systemising capital components as measures of social impact and performed a consolidation of the capitals, their components and specific measures, which should be probed and extended in future research.

To summarise, my research relates to the concept of economic productivity but extends its relevance beyond the economy by addressing broader societal transformation processes. It connects a variety of social sciences, including sociology, political science, culture studies, psychology and economics. The capital based approach I propose taps several detached research strands on capitals and integrates them into a coordinated concept that aids our understanding of organisations' social impact. My research responds to the question of what organisations are good for by considering how they contribute to capitals that are vital within the system of social productivity. Thereby the approach explicitly takes into account the inherent productive capacity of norms, networks or virtues. It posits, even if two strategies result in the same state of individual functionings or utilities, the one option shall be preferred that unleashes future potential at the collective level, and thus suggests future productivity. The reasoning behind this rationale is that a positive value is ascribed to the capitals (be they social, cultural, political or otherwise) per se. The approach sug-

gests that capitals enhance society's capacity to act and that therefore capital formation equals social impact.

9.2 Relations between the different conceptual approaches

The way by which I have made a contribution to theorising social impact is that of 'noncoherence' (Locke & Golden-Biddle, 1997), that is I have approached the subject in a way which in parts is in contrast to existing theory. I have done so, not because there are no common and useful points in existing approaches or because I see them as incommensurable, but because existing approaches are marked by gaps on social impact that could not be solved despite their long tradition. They can still be more useful than the capital based approach in some areas. And a concerted application of the different approaches will be needed in some fields of application. I will outline in the following when either is the case.

As discussed in the methodology chapter the capital based approach or in fact any more complex attempt to capture social impact, may not be necessary at all for particular interventions. By relating to previous research (Ebrahim & Rangan, 2010) I have outlined that certain activities would be sufficiently assessed by looking at performed activities. This is the case when both, the impact chain and the operational strategy are "focussed." Examples have been discussed that ranged from soup kitchens, to vaccination initiatives to drug withdrawal programmes. The majority of activities performed by social purpose organisations, however, will be more complex and would thus benefit from the capital based approach.

I want to stress that complexity as assessed in relation to the given examples only refers to the provision of the service and the impact mechanism underlying it when the intervention is considered in isolation. Attesting a low complexity does not suggest that managing a single food kitchen or an organisation running several food kitchens is a simple task. The contrary is likely true, in particular when considering the addressed target group. Soup kitchens often serve people without voice, a personal safety network or even shelter, and it would be ignorant to assume that it is easy to deal with such individuals. Yet, the means by which a soup kitchen helps this target group *primarily* is based on a simple need-response pair (hunger-food). If we extended the view beyond the immediate intervention and took into account potential follow-up effects, this might change. For example if we supposed that the same people use the service repeatedly and do not only visit the soup kitchen to get something to eat, but also to chat with or seek help from other users or the staff running the organisations, we could find the social, cultural or political capital of the target group affected. This would justify the application of a more comprehensive impact assessment by the capital based approach. The main message behind stressing the preliminary step of checking for complexity is to be clear on the mechanisms by which an organisation supposedly works and to decide whether more complex modes of assessment are necessary to test the effects produced by the intervention.

I have outlined the benefits of the capital based approach over the other three conceptual approaches to social impact. However, even when the application of the capital based approach is indicated, it might be useful not to apply it in isolation. Capital measures can and should be combined with measures from the other conceptual approaches where meaningful. Benefits, satisfaction and capabilities will matter in different ways to interventions and provide additional explanatory potential, especially when they are combined with capital measures.

An additional focus on individual freedom in realising one's capabilities might remain necessary in contexts marked by multidimensional deprivation (Alkire & Foster, 2011a), for instance in regions where people suffer from poverty. The enhancement of a choice between multiple options might be the main source of social impact created by an intervention. The enabling of choice through an intervention may have unexpected, potentially deteriorating, effects on the level of capitals but be valuable in itself. Blattman, Emeriau and Fiala (2017) have studied the effects of an entrepreneurship funding programme in Uganda with beneficial effects for employment and earnings. Instead of rewarding the initiating policy makers, however, the funded young entrepreneurs chose to support the opposition. Blattman and colleagues interpret this as a sign that voters, when gaining in material terms, become also freer to express their real political preferences publicly. On the other hand swaying voters might have counter-effects and lead to polarised public debate (potential decrease in cultural capital) and social instability (potential decrease in social capital). The element of freedom might help weigh effects against each other. Decision makers would then likely prefer a state with more personal freedom, and less suppression for that matter, despite the distortions in capitals that could come with it.

A cost-benefit rationale and an analysis of satisfaction measures would be useful in other cases, in particular where clashes between these two and capitals occur. Take for instance an initiative furthering social capital within a community setting. The resulting more intense utilization of public places by residents of a neighbourhood might cause asset depreciation—and thereby decrease monetary benefits—or inhibit the life satisfaction of other dwellers in the vicinity. Both would need to be offset against fortified social capital.

Satisfaction measures can generally be useful complementary indicators in judging whether an intervention is not only objectively beneficial, for instance as assessed by increases in

capitals, but also ‘felt to be’ beneficial by the individuals affected. There is a reason why subjective satisfaction and quality-of-life play a role even in health care, where intervention effects can be measured metrically and in a standardised way. ‘Quality adjusted life years’ for instance move beyond the investigation of therapeutic results and are used to display whether a treatment effectively leads to improved patient well-being (for example Fanshel & Bush, 1970; Nord, 1994, and Dolan, 2011 for an extension of the concept with a focus on happiness). A lack of differences in satisfaction, however, is not always suggestive of no substantial difference between the effects of two interventions as per my investigation of the elderly living facilities. Nor does decreasing satisfaction automatically mean that an intervention is ineffective. Recall the school truancy or addiction examples in my discussion of limitations to the application of satisfaction measures.

I have outlined earlier that where direct and indirect monetary benefits are (1) substantial areas in which impact is expected or (2) where these can be derived with a high degree of confidence from other capital measures, cost-benefit analysis will at least be a useful addition and more likely even the prime instrument of analytic choice. The former applies for instance to work integration interventions. The latter applies for example to human capital measured by improved educational degrees or in the protection of natural capital, both of which may contain an inherent economic value. The fact that costs always matter in the provision of organisational activities (input) does not mean that cost-benefit analysis always needs to be applied. The costs produced by the intervention can just as well be weighed against the capitals created (outcomes) without trying to monetise the latter—that is unless as just said monetary benefits represent an important analytic category of effects (compare to Marée & Mertens, 2012). There are cases where costs might not be of relevance at all, namely where two discrete interventions are to be compared, which do not differ fundamentally in the investment they require.

The criteria for when a combined assessment is useful need to emerge in the process of reasoning in the academic and public discourse I outlined in relation to what counts as capital and how capitals enable social productivity. However, there is a heuristic procedure which can be applied in prioritising between and potentially combining measures. A first question should be used to decide on whether a capitals based approach would be useful: does the intervention fundamentally affect relations between individuals, the positioning of individuals towards others, or the engagement of individuals within social settings? A second question should be applied to consider whether the capital based approach needs to be combined with other approaches: are any direct or indirect monetary benefits produced, or is the satisfaction and well-being of individuals directly affected, or does the enabling of individual choice and freedoms play a prominent role? Even if none of the other conceptual approaches is identified as particularly relevant through the second question, we would thirdly have to ask: will the enhancement of one capital through the interventions likely have any positive or negative interactions with other capitals or measures from the other approaches? Fourth, in order to prioritise between approaches and their respective measures impact analysists need to clarify: which of the identified effects is most fundamental for deciding whether the organisation has achieved its postulated mission?

In other words, we need to find out which measures will be central to decision makers, be they executives of the organisation, potential investors or regulators. By asking themselves these three questions impact analysts are going to be able to calibrate their conceptual grounding even in the absence of a definitive set of criteria which still need to emerge. My suggestion is that the capital based approach will be central in a large number of cases—any in which the first question is answered positively—and often will not need to be combined with the other approaches, yet the combination of approaches can be beneficial for a deeper understanding of the intervention.

9.3 Lessons learnt for measuring social impact

Now I want to briefly outline which lessons could be learnt through my research on the measurement of social impact more generally and how the insights generated differ from established practice.

First of all, it should have become clear that assessing social impact involves a considerable amount of effort, in particular now that so many tools and approaches are available and common standards are missing. This is in clear contrast to for instance online tools that are offered by which organisations can try to perform a quick and easy estimation of their impact remotely. While these tools might help organisations get insights into relevant perspectives of impact or develop an understanding for their 'theory of change' or 'impact chain' they can hardly ever serve as an assessment of actual impact. A detailed assessment of social impact along the capital based approach is also clearly different from analyses that judge organisational capacities and organisations' potential for creating impact. My work thus promotes the message that if social impact is to be assessed, it should be done thoroughly and no easy gains are to be expected. If organisations are not willing to take the effort involved, they should not proclaim their alternative analyses as 'impact measurement'. In effect against the experiences made, I would advocate clearer distinctions between other dimensions of organisational performance on the one side, such as 'service quality', 'operational efficiency', or 'accountability', and 'impact' on the other side.

Second, there is little social impact that cannot be systematically studied. My research has shown that even in cases where the main impact lies in areas that are typically considered hard or impossible to assess we can measure which changes have been achieved. The ability to do so rests on the grounding of the capital based approach in previous research and its utilisation of established measures. Beyond the observed messiness of the current land-

scape, this is a strong argument against the separate development of tools by individual stakeholders, in particular if these tools are guided by pragmatism only and lack grounding in existing research or conceptual literatures. The fact that so many tools have emerged is probably driven by the 'creators' ambition to 'brand' the new assessment method, or by the promotional role the tool is often supposed to fulfil for the organisation applying it (see Tuan, 2008, p. 22 for the latter). It might also have to do with the scepticism of practitioners about the usefulness of applying the detailed and lengthy processes academic research requires. My counter argument would be only by the later will we be able to steadily advance our ability to grasp what is currently thought of as unfathomable.

Third, an essential part of the capital based approach, if not an obvious one at face value, is a focus of the analysis on the organisation's target group. This does not equal a denial that other stakeholders are important. Instead it is a plea for recognising that as long as we have no unequivocal account what has improved for that very group and how the triggered effects have come about, it is rather inconsequential to explore effects on other stakeholders. Effects on the latter should thus only be traced subsequent to a thorough analysis of changes in the capitals of the target group. This focus is derived from the capital based approach's strong link to organisational mission. When we scrutinise the mission statements of social purpose organisations we will mostly be provided with a stress of a particular target group, often a needy one. And this is unsurprising, since social purpose organisations do not exist to save the state money or to generate income on markets. They exist to help those in need (broadly defined), and at least in part because both, the market and the state have been unable to address these needs or even produced them (see the arguments of state failure in Weisbrod, 1975 and market failure in Hansmann, 1987). Impact analysts should therefore calibrate their analysis along what and who is stressed in the organisation's mission.

A fourth learning is closely connected to the last one. Recognising that the target group needs to be at the centre of any impact assessment in combination with the prominent role of mission statements also provides clear arguments for differentiating by interventions when assessing social impact. Most organisations will not perform a single activity or have only one target group. However, when variation and multitude comes in it is unrealistic that impact can be covered in one sweep. Instead, when organisations contribute to several separate impact chains that contribute to social productivity and with different strategies or activities, each of them needs to be assessed individually. It might be tempting for organisations to aspire demonstrating all the positive effects they produce, but any integrated attempt will lack the rigour and thoroughness needed to critically assess whether, which and how impact is created by the different organisational activities. The multi-capitals framework underlines that connections and potentially trade-offs between different capitals exist even in relation to one discrete or closely related activities. In relation to the political web platform and AW's other activities for example, it is already challenging to clearly show and attribute achieved effects. It will be nearly impossible to maintain the necessary degree of differentiation if multiple activities, target groups or even fields of activity are analysed in combination. The complexity of analysing social impact thus asks for a careful delineation of the explicandum.

Fifth, one of the 'negative' or relativizing lessons learnt is that we might need to take a step back in the ambitions promoted in the social impact debate. The emergent state of classifications such as the capitals framework shows that we are far from a state of consolidation that would enable straightforward comparisons of discrete organisational interventions, which are needed to assess changes at a broader level. This ambition is often found and ultimately grounded in the broader definition of social impact as 'change within society'. It is directly connected to the current focus on a multitude of stakeholders or the desire to

assess the impact of entire organisations rather than interventions. My analysis has demonstrated that even focussed, multi-facetted and deep investigations of impact come with uncertainties, ambiguities and open questions. It is therefore illusionary to think that the benefits of applying a broad society-oriented understanding of impact at present would be capable of providing an additional added-value big enough to counterbalance the biases and confounding influences provoked by such a move. This confirms the usefulness of the 'attributable outcome' version in studying organisational social impact and demands a measured approach to the ambitions and claims of what social impact measurement can do at present.

Now that we've seen what kind of learnings my research has enabled, it remains to be discussed where its limitations lie.

9.4 Limitations

The capital based approach has only in a limited sense been discussed in relation to detecting negative social impacts (Nicholls & Murdock, 2012b), be they intended or unintended. It is obviously capable of assessing decreases in the capital forms assessed, if a pre-observation and a post-observation are applied. Although not encountered in my own empirical studies, such decreases may be indirect, resulting from interaction effects between the capital forms at play, or direct, that is produced by a misled intervention strategy. For instance, furthering social capital within a multi-stakeholder network will not be productive when an ideological clash occurs in the cultural capital of the involved parties (interaction). Or an intervention might inadvertently lower political participation by making too much or inaccurate information available in a political discourse (defective intervention). The capital based approach would need to be tested in relevant empirical settings to deepen insights about possible interactions or defective interventions. It would also require more

reasoning and testing to find out whether the approach is generally useful in studying negative effects of interventions apart from these particular constellations.

One of the major downsides of my approach is the amount of effort, expertise and resources it requires, and thus the transaction costs it produces. For assessing capital forms, we need to first perform a careful drafting of the respective capital(s) under investigation. As mentioned there are links in previous research that can be used, but the amount of work that needs to be invested into adapting them for the measurement of social impact and for providing a good match to the specifics of the organisation is considerable. It is also not always obvious where to look and how measures can be adapted. In any case they would need to be tested for validity and reliability, which requires specific expertise. As has also been mentioned, rigorous designs need to be applied.

The conceptual challenges can be mastered through involving the right kinds of experts. The need for realising rigorous designs in contrast will produce transaction costs of the literal kind, that is require considerable budgets, which are harder to keep in check. As a rule of thumb, the broader the issues and the more challenging the target group the more costly will the performed research be. Online surveys can be performed at little costs. Interviews for gathering qualitative data will be considerably more expensive. Yet, since typically a lower number of participants is needed for such research, they will probably be at a medium level when it comes to costs. The most expensive form of a study is a quantitative survey that needs to be administered in person, that is where one-on-one questioning is needed (as has been the case for the study on community oriented housing). In the latter case a single study can result in a medium range five-digit budget. This would be a huge obstacle for many social purpose organisations. However, such challenges are not overly specific to the capital based approach. Sophisticated designs will always be costly, regard-

less of the conceptual approach taken. Costs and how to deal with them still need to be considered when thinking about the practical and policy use of applying the capital based approach empirically.

9.5 Implications for management theory, policy and practice

My approach has major implications for the practice of social impact measurement. First of all, it offers an alternative measurement that can help organisations, investors and policy makers alike to make an informed decision about interventions by focusing on the main sources of their social impact: the capital(s) they produce. Thereby it not only ensures that the performed analysis is 'material' to the served target group, but also helps calibrating the analysis towards the organisation's core mission. Studying social impact per intervention and by analysing key capital components promotes coherence and lucidity in understanding and communicating social impact mechanisms, as increasingly demanded for and by non-profits, companies and impact investors alike (Epstein & Yuthas, 2014).

The practice and research of non-profit and public sector performance management is currently often limited to classifying performance frameworks or discussing standard measures of performance and issues related to it, instead of focussing on how to measure social impact (Moynihan, 2008). An example of the current focus is to be found in research looking at standard performance measures such as the number of beneficiaries for example; another example is research that is interested in the legitimacy arising from providing evidence (Allan, 2016; Arena et al., 2015; Chatterji & Levine, 2006; Cho & Trent, 2009; Ebrahim & Rangan, 2014; Hatzfeld, 2014; Nicholls, 2009b). The capital based approach would help increase the conceptual scope and depth of measures used. At the same time, it could help condensing the effort which would otherwise be invested into tracing a magnitude of follow-up effects. It can be combined with the alternative conceptual approaches

discussed in my work or other performance tools (Berman, 2006; Poister et al., 2015 for an overview). Complementary approaches and tools need to be included for example when assessing additional performance criteria such as cost effectiveness, efficiency etc.

The capital based approach ideologically is close to a 'dash-board' format for reporting the social impact achieved. When the levels of or shifts in capital components are displayed in a radar chart for instance, the latter helps maintain informational richness by enabling benchmarking through a systemised way of illustrating impacts, which enables the establishment of links to emerging reporting standards (Nicholls, 2009a). For the individual measures to effectively perform the function of benchmarking common standards would need to emerge based on further empirical studies. The capitals framework presented in the methods section suggests that consolidation and standard setting would then be likely.

As mentioned in the limitations, all the positive aspects shall not hide the fact that establishing a link to research in a variety of fields, which is necessary for a sound construction of capital measures, is a demanding task. The good news for academia is that it can play a key role in this and enrich a social impact measurement landscape which is currently dominated by commercial providers (see Krlev et al., 2013 on SROI). If policy makers are interested in evidence of impact, which they probably should be, given new financing vehicles such as social impact bonds, they would need to try to promote the benchmarking process. For to move from the 'attributable outcome' to the 'societal change' definition of social impact we would first require a consolidation of knowledge and second a combination of organisational level studies with large scale research that traces follow-up and long-term effects, in particular when prevention is involved. Policy makers or also investors could get engaged in financing studies on interventions that are advocated as new and particularly effective, especially those that could transform the way services are provided at

the field level. More individual and less distinct approaches promoted by organisations would need to be financed by the organisations themselves. The mix of designs applied to realise the capital based approach shows that there is sufficient flexibility in making the approach practicable for smaller organisations too—within the restraints of what is needed to produce reliable evidence of capital formation.

9.6 Prompts for future research

My research has implications for shaping social science theory beyond the immediate realm of measuring organisational impact. I suggest its primary links are (1) the debate about capitals and (2) the study of (social) innovation.

The approach enriches the study of productive systems by extending the classical economic view on productivity towards society as a productive system. It establishes a link between the study of social welfare and that of how economic and other capital forms affect society. This needs to be explored further, including the discussion within the wider social sciences of whether social capital, human capital and others should actually be regarded as capital at all, and if so which benefit this brings about (Baron, Field, & Schuller, 2000; Milanovic, 2015). It may even be extended to the controversies concerning economic capital itself, for example the question in which units capital should be measured (Birner, 2002; Cohen & Harcourt, 2003). My reasoning in relation to social, cultural and political capital in the context of social productivity may provide some arguments in favour of embracing the ‘capital concept’. It also outlines how capitals, albeit other than economic in nature, can be measured.

At any rate, applying the different capitals under the specific lens of productivity could help determining their boundaries and address the increasing criticism that their meaning

has been extended so far as to mean all and nothing (see Fine, 2010 on social capital). The application of capitals in social impact measurement brings discipline into drafting and operationalising capital forms, not least since their formation is linked to the organisation's performance in achieving its mission. As of today we have certainly made a step beyond that "[t]here is far more speculation than measurement" (Paldam, 2000 on social capital). However, measurement to date mostly refers to secondary survey data (Bjørnskov, 2006b; Elgar et al., 2011; Guillen, Coromina, & Saris, 2011; Rothstein & Stolle, 2008), which forces scholars to compile measures that match available data rather than the other way around. Starting to gather data only once concepts are defined is even more important for political and cultural capital than for social capital, since their study is far less advanced.

Alongside the links to the capitals debate, my approach could also be developed further by an exploration of how it could serve to study social innovation or organisations' social innovativeness as the ability to address social needs and potentially spur transformative social change (Anheier et al., 2014, p. 33). For this it might be useful to refer back to the discussion of Salter's work on economic productivity and the proposed concept of social productitvity. Salter (1969, p. 5) describes innovation as a push in productivity through a major advancement in state-of-the-art techniques that determine the outfit of capitals employed in the production process. The advancement of techniques through innovation can cause major leaps forward, but as Salter remarks, the cumulative effects of smaller improvements often match the effects of the greater shifts. Hence the questions arises whether the concept of social productivity proposed, on which the capitals approach rests, is compatible with the implications given by the study of technological change and economic productivity.

In Salter's study, the enhancement of techniques designates technological progress in machinery and equipment across a large variety of industries, including the production of bricks, cement, houses, shoes, beet sugar, cotton or jute. These shifts (with delays due to time lags in the diffusion of innovation) have levered economic productivity significantly. Let's consider a prominent social innovation, namely the foundation of the modern university promoted by Humboldt (Drucker, 1985, p. 21). We will see that the way social innovations lever social productivity is similar to the mechanisms found in technological innovation. In both cases, the one of economic and social productivity, capital and techniques are closely interrelated.

When considering the example of the modern university, it could be argued that the innovation that led to a push in social productivity was grounded in increasing the level of human capital fostered by the unity of research and teaching. It could also be argued that it lay in the strengthening of the influence of universities on public policy or the initiation of closer, cross-border collaboration, in geographic terms as well as regarding academic disciplines; the former resulting from the constituted independence of science, the latter driven by the effort to integrate formerly separate disciplinary traditions. In all of these cases, the enhancement of social productivity through innovation can be assessed by taking account of the formation of the different capital forms, more explicitly by (1) the levering of skills and expertise, (2) the strengthening of political participation, (3) the initiation of social connections and (4) the reduction of pillarised thinking. If this thought experiment works with a very fundamental and broad incident, such as the one just used, it is to be suspected that it will work with more fine-grained and more narrowly defined innovation activities at the organisational level too. However, this issue would require more scrutiny and would need to be tested along a range of different innovations.

9.7 Outlook

Although the research has focussed on social purpose organisations, because the question about social impact is closest to their reason for being, it is to be highlighted that the capital based approach is in principle applicable to all kinds of organisations. I have presented in the introduction that there are numerous factors that urge firms to care about their social impact. The latter has recently been promoted as a core component of "big strategy" in firms (Whittington, 2012).

Firms, supposedly more than other organisations, follow the imperative of "what cannot be measured, cannot be managed" or some variation thereof—a statement whose origin is related back to the work of various scholars ranging from Peter Drucker (Dumay & Rooney, 2011) to The Lord Kelvin (van der Zee, 2002). It expresses a desire to know if intended goals are met and if not, what can be done about it. Against this background it appears strange that firms are increasingly performing deeds (CSR, philanthropic engagement, corporate volunteering etc.) of which they might not know or care which effects they have on those affected (Rangan, Chase, & Karim, 2015), or that they seem preoccupied with narrow versions of performance (efficiency) while neglecting the bigger picture (impact) (Whittington, 2012).

The applicability of the capital based approach to for-profits is not limited to actions intended to promote the 'social good' such as the ones just mentioned. It can also be applied in the context of what is sometimes referred to as social accounting or reporting (Gray, 2002; Gray, Kouhy, & Lavers, 1995), meaning a concerted effort of assessing and reporting the effects that firms have beyond the financial figures they produce. One could for instance assess the effects of establishing a multinational corporation's factory in a deprived area in a developing country on the role of women in the rural community (cultural

capital). One could also investigate whether such an action affects the ability of the population to make their voices heard to the regional government (political capital). The capital based approach could also help investigate the reverse situation, namely what happens when the firm decides to move their production plant to another part of the country or abroad? These situations link up to that the approach can be used to study productive as well as destructive effects of organisational activity, be they intended or unintended. An example of existing research related to this kind of thinking is Banarjee, Jackson and colleagues' work (2013; 2012) on the effects of a microfinance programme on social networks, as one part of social capital, in rural Indian communities. Microfinance is more closely connected to the social purpose area, but the study I believe outlines the general applicability of the capital based approach to a variety of interventions and effect investigations, including the consequences of for-profits' actions.

Just as to external effects, my approach could in principle be used to study social impact within for-profit organisations. An example would be the effects of flexible work arrangements on the social capital of employees. Flexible work patterns or their reverse, namely work place control, have been at the centre of research. However, existing research mostly seems to focus on their effects on job satisfaction or career choices (see for instance Irak & Mantler, 2017; Tomlinson, Baird, Berg, & Cooper, 2017 for flexible arrangements and Briscoe, 2006 for rigid ones) rather than employees' social contacts.

Thus, in addition to the value the capital based approach is likely to unfold for social purpose organisations in future research, it seems a promising option for thinking more explicitly about how to assess the effects of organisations on their surrounding human and natural environment more generally, the negative as well as the positive ones.

10 Bibliography

Abbott, A. (2015). *The Future of the Social Sciences*. 37th Marc Bloch Lecture. École des Hautes Études en Sciences Sociales. Paris, June 18. Retrieved from http://home.uchicago.edu/aabbott/Papers/Marc%20Bloch%20Lecture%20Pre%20Trans.pdf

Abdelal, R., Herrera, Y. M., Johnston, A. I., & McDermott, R. (2006). Identity as a Variable. *Perspectives on Politics*, *4*(4), 695–711.

Abgeordnetenwatch. (2017). *Das Ziel: Was wir wollen*. Retrieved from https://www.abgeordnetenwatch.de/was-wir-wollen-das-ziel

Adler, P. S., & Kwon, S.-W. (2002). Social Capital: Prospects for a New Concept. *The Academy of Management Review*, *27*(1), 17–40. doi:10.2307/4134367

Affonso, D. D., Mayberry, L., Shibuya, J. Y., Archambeau, O. G., Correa, M., Deliramich, A. N., & Frueh, B. C. (2010). Cultural Context of School Communities in Rural Hawaii to Inform Youth Violence Prevention. *Journal of School Health*, *80*(3), 146–152. doi:10.1111/j.1746-1561.2009.00478.x

Albrecht, S. (2006). Whose voice is heard in online deliberation?: A study of participation and representation in political debates on the internet. *Information, Communication & Society*, *9*(1), 62–82. doi:10.1080/13691180500519548

Alexander, J. C. (2006). *The civil sphere*. New York: Oxford University Press.

Alkire, S. (2005). Why the Capability Approach? *Journal of Human Development*, *6*(1), 115–135. doi:10.1080/146498805200034275

Alkire, S., & Foster, J. (2011a). Counting and Multidimensional Poverty Measurement. *Journal of Public Economics*, *95*((7-8)), 476–487. Retrieved from doi:10.1016/j.jpubeco.2010.11.006

Alkire, S., & Foster, J. (2011b). *Understandings and misunderstandings of multidimensional poverty measurement*. *Working paper / Oxford Poverty & Human Development Initiative: Vol. 43*. Oxford: Oxford Poverty & Human Development Initiative.

Allan, W. L. (2016). Factors that impact how civil society intermediaries perceive evidence. *Evidence & Policy: A Journal of Research, Debate and Practice*. doi:10.1332/174426416X14538259555968

Allison, K. W., Edmonds, T., Wilson, K., Pope, M., & Farrell, A. D. (2011). Connecting Youth Violence Prevention, Positive Youth Development, and Community Mobilization. *American Journal of Community Psychology*, *48*(1-2), 8–20. doi:10.1007/s10464-010-9407-9

Almond, G. A., & Verba, S. (1989). *The civic culture revisited: Political attitudes and democracy in five nations*. Newbury Park, Calif: Sage Publications.

Ames, D. (1990). Depression among elderly residents of local-authority residential homes. Its nature and the efficacy of intervention. *The British Journal of Psychiatry*, *156*(5), 667–675. doi:10.1192/bjp.156.5.667

Andersen, J. G., van Deth, J. W., Geurts, P., Viegas, J. M. L., Badescu, G., Selle, P., … (2007). *Citizenship, Involvement, Democracy*. GESIS Data Archive, Cologne. ZA4492 Data file Version 1.0.0. Retrieved from https://dbk.gesis.org/dbksearch/sdesc2.asp?no=4492

Andriof, J., Waddock, S., Husted, B., & Sutherland Rahman, S. (Eds.). (2003). *Unfolding stakeholder thinking 2: Relationships, Communication, Reporting and Performance*. Sheffield: Greenleaf.

Anheier, H. K. (2014). *Nonprofit organizations: Theory, management, policy* (2nd). London: Routledge.

Anheier, H. K., Krlev, G., Preuss, S., Mildenberger, G., Bekkers, R., Mensink, W., … (2014). *Social innovation as Impact of the Third Sector*. A deliverable of the project: "Impact of the Third Sector as Social Innovation" (ITSSOIN), European Commission - 7th Framework Pro-gramme, Brussels: European Commission, DG Research.

Arena, M., Azzone, G., & Bengo, I. (2015). Performance Measurement for Social Enterprises. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, *26*(2), 649–672. doi:10.1007/s11266-013-9436-8

Argyris, C., & Schön, D. A. (1978). *A theory of action perspective*. *Organizational learning*. Reading, Mass: Addison-Wesley Pub. Co.

Arrow, K. J. (1970). *Social choice and individual values* (2d ed). *Yale University. Cowles Foundation for Research in Economics. Monograph: Vol. 12*. New Haven: Yale University Press.

Arvidson, M., Battye, F., & Salisbury, D. (2014). The social return on investment in community befriending. *International Journal of Public Sector Management*, *27*(3), 225–240. doi:10.1108/IJPSM-03-2013-0045

Arvidson, M., & Lyon, F. (2013). Social Impact Measurement and Non-profit Organisations: Compliance, Resistance, and Promotion. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, *25*(4), 869–886. doi:10.1007/s11266-013-9373-6

Ashoka Germany. (2010). *Ashoka Fellows: Heather Cameron*. Retrieved from http://germany.ashoka.org/fellow/heather-cameron

Austin, P. C. (2011a). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, *46*(3), 399–424. doi:10.1080/00273171.2011.568786

Austin, P. C. (2011b). Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharmaceutical statistics*, *10*(2), 150–161. doi:10.1002/pst.433

Babüroglu, O. N., & Ravn, I. (1992). Normative Action Research. *Organization Studies*, *13*(1), 19–34. doi:10.1177/017084069201300104

Bagnoli, L., & Megali, C. (2009). Measuring Performance in Social Enterprises. *Nonprofit and Voluntary Sector Quarterly*, *40*(1), 149–165. doi:10.1177/0899764009351111

Baier, D., Pfeiffer, C., Simson, J., & Rabold, S. (2009). *Jugendliche in Deutschland als Opfer und Täter von Gewalt: Erster Forschungsbericht zum gemeinsamen Forschungsprojekt des Bundesministeriums des Innern und des KFN*. Retrieved from kfn.de/wp-content/uploads/Forschungsberichte/FB_107.pdf

Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2013). The diffusion of microfinance. *Science (New York, N.Y.)*, *341*(6144), 1236498. doi:10.1126/science.1236498

Bang, H. P., & Sørensen, E. (1999). The Everyday Maker: A New Challenge to Democratic Governance. *Administrative Theory & Praxis*, *21*(3), 325–341. Retrieved from http://www.jstor.org/stable/25611362

Bang, H. P., & Sørensen, E. (2003). The Everyday Maker: Building political rather than social capital. In P. Dekker & E. M. Uslaner (Eds.), *Routledge/ECPR studies in Euro-*

pean political science: Vol. 23. Social capital and participation in everyday life (pp. 148–161). London: Routledge.

Barman, E. (2016). *Caring capitalism: The meaning and measure of social value*. New York: Cambridge University Press.

Barnes, J., Sampson, R. J., Kindlon, D., & Reiss, A. J. (1997). Chapter 5. A Community Survey Approach to Ecological Assessment. Results From a Pilot Study. In National Institute of Justice (Ed.), *Project on Human Development in Chicago Neighborhoods. Technical report I* (pp. 34–45).

Barnett, E., & Casper, M. (2001). A definition of "social environment". *American Journal of Public Health*, *91*(3), 465.

Baron, S., Field, J., & Schuller, T. (Eds.). (2000). *Social capital: Critical perspectives*. Oxford, New York: Oxford University Press.

Bassett, R., Beagan, B. L., Ristovski-Slijepcevic, S., & Chapman, G. E. (2008). Tough Teens: The Methodological Challenges of Interviewing Teenagers as Research Participants. *Journal of Adolescent Research*, *23*(2), 119–131. doi:10.1177/0743558407310733

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis with special reference to education*: University of Chicago Press.

Becker, G. S. (1974). A Theory of Social Interactions. *Journal of Political Economy*, *82*(6), 1063–1093. doi:10.2307/1830662

Beresford, P., & Carr, S. (Eds.). (2012). *Research highlights in social work series: Vol. 55. Social care, service users and user involvement*. London, Philadelphia: Jessica Kingsley Publishers.

Berger-Schmitt, R. & Noll, H.-H. (2000). *Conceptual Framework and Structure of a European System of Social Indicators: Towards a European System of Social Reporting and Welfare Measurement, A TSER-Project financed by the European Commission*. EuReporting Working Paper No. 9. Retrieved from http://www.gesis.org/fileadmin/upload/dienstleistung/daten/soz_indikatoren/eusi/paper9.pdf

Bergmann, J. (2014). Der sanfte Kämpfer. *Brand Eins*, (11), 132–136. Retrieved from https://www.brandeins.de/archiv/2014/scheitern/joachim-koerkel-der-sanfte-kaempfer-kontrolliertes-trinken-abstinenz-dogma/

Berk, R. A. (1983). An Introduction to Sample Selection Bias in Sociological Data. *American Sociological Review*, *48*(3), 386–398. doi:10.2307/2095230

Berman, E. M. (2006). *Performance and productivity in public and nonprofit organizations* (2nd ed.). Armonk, N.Y.: M.E. Sharpe.

Biesta, G. (2007). Why "what works" won't work: evidence-based practice and the democratic deficit in educational research. *Educational Theory*, *57*(1), 1–22. doi:10.1111/j.1741-5446.2006.00241.x

Birner, J. (2002). *The Cambridge controversies in capital theory: A study in the logic of theory development. Routledge studies in the history of economics: Vol. 47*. London, New York: Routledge.

Bjørnskov, C. (2006a). Determinants of generalized trust: A cross-country comparison. *Public Choice*, *130*(1-2), 1–21. doi:10.1007/s11127-006-9069-1

Bjørnskov, C. (2006b). The multiple facets of social capital. *European Journal of Political Economy*, *22*(1), 22–40. doi:10.1016/j.ejpoleco.2005.05.006

Blattman, C., Emeriau, M., & Fiala, N. (2017). *Do Anti-Poverty Programs Sway Voters? Experimental Evidence from Uganda*. National Bureau of Economic Research. Cambridge, MA.

Bliss, C. J. (1993). Life-Style and the Standard of Living. In M. Nussbaum & A. Sen (Eds.), *The Quality of Life* (pp. 417–436). Oxford University Press.

Bohman, J. (1999). Critical theory and democracy. In D. M. Rasmussen (Ed.), *The handbook of critical theory* (1st ed., pp. 190–219). Oxford: Blackwell.

Bonini, S. & Emerson, J. (2005). *Maximizing Blended Value: Building Beyond the Blended Value Map to Sustainable Investing, Philanthropy and Organizations*. Retrieved from http://www.blendedvalue.org/media/pdf-max-blendedvalue.pdf

Bornstein, D. (2007). *How to change the world: social entrepreneurs and the power of new ideas* (Updated ed.). Oxford: Oxford Univ. Press.

Bosworth, K., Espelage, D., Du Bay, T., Dahlberg, L. L., & Daytner, G. (1996). Using multimedia to teach conflict-resolution skills to young adolescents. *American Journal of Preventive Medicine*, *12*(5 (Supplement)), 65–74.

Bourdieu, P. (1973). Cultural Reproduction and Social Reproduction. In B. Richard (Ed.), *Knowledge, Education, and Cutlural Change* (pp. 71–112). London: Tavistock.

Bourdieu, P. (1977). *Outline of a theory of practice*. Cambridge: Cambridge University Press.

Bourdieu, P. (1986). The Forms of Capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). New York: Greenwood Press.

Bowi, U., Ott, G., & Tress, W. (2008). Faustlos – Gewaltprävention in der Grundschule. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, *57*(7), 509–520.

BRAC. (2017). *Our approach*. Retrieved from http://www.brac.net/our-approach

Bradley, E. H., Curry, L. A., & Devers, K. J. (2007). Qualitative data analysis for health services research: developing taxonomy, themes, and theory. *Health services research*, *42*, 1758–1772. Retrieved from 10.1111/j.1475-6773.2006.00684.x

Brady, H. S., Verba, S., & Schlozman, K. L. (1995). Beyond Ses: A Resource Model of Political Participation. *The American Political Science Review*, *89*(2), 271–294.

Brest, P., Harvey, H., & Low, K. (2009). Calculated Impact. *Stanford Social Innovation Review*, (Winter), 50–56. Retrieved from http://www.ssireview.org/articles/entry/calculated_impact

Briscoe, F. (2006). Temporal Flexibility and Careers: The Role of Large-Scale Organizations for Physicians. *ILR Review*, *60*(1), 88–104. doi:10.1177/001979390606000105

Brown, T. A. (2006). *Confirmatory factor analysis for applied research. Methodology in the social sciences*. New York: Guilford Press.

Bunge, M. A. (1996). *Finding philosophy in social science*. New Haven, Conn.: Yale University Press.

Burchell, S., Clubb, C., Hopwood, A. G., Hughes, J., & Nahapiet, J. (1980). The role of accounting in organizations and society. *Accounting, Organizations and Society*, *5*(1), 5–27.

Burt, R. S. (2000). The Network Structure of Social Capital. In R. Sutton & B. Staw (Eds.), *Research in Organizational Behavior* (pp. 345–423). Greenwich, CT: JAI Press.

Bushe, G. R. (2013). Generative process, generative outcome: The transformational potential of appreciative inquiry. In D. L. Cooperrider, D. P. Zandee, L. N. Godwin, M. Avital, & B. Boland (Eds.), *Advances in appreciative inquiry: volume 4. Organizational*

generativity. The appreciative inquiry summit and a scholarship of transformation (pp. 89–113). Bingley, United Kingdom: Emerald.

Byford, S., & Sefton, T. (2003). Economic Evaluation of Complex Health and Social Care Interventions. *National Institute Economic Review*, *186*, 98–108.

CamP Group. (2015). *RespAct Baseline*. Retrieved from http://www.respact.org/home-page-2/the-project/1974-2/

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*: SAGE Publications Limited.

Chatterji, A., & Levine, D. (2006). Breaking Down the Wall of Codes: Evaluating Non-Financial Performance Measurement. *California Management Review*, *48*(2), 29–51.

Cho, J., & Trent, A. (2006). Validity in qualitative research revisited. *Qualitative Research*, *6*(3), 319–340.

Cho, J., & Trent, A. (2009). Validity Criteria for Performance-Related Qualitative Work: Toward a Reflexive, Evaluative, and Co-constructive Framework for Performance in/as Qualitative Inquiry. *Qualitative Inquiry*, *15*(6), 1013–1041. doi:10.1177/1077800409334224

Chowdhury, A. M. R., & Bhuiya, A. (2004). The wider impacts of BRAC poverty alleviation programme in Bangladesh. *Journal of International Development*, *16*(3), 369–386. doi:10.1002/jid.1083

Clark, C., Rosenzweig, W., Long, D., & Olsen, S. (2004). *Double bottom line project: Assessing social impact in double bottom line ventures*. Working Paper Series, Center for Responsible Business, UC Berkeley.

Clarkson, M. B. E. (1995). A Stakeholder Framework for Analyzing and Evaluating Corporate Social Performance. *Academy of Management Review*, *20*(1), 92–117.

Clemens, M. A., & Demombynes, G. (2011). When does rigorous impact evaluation make a difference? The case of the Millennium Villages. *Journal of Development Effectiveness*, *3*(3), 305–339. doi:10.1080/19439342.2011.587017

Clifford, J. (2014). *Proposed approaches to social impact measurement in European Commission legislation and in practice relating to EuSEFs and the EaSI: GECES sub-group on impact measurement 2014*. Luxembourg: Publications Office.

Coase, R. (1937). The Nature of the Firm. *Economica*, *4*(16), 386–405. doi:10.1111/j.1468-0335.1937.tb00002.x

Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics*, *3*(October), 1–44.

Coase, R. H. (1992). The New Institutional Economics. *The American economic review*, *88*(2), 72–74.

Cohen, A. J., & Harcourt, G. C. (2003). Retrospectives: Whatever Happened to the Cambridge Capital Theory Controversies? *The journal of economic perspectives*, *17*(1), 199–214. doi:10.2307/3216846

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. doi:10.1037/0033-2909.112.1.155

Cohen, S., & Hoberman, H. M. (1983). Positive Events and Social Supports as Buffers of Life Change Stress1. *Journal of Applied Social Psychology*, *13*(2), 99–125. doi:10.1111/j.1559-1816.1983.tb02325.x

Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. *The American journal of sociology*, *94*(Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure), 95–120.

Commission on 2020 Public Services. (2010). *From social security to social productivity: A vision for 2020 public services : the final report of the Commission on 2020 Public Services*. London: 2020 Public Services Trust. Retrieved from http://clients.squareeye.net/uploads/2020/documents/PST_final_rep.pdf

Cooperrider, D. L., & Srivastva, S. (1987). Appreciative inquiry in organizational life. *Research in Organizational Change and Development*, *1*, 129–169.

Costanza, R., d'Arge, R., Groot, R. de, Farber, S., Grasso, M., Hannon, B., … (1997). The value of the world's ecosystem services and natural capital. *Nature*, *387*(6630), 253–260. doi:10.1038/387253a0

Crepaldi, C., Rosa, E. de, & Pesce, F. (2012). *Workpackage 1: Literature Review on innovation in social services in Europe (sectors of Health, Education and Welfare Services) Report*. IRS. A deliverable of the project INNOSERV, European Commission - 7th Framework Programme, Brussels: European Commission, DG Research.

Crick, N. R., & Bigbee, M. A. (1998). Relational and overt forms of peer victimization: A multiinformant approach. *Journal of Consulting and Clinical Psychology*, *66*(2), 337–347. doi:10.1037/0022-006X.66.2.337

Crick, N. R., & Grotpeter, J. K. (1995). Relational Aggression, Gender, and Social-Psychological Adjustment. *Child Development*, *66*(3), 710–722. doi:10.2307/1131945

Cronbach, L. J. (1990). *Essentials of psychological testing. 5th ed*. New York: HarperCollins.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. doi:10.1037/h0040957

Cronin, M. J., & Dearing, T. C. (2017). *Managing for social impact: Innovations in responsible enterprise*. Cham, Switzerland: Springer.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*(4), 349–354. doi:10.1037/h0047358

Cunningham, C. E., Cunningham, L. J., Ratcliffe, J., & Vaillancourt, T. (2010). A Qualitative Analysis of the Bullying Prevention and Intervention Recommendations of Students in Grades 5 to 8. *Journal of School Violence*, *9*(4), 321–338. doi:10.1080/15388220.2010.507146

Cyranowski, J. M., Zill, N., Bode, R., Butt, Z., Kelly, Morgen A R, Pilkonis, P. A., … (2013). Assessing social support, companionship, and distress: National Institute of Health (NIH) Toolbox Adult Social Relationship Scales. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, *32*(3), 293–301. doi:10.1037/a0028586

Dacin, M. T., Goodstein, J., & Scott, W. R. (2002). Institutional Theory and Institutional Change: Introduction to the Special Research Forum. *The Academy of Management Journal*, *45*(1), 43. doi:10.2307/3069284

Dahlberg, L. L., & Butchart, A. (2005). State of the science: violence prevention efforts in developing and developed countries. *International Journal of Injury Control and Safety Promotion*, *12*(2), 93–104. doi:10.1080/15660970500086239

Dahlberg, L. L., Toal, S. B., Swahn, M., & Behrens, C. B. (2005). *Measuring Violence-Related Attitudes, Behaviors, and Influences Among Youths: A Compendium of Assessment Tools*. Atlanta, Georgia.

Dalton, R. J. (2008). Citizenship Norms and the Expansion of Political Participation. *Political Studies*, *56*(1), 76–98.

Dart, R. (2004). Being “Business-Like” in a Nonprofit Organization: A Grounded and Inductive Typology. *Nonprofit and Voluntary Sector Quarterly*, *33*(2), 290–310. doi:10.1177/0899764004263522

Davis, J. B., & Marin, S. R. (2009). Identity and Democracy: Linking individual and social reasoning. *Development*, *52*, 500–508.

Deaton, A., & Cartwright, N. (2016). *Understanding and Misunderstanding Randomized Controlled Trials*. Cambridge, MA: National Bureau of Economic Research.

Dees, J. G. (2001). *The meaning of Social Entrepreneurship*. Occasional Paper. Durham, NC.

Dess, G. G., & Davis, P. S. (1984). Porter's (1980) Generic Strategies as Determinants of Strategic Group Membership and Organizational Performance. *Academy of Management Journal*, *27*(3), 467–488. doi:10.2307/256040

Dess, G. G., & Shaw, J. D. (2001). Voluntary turnover, social capital, and organizational performance. *Academy of Management Review*, *26*(3), 446–456. doi:10.5465/AMR.2001.4845830

DeVellis, R. F. (2016). *Scale development: Theory and applications* (Fourth edition). *Applied social research methods series: Volume 26, 4th edition*. Thousand Oaks: SAGE Publications, Inc.

Di Gennaro, C. (2006). The Internet and the Public: Online and Offline Political Participation in the United Kingdom. *Parliamentary Affairs*, *59*(2), 299–313. doi:10.1093/pa/gsl004

Diefenbacher, H., & Zieschank, R. (2010). *Wohlfahrtsmessung in Deutschland: Ein Vorschlag für einen nationalen Wohlfahrtsindex*: Forschungsstätte der Ev. Studiengemeinschaft.

Diener, E., Helliwell, J. F., & Kahneman, D. (Eds.). (2010). *Series in positive psychology. International differences in well-being*. Oxford, New York: Oxford University Press.

Diener, E., Inglehart, R., & Tay, L. (2013). Theory and Validity of Life Satisfaction Scales. *Social Indicators Research*, *112*(3), 497–527. doi:10.1007/s11205-012-0076-y

Diener, E., Lucas, R. E., & Scollon, C. N. (2006). Beyond the hedonic treadmill: Revising the adaptation theory of well-being. *American Psychologist*, *61*(4), 305–314. doi:10.1037/0003-066X.61.4.305

Diener, E., Shigehiro, O., & Lucs, R. E. (2009). Subjective Well-Being: The Science of Happiness and Life Satisfaction. In C. R. Snyder (Ed.), *Handbook of positive psychology* (pp. 187–194). Oxford: Oxford Univ. Press.

DiMaggio, P. (1982). Cultural Capital and School Success: The Impact of Status Culture Participation on the Grades of U.S. High School Students. *American Sociological Review*, *47*(2), 189–201. doi:10.2307/2094962

Dolan, P. (2011). *Using Happiness to Value Health*. London.

Dolan, P., Fujiwara, D., & Metcalfe, R. (2011). *A step towards valuing utility the marginal and cardinal way*. CEP Discussion Paper No 1062.

Donaldson, T., & Preston, L. E. (1995). The Stakeholder Theory of the Corporation: Concepts, Evidence, and Implications. *The Academy of Management Review*, *20*(1), 65–91. doi:10.2307/258887

Drèze, J., & Stern, N. (1985). The theory of cost-benefit analysis. In A. J. Auerbach & M. S. Feldstein (Eds.), *Handbook of public economics* (pp. 909–990). Amsterdam: North-Holland Pub. Co.

Drucker, P. F. (1985). *Innovation and entrepreneurship: Practice and principles*. New York: HarperBusiness.

Dumay, J., & Rooney, J. (2011). "Measuring for managing?: " An IC practice case study. *Journal of Intellectual Capital*, *12*(3), 344–355. doi:10.1108/14691931111154670

Ebrahim, A., & Rangan, V. K. (2010). Putting the brakes on impact: A contingency framework for measuring social performance. *Academy of Management Proceedings*, *2010*(1), 1–6. doi:10.5465/AMBPP.2010.54500944

Ebrahim, A., & Rangan, V. K. (2014). What Impact? A Framework for Measuring the Scale and Scope of Social Performance. *California Management Review*, *56*(3), 118–141.

Edelstein, I. (2016). *Measuring Pathways to Youth Violence and the Possible Effects of a Sport-based Development Intervention on Youth in South Africa* (Doctoral dissertation). University of Cape Town, Cape Town.

Edery, Y. (2006). A new model for supporting social enterprise through sustainable investment. *Social Enterprise Journal*, *2*(1), 82–100. doi:10.1108/17508610680000715

Edmiston, D., & Nicholls, A. (2017). Social Impact Bonds: The Role of Private Capital in Outcome-Based Commissioning. *Journal of Social Policy*, *29*, 1–20. doi:10.1017/S0047279417000125

Eisner, M., & Malti, T. (2012). The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section. *International Journal of Conflict and Violence*, *6*(2), 167–175.

Elgar, F. J., Davis, C. G., Wohl, M. J., Trites, S. J., Zelenski, J. M., & Martin, M. S. (2011). Social capital, health and life satisfaction in 50 countries. *Health & Place*, *17*(5), 1044–1053. doi:10.1016/j.healthplace.2011.06.010

Elkington, J. (1998). Partnerships from Cannibals with Forks: The Triple Bottom Line of 21st-Century Business. *Environmental Quality Management*, *8*(1), 37–51.

Elkington, J. (2010, April 6). A new paradigm for change. *Mc Kinsey Quarterly - What matters*. Retrieved from http://whatmatters.mckinseydigital.com/social_entrepreneurs/a-new-paradigm-for-change

Elkington, J., & Hartigan, P. (2008). *The power of unreasonable people: How social entrepreneurs create markets that change the world*. Boston, Mass: Harvard Business Press.

Emerson, J. (2003). The Blended Value Proposition: Integrating Social and Financial Returns. *California Management Review*, *45*(4), 35–51.

Epstein, M. J., & Yuthas, K. (2014). *Measuring and improving social impacts: A guide for nonprofits, companies, and impact investors*. San Francisco: Berrett-Koehler Publishers.

Erdley, C. A., & Asher, S. R. (1998). Linkages between Children's Beliefs about the Legitimacy of Aggression and Their Behavior. *Social Development*, *7*(3), 321–339. doi:10.1111/1467-9507.00070

Esmer, Y. R., & Pettersson, T. (2007). *Measuring and mapping cultures: 25 years of comparative value surveys*. *International studies in sociology and social anthropology*. Leiden, Boston: Brill.

European Social Survey. (2014a). *Documentation of ESS post-stratification weights*. Retrieved from http://www.europeansocialsurvey.org/

European Social Survey. (2014b). *ESS Round 7 Source Questionnaire*. London.

European Social Survey. (2014c). *Gesellschaft und Demokratie in Europa: Deutsche Teilstudie im Projekt „European Social Survey" (Welle 7)*. London.

European Social Survey. (2014d). *Weighting European Social Survey Data*. Retrieved from http://www.europeansocialsurvey.org/

Evans, I., Chalmers, I., & Thornton, H. (2006). *Testing treatments: Better research for better healthcare*. London: British Library.

Fanshel, S., & Bush, J. W. (1970). A Health-Status Index and its Application to Health-Services Outcomes. *Operations Research*, *18*(6), 1021–1066. doi:10.1287/opre.18.6.1021

Farrell, A. D., Erwin, E. H., Allison, K. W., Meyer, A., Sullivan, T., Camou, S., … (2007). Problematic Situations in the Lives of Urban African American Middle School Students: A Qualitative Study. *Journal of Research on Adolescence*, *17*(2), 413–454. doi:10.1111/j.1532-7795.2007.00528.x

Farrell, A. D., Erwin, E. H., Bettencourt, A., Mays, S., Vulin-Reynolds, M., Sullivan, T., … (2008). Individual Factors Influencing Effective Nonviolent Behavior and Fighting in Peer Situations: A Qualitative Study with Urban African American Adolescents. *Journal of Clinical Child & Adolescent Psychology*, *37*(2), 397–411. doi:10.1080/15374410801955821

Ferraro, F., Etzion, D., & Gehman, J. (2015). Tackling Grand Challenges Pragmatically: Robust Action Revisited. *Organization Studies*, *36*(3), 363–390. doi:10.1177/0170840614563742

Fine, B. (2010). *Theories of social capital: Researchers behaving badly*. *Political economy and development*. London, New York, New York: Pluto Press.

Fioramonti, L. (2013). *Gross domestic problem: The politics behind the world`s most powerful number*. *Economic Controversies*. London: Zed Books Ltd.

Flap, H., & Völker, B. (2004). Academy Colloquium on "Creation and Returns of Social Capital. the State of the Art. *Bulletin de Methologie Sociologique*, *81*(1), 66–80. doi:10.1177/075910630408100106

Fox, C., & Albertson, K. (2011). Payment by results and social impact bonds in the criminal justice sector: New challenges for the concept of evidence-based policy? *Criminology and Criminal Justice*, *11*(5), 395–413. doi:10.1177/1748895811415580

Frączek, A., & Zumkley, H. (Eds.). (1992). *Recent research in psychology*. *Socialization and aggression*. Berlin, New York: Springer.

Freeman, R. E. (2010). *Strategic management: A stakeholder approach*. Cambridge: Cambridge University Press.

Freeman, R. E., Harrison, J. S., Wicks, A. C., Parmar, B. L., & Colle, S. de. (2010). *Stakeholder Theory: The State of the Art*: Cambridge University Press.

Freise, M., & Hallmann, T. (Eds.). (2014). *Modernising Democracy: Associations and associating in the 21st century*. New York: Springer.

French, R. D. (2011). Political Capital. *Representation*, *47*(2), 215–230. doi:10.1080/00344893.2011.581086

Frey, B. S. (2010). *Happiness: A revolution in economics*. Cambridge, Mass: Cambridge University Press; MIT Press.

Frey, B. S., & Stutzer, A. (2002a). *Happiness and economics: How the economy and institutions affect well-being*. Princeton, N.J.: Princeton University Press.

Frey, B. S., & Stutzer, A. (2002b). The Economics of Happiness. *World Economics*, *3*(1), 2–17.

Friedman, M. (1970, September 13). The Social Responsibility of Business is to Increase its Profits. *New York Times Magazine*. Retrieved from http://highered.mcgraw-hill.com/sites/dl/free/0073524697/910345/Appendices.pdf

Garriga, E., & Melé, D. (2004). Corporate Social Responsibility Theories: Mapping the Territory. *Journal of Business Ethics*, *53*, 51–71.

Gefangene helfen Jugendlichen (GHJ). (2017). *Über GHJ*. Retrieved from https://www.gefangene-helfen-jugendlichen.de/ueber-ghj/

Gil De Zúñiga, H., Puig-I-Abril, E., & Rojas, H. (2009). Weblogs, traditional sources online and political participation: An assessment of how the internet is changing the political environment. *New Media & Society*, *11*(4), 553–574. doi:10.1177/1461444809102960

Ginev, D. (1999). On the hermeneutic fore-structure of scientific research. *Continental Philosophy Review*, *32*(2), 143–168. doi:10.1023/A:1006966329046

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New Jersey: AldineTransaction.

Gläser, J., & Laudel, G. (2013). Life With and Without Coding: Two Methods for Early-Stage Data Analysis in Qualitative Research Aiming at Causal Explanations. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, *14*(2), Art. 5.

Glasziou, P., & Heneghan, C. (2009). A spotter's guide to study designs. *Evidence-based medicine*, *14*(2), 37–38. doi:10.1136/ebm.14.2.37-a

Global Impact Investing Network. (2010). *Impact Reporting and Investment Standards (IRIS)*. Retrieved from http://socialvalueint.org/wp-content/uploads/2012/05/IRIS-Overview-2010.pdf

Gluckler, J., & Armbruster, T. (2003). Bridging Uncertainty in Management Consulting: The Mechanisms of Trust and Networked Reputation. *Organization Studies*, *24*(2), 269–297. doi:10.1177/0170840603242004

Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report*, *8*(4), 597–607.

Goldin, C. (2016). Human Capital. In C. Diebolt & M. J. Haupert (Eds.), *Handbook of cliometrics* (pp. 55–86). Heidelberg: SpringerReference.

Goldthorpe, J. (2007). "Cultural Capital": Some Critical Observations. *Sociological Inquiry*, (2), 1–23.

Graham, C. (2005). The Economics of Happiness: Insights on globalization from a novel approach. *World Economics*, *6*(3), 41–55.

Gray, R. (2002). The social accounting project and Accounting Organizations and Society: Privileging engagement, imaginings, new accountings and pragmatism over critique? *Accounting, Organizations and Society*, *27*, 687–708.

Gray, R., Kouhy, R., & Lavers, S. (1995). Corporate social and environmental reporting: a review of the literature and a longitudinal study of UK disclosure. *Accounting, Auditing & Accountability Journal*, *8*(2), 47–77. doi:10.1108/09513579510146996

Gray-Little, B., Williams, V. S. L., & Hancock, T. D. (1997). An Item Response Theory Analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, *23*(5), 443–451. doi:10.1177/0146167297235001

Grootaert, C., Narayan, D., Nyhan Jones, V., & Woolcock, M. (2004). *Measuring social capital: An integrated questionnaire*. *World Bank working paper: Vol. 18*. Washington, D.C: World Bank.

Grootaert, C. & van Bastelaer, T. (2001). *Understanding and measuring social capital: A synthesis of findings and recommendations from the Social Capital initiative*. Social Capital Initiative Working Paper No. 24. Washington, DC.

Grootaert, C., & van Bastelaer, T. (2002). *Understanding and measuring social capital: A multidisciplinary tool for practitioners*. *Directions in development (Washington, D.C.)*. Washington, D.C: World Bank. Retrieved from http://elibrary.worldbank.org/content/book/9780821350683

Grossman, A., Appleby, S., & Reimers, C. (2013). *"Venture Philanthropy: Its Evolution and Its Future."*. Harvard Business School Background Note 313-111. Retrieved from http://evpa.eu.com/wp-content/uploads/2010/11/VP-Its-Evolution-and-Its-Future-Final.pdf

Gubrium, J. F., & Holstein, J. A. (Eds.). (2001). *Handbook of interview research: Context & method*. Thousand Oaks, Calif: Sage Publications.

Guerra, N. G., Boxer, P., & Cook, C. R. (2006). What works (and what does not) in youth violence prevention: Rethinking the questions and finding new answers. *New Directions for Evaluation*, *2006*(110), 59–71. doi:10.1002/ev.187

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, *18*, 59–82. doi:10.1177/1525822X05279903

Guillen, L., Coromina, L., & Saris, W. E. (2011). Measurement of Social Participation and its Place in Social Capital Theory. *Social Indicators Research*, *100*(2), 331–350. doi:10.2307/41476397

Gyrd-Hansen, D. (2003). Willingness to pay for a QALY. *Health Economics*, *12*(12), 1049–1060. doi:10.1002/hec.799

Haas, M. de, & Kleingeld, A. (1999). Multilevel design of performance measurement systems: Enhancing strategic dialogue throughout the organization. *Management Accounting Research*, *10*(3), 233–261. doi:10.1006/mare.1998.0098

Habermas, J. (1984). *The theory of communicative action*. Tranlated by MacCarthy, Thomas A. Boston: Beacon.

Hagerty, B. M. K., & Patusky, K. (1995). Developing a Measure Of Sense of Belonging. *Nursing Research*, *44*(1), 9–13. doi:10.1097/00006199-199501000-00003

Hansmann, H. (1980). The Role of Nonprofit Enterprise. *The Yale Law Journal*, *89*(8), 835–902.

Hansmann, H. (1987). Economic Theories of Nonprofit Organization. In W. W. Powell (Ed.), *The Nonprofit Sector. A Research Handbook* (pp. 27–42). New Haven: Yale University Press.

Hanson, H. M., Ashe, M. C., MacKay, H. A., & Winters, M. G. (2012). *Intersection between the Built and Social Environments and Older Adults' Mobility: An Evidence Review*: NCCEH.

Hanson, H. M., Schiller, C., Winters, M., Sims-Gould, J., Clarke, P., Curran, E., … (2013). Concept mapping applied to the intersection between older adults' outdoor walking and the built and social environments. *Preventive Medicine*, *57*(6), 785–791. doi:10.1016/j.ypmed.2013.08.023

Hansson, B. (2008). *Job-Related Training and Benefits for Individuals: A Review of Evidence and Explanations*. OECD Education Working Papers (Vol. 19).

Hatzfeld, C. A. (2014). Performance Measurement in Non-Profit Organisations: Exploring the Commonalities between Impact, Outcome and Performance Measurement in Open Youth Work Providers. *International Journal of Management Cases*, *16*(4), 76–92.

Hawes, C., Phillips, C. D., Rose, M., Holan, S., & Sherman, M. (2003). A National Survey of Assisted Living Facilities. *The Gerontologist*, *43*(6), 875–882. doi:10.1093/geront/43.6.875

Hawken, P., Lovins, A. B., & Lovins, L. H. (1999). *Natural capitalism: The next industrial revolution*. London: Earthscan.

Helm, S., & Baker, C. (2011). The Need to Consider Ethnocultural Context in Prevention Programming: A Case Example from Hawai'i. *Journal of Ethnic And Cultural Diversity in Social Work*, *20*(2), 131–149. doi:10.1080/15313204.2011.570125

Hennings, K. H. (2005). Capital as a factor of production. In C. J. Bliss, A. J. Cohen, & G. C. Harcourt (Eds.), *An Elgar reference collection. Capital theory* (pp. 3–17). Cheltenham, UK, Northampton, MA: Edward Elgar Pub.

Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., … (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*, *20*(10), 1727–1736. doi:10.1007/s11136-011-9903-x

Hernes, T. (2009). *Understanding organization as process: Theory for a tangled world*. *Routledge studies in management, organisation and society: Vol. 2*. London: Routledge.

Hockerts, K. (2015). How Hybrid Organizations Turn Antagonistic Assets into Complementarities. *California Management Review*, *57*(3), 83–106. doi:10.1525/cmr.2015.57.3.83

Høgsbro, K., & Shaw, I. (2017). *Social work and research in advanced welfare states*. *Routledge advances in social work*. Abingdon, Oxon, New York, NY: Routledge.

Homish, G. G., Edwards, E. P., Eiden, R. D., & Leonard, K. E. (2010). Analyzing family data: A GEE approach for substance use researchers. *Addictive behaviors*, *35*(6), 558–563. doi:10.1016/j.addbeh.2010.01.002

Howard, M. M., Gibson, J. L., & Stolle, D. (2006). *United States Citizenship, Involvement, Democracy (CID) Survey: ICPSR04607-v2*. Ann Arbor. Retrieved from http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/4607

Hubbard, A. E., Ahern, J., Fleischer, N. L., van der Laan, M., Lippman, S. A., Jewell, N., … (2010). To GEE or Not to GEE. *Epidemiology*, *21*(4), 467–474. doi:10.1097/EDE.0b013e3181caeb90

Hume, D. (1978). *A treatise of human nature* (Second edition with an analytical index by L. A. Selby-Biggewith, and with text revised and variant readings by P.H. Nidditch). Oxford: Clarendon Press.

Huntington, S. P., & Nelson, J. M. (1984). *No Easy Choice: Political Participation in Developing Countries*: Harvard University Press. Retrieved from http://www.lob.de/cgi-bin/work/suche2?titnr=211817308&flag=citavi

Huppes, G., & Ishikawa, M. (2005). A framework for quantified eco-efficiency analysis. *Journal of Industrial Ecology*, *9*(4), 25–41.

Hwang, H., & Powell, W. W. (2009). The Rationalization of Charity: The Influences of Professionalism in the Nonprofit Sector. *Administrative Science Quarterly*, *54*(2), 268–298. doi:10.2189/asqu.2009.54.2.268

Iacus, S. M., King, G., & Porro, G. (2011). Causal Inference Without Balance Checking: Coarsened Exact Matching. *Political Analysis*, *20*(1), 1–24.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of economic literature*, *47*(1), 5–86. doi:10.1257/jel.47.1.5

Irak, D. U., & Mantler, J. (2017). The role of temporal flexibility on person–environment fit and job satisfaction. *Journal of Management & Organization*, *15*, 1–17. doi:10.1017/jmo.2017.50

Israel, J. (1972). Is a Non-Normative Social Science Possible? *Acta Sociologica*, *15*(1), 69–89.

Jackson, M. O., Rodriguez-Barraquer, T., & Tan, X. (2012). Social Capital and Social Quilts: Network Patterns of Favor Exchange. *American Economic Review*, *102*(5), 1857–1897. doi:10.1257/aer.102.5.1857

Jacobi, N. von, Edmiston, D., & Ziegler, R. (2017). Tackling Marginalisation through Social Innovation?: Examining the EU Social Innovation Policy Agenda from a Capabilities Perspective. *Journal of Human Development and Capabilities*, *18*(2), 148–162. doi:10.1080/19452829.2016.1256277

Jacobs, D. C., & Getz, K. A. (1995). Dialogue on the Stakeholder Theory of the Corporation: Concepts, Evidence, and Implications. *The Academy of Management Review*, *20*(4), 793–795.

Jones, C., Hesterly, W. S., & Borgatti, S. P. (2005). A General Theory of Network Governance: Exchange Conditions and Social Mechanisms. *Academy of Management Review*, *22*(4), 911–945.

Kahneman, D., & Thaler, R. H. (2006). Anomalies: Utility Maximization and Experienced Utility. *Journal of Economic Perspectives*, *20*(1), 221–234. doi:10.1257/089533006776526076

Kalaycioglu, E., & Turan, I. (1981). Measuring Political Participation: A Cross-Cultural Application. *Comparative Political Studies*, *14*, 123–135.

Kaplan, R. S. (2001). Strategic Performance Measurement and Management in Nonprofit Organizations. *Nonprofit Management and Leadership*, *11*(3), 353–370. doi:10.1002/nml.11308

Kapteyn, A., Smith, J. P., & van Soest, A. (2010). Life Satisfaction. In E. Diener, J. F. Helliwell, & D. Kahneman (Eds.), *Series in positive psychology. International differences in well-being* (pp. 70–104). Oxford, New York: Oxford University Press.

Kauppi, N. (2003). Bourdieu's political sociology and the politics of European integration. *Theory and Society*, *32*(5/6), 775–789. doi:10.1023/B:RYSO.0000004919.28417.7c

Kehl, K., Krlev, G., Then, V., & Mildenberger, G. (2018). Adapting the measuring rod for social returns in advanced welfare states: A critique of SROI. In A. Lindgreen, B. Hirsch, C. Vallaster, & S. Yousafzai (Eds.), *Measuring and controlling sustainability. Spanning theory and practice.* Routledge.

Kehl, K., & Then, V. (2013). Community and Civil Society Returns of Multi-generation Cohousing in Germany. *Journal of Civil Society*, *9*(1), 41–57. doi:10.1080/17448689.2013.771084

Kehl, K., Then, V., & Münscher, R. (2012). Social Return on Investment: Auf dem Weg zu einem integrativen Ansatz der Wirkungsforschung. In H. K. Anheier, A. Schröer, & V. Then (Eds.), *Soziale Investitionen. Interdisziplinäre Perspektiven.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Keller, R. (2011). The Sociology of Knowledge Approach to Discourse (SKAD). *Human Studies*, *34*(1), 43. doi:10.1007/s10746-011-9175-z

Kelly, T. (2007). Evidence and Normativity: Reply to Leite. *Philosophy and Phenomenological Research*, *75*(2), 465–474. doi:10.1111/j.1933-1592.2007.00085.x

Kendall, J., & Knapp, M. (2000). Measuring the Performance of Voluntary Organizations. *Public Management Review*, *2*(1), 105–132. doi:10.1080/14719030000000006

Khawaja, M., & Mowafi, M. (2006). Cultural Capital and Self-Rated Health in Low Income Women: Evidence from the Urban Health Study, Beirut, Lebanon. *Journal of Urban Health*, *83*(3), 444–458. doi:10.1007/s11524-006-9051-8

King, G., Keohane, R. O., & Verba, S. (1995). *Designing social inquiry: Scientific inference in qualitative research* (3rd). Princeton, N.J.: Princeton University Press.

King, G., & Roberts, M. E. (2015). How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What to Do About It. *Political Analysis*, *23*(2), 159-179.

Kippin, H., Hauf, H., & Sadique, A. (2012). *Business, society and public services: a social productivity framework*. London. Retrieved from http://www.thersa.org/__data/assets/pdf_file/0004/583177/2020PSH_business_in_society_report.pdf

Kippin, H. & Lucas, B. (2010). *From Big Society to Social Productivity*. London. Retrieved from http://www.thersa.org/__data/assets/pdf_file/0013/1530310/RSA_2020_BigSociety.pdf

Kjellberg, H., Mallard, A., Arjaliès, D.-L., Lamont, M., Aspers, P., Beljean, S., & Bidet, A. (2013). Valuation Studies? Our Collective Two Cents. *Valuation Studies*, *1*(1), 11–30.

Kolb, R. W. (2008). *Encyclopedia of business ethics and society*. Thousand Oaks: Sage Publications.

Koppell, J. G. S. (2005). Pathologies of Accountability: ICANN and the Challenge of "Multiple Accountabilities Disorder". *Public Administration Review*, *65*(1), 94–108.

Krlev, G., Münscher, R., & Mülbert, K. (2013). *Social Return on Investment (SROI): State-of-the-Art and Perspectives*. CSI Report. Retrieved from https://www.csi.uni-heidelberg.de/downloads/CSI_SROI_Meta_Analysis_2013.pdf

Kroeger, A., & Weber, C. (2014). Developing a Conceptual Framework for Comparing Social Value Creation. *Academy of Management Review*, *39*(4), 513–540. doi:10.5465/amr.2012.0344

Lamont, M., & Lareau, A. (1988). Cultural Capital: Allusions, Gaps and Glissandos in Recent Theoretical Developments. *Sociological Theory*, *6*(2), 153. doi:10.2307/202113

Langley, A., Smallman, C., Tsoukas, H., & van de Ven, A. H. (2013). Process Studies of Change in Organization and Management: Unveiling Temporality, Activity, and Flow. *Academy of Management Journal*, *56*(1), 1–13. doi:10.5465/amj.2013.4001

Lareau, A., & Weininger, E. B. (2005). Cultural capital in educational research: A critical assessment. In D. L. Swartz & V. L. Zolberg (Eds.), *SpringerLink : Bücher. After Bourdieu. Influence, Critique, Elaboration* (pp. 105–144). Dordrecht: Springer Science + Business Media.

Layard, R., & Glaister, S. (1994). *Cost-benefit analysis* (2nd). Cambridge: Cambridge University Press.

Ledogar, R. J., & Fleming, J. (2008). Social Capital and Resilience: A Review of Concepts and Selected Literature Relevant to Aboriginal Youth Resilience Research. *Pimatisiwin*, *6*(2), 25–46.

Lee, B. K., Lessler, J., & Stuart, E. A. (2011). Weight trimming and propensity score weighting. *PLoS ONE*, *6*(3), e18174. doi:10.1371/journal.pone.0018174

Li, Y., & Marsh, D. (2008). New Forms of Political Participation: Searching for Expert Citizens and Everyday Makers. *British Journal of Political Science*, *38*(02). doi:10.1017/S0007123408000136

Liket, K. C., Rey-Garcia, M., & Maas, K. E. H. (2014). Why Aren't Evaluations Working and What to Do About It. *American Journal of Evaluation*, *35*(2), 171–188. doi:10.1177/1098214013517736

Lillbacka, R. (2006). Measuring Social Capital: Assessing Construct Stability of Various Operationalizations of Social Capital in a Finnish Sample. *Acta Sociologica*, *49*(2), 201–220.

Lin, H., Zeng, S., Ma, H., Qi, G., & Tam, V. W. (2014). Can political capital drive corporate green innovation? Lessons from China. *Journal of Cleaner Production*, *64*, 63–72. doi:10.1016/j.jclepro.2013.07.046

Lin, N. (1999). Building a Network Theory of Social Capital'. *Connections*, *22*(1), 28–51.

Lin, N. (2001). *Social capital: A theory of social structure and action*. *Structural analysis in the social sciences: Vol. 19*. Cambridge, New York: Cambridge University Press.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, Calif.: Sage.

Lindgreen, A., Hirsch, B., Vallaster, C., & Yousafzai, S. (Eds.). (2018). *Measuring and controlling sustainability: Spanning theory and practice*: Routledge.

Linz, J. J., & Stepan, A. (1996). *Problems of democratic transition and consolidation: Southern Europe, South America, and post-communist Europe*. Baltimore and London: Johns Hopkins.

Locke, K., & Golden-Biddle, K. (1997). Constructing Opportunities for Contribution: Structuring Intertextual Coherence and "Problematizing" in Organizational Studies. *The Academy of Management Journal*, *40*(5), 1023–1062. doi:10.2307/256926

Lösel, F., & Stemmler, M. (2012). Preventing Child Behavior Problems in the Erlangen-Nuremberg Development and Prevention Study: Results from Preschool to Secondary School Age. *International Journal of Conflict and Violence*, *6*(2), 215–224.

Lubben, J., Blozik, E., Gillmann, G., Iliffe, S., Renteln Kruse, W. von, Beck, J. C., & Stuck, A. E. (2006). Performance of an Abbreviated Version of the Lubben Social Network Scale Among Three European Community-Dwelling Older Adult Populations. *The Gerontologist*, *46*(4), 503–513. doi:10.1093/geront/46.4.503

Lund-Chaix, A. A., & Gelles, E. (2013). A Cultural Capital Perspective of the Effect of a Government–Voluntary Sector Partnership for Enhancing Access to Postsecondary Education. *Nonprofit and Voluntary Sector Quarterly*, *43*(3), 436–454. doi:10.1177/0899764012470028

Maas, K. (2009). *Corporate Social Performance. From Output Measurement to Impact Measurement* (Doctoral Thesis). Erasmus University Rotterdam, Rotterdam.

Mackor, A. R. (2011). Explanatory non-normative legal doctrine: Taking the distinction between theoretical and practical reason seriously. In M. van Hoecke (Ed.), *European Academy of Legal Theory Series. Methodologies of Legal Research. Which Kind of Method for What Kind of Discipline?* (pp. 45–70). London: Bloomsbury Publishing.

Mahoney, L. S., & Roberts, R. W. (2004). Corporate Social Performance. In C. Jeffrey (Ed.), *Research on Professional Responsibility and Ethics in Accounting: Vol. 9. Research on Professional Responsibility and Ethics in Accounting* (pp. 73–99). Emerald Group Publishing Limited.

Maier, F., Schober, C., Simsa, R., & Millner, R. (2014). SROI as a Method for Evaluation Research: Understanding Merits and Limitations. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations.* doi:10.1007/s11266-014-9490-x

Mair, J., Battilana, J., & Cardenas, J. (2012). Organizing for Society: A Typology of Social Entrepreneuring Models. *Journal of Business Ethics*, *111*(3), 353–373. doi:10.1007/s10551-012-1414-3

Mair, J., & Hehenberger, L. (2014). Front-Stage and Backstage Convening: The Transition from Opposition to Mutualistic Coexistence in Organizational Philanthropy. *Academy of Management Journal*, *57*(4), 1174–1200. doi:10.5465/amj.2012.0305

Marée, M., & Mertens, S. (2012). The limits of economic value in measuring performance of social innovation. In A. Nicholls & A. Murdock (Eds.), *Social innovation. Blurring boundaries to reconfigure markets* (pp. 114–138). Houndmills, Basingstoke: Palgrave Macmillan.

Maring, E. F., & Koblinsky, S. A. (2013). Teachers' Challenges, Strategies, and Support Needs in Schools Affected by Community Violence: A Qualitative Study. *Journal of School Health*, *83*(6), 379–388. doi:10.1111/josh.12041

Martin, G. P. (2011). The Third Sector, User Involvement and Public Service Reform: A case study in the co-governance of health service provision. *Public Administration*, *89*(3), 909–932. doi:10.1111/j.1467-9299.2011.01910.x

Maxwell, J. A. (2012). *Qualitative research design: An interactive approach* (3. ed.). *Applied social research methods series: Vol. 41*. London: Sage.

McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., & Burgette, L. F. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in medicine*, *32*(19), 3388–3414. doi:10.1002/sim.5753

McCallister, L., & Fischer, C. S. (1978). A Procedure for Surveying Personal Networks. *Sociological Methods & Research*, *7*(2), 131–148. doi:10.1177/004912417800700202

McDonald, R. E. (2007). An Investigation of Innovation in Nonprofit Organizations: The Role of Organizational Mission. *Nonprofit and Voluntary Sector Quarterly*, *36*(2), 256–281. doi:10.1177/0899764006295996

McWilliams, A., & Siegel, D. (2001). Corporate Social Responsibility: A theory of the firm perspective. *Academy of Management Review*, *26*(1), 117–127. doi:10.5465/AMR.2001.4011987

Mekouar, M. (2016). *Protest and mass mobilization: Authoritarian collapse and political change in North Africa*. London and New York: Routledge, Taylor et Francis Group.

Meppem, T., & Gill, R. (1998). Planning for sustainability as a learning concept. *Ecological Economics*, *26*(2), 121–137. doi:10.1016/S0921-8009(97)00117-1

Merriam, S. B. (1995). What can you tell from an N of 1?: Issues of validity and reliability in qualitative research. *PACE Journal of Life Long Learning*, *4*, 51–60.

Meyer, R. P., & Schuyler, D. (2011). Old age and loneliness. *The primary care companion for CNS disorders*, *13*(2), e1-2. doi:10.4088/PCC.11f01172

Micheli, P., & Kennerley, M. (2005). Performance measurement frameworks in public and non-profit sectors. *Production Planning & Control*, *16*(2), 125–134. doi:10.1080/09537280512331333039

Miettinen, O. (1974). Confounding and effect modification. *American Journal of Epidemiology*, *100*(5), 350–353. doi:10.1093/oxfordjournals.aje.a112044

Migration Advisory Committee. (2012). *Analysis of the impact of migration*. London. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/257235/analysis-of-the-impacts.pdf

Milanovic, B. (2015). *Junk the phrase 'human capital'*. Retrieved from http://america.aljazeera.com/opinions/2015/2/junk-the-phrase-human-capital.html

Mildenberger, G., Münscher, R., & Schmitz, B. (2012). Dimensionen der Bewertung gemeinnütziger Organisationen und Aktivitäten. In H. K. Anheier, A. Schröer, & V. Then

(Eds.), *Soziale Investitionen. Interdisziplinäre Perspektiven* (pp. 279–312). Wiesbaden: VS Verlag für Sozialwissenschaften.

Miles, R. E., Snow, C. C., Meyer, A. D., & Coleman, H. J. (1978). Organizational Strategy, Structure, and Process. *Academy of Management Review, 3*(3), 546–562.

Millar, R., & Hall, K. (2013). Social Return on Investment (SROI) and Performance Measurement. *Public Management Review, 15*(6), 923–941. doi:10.1080/14719037.2012.698857

Miller, D. C., & Salkind, N. J. (2002). *Handbook of research design & social measurement* (6. ed.). Thousand Oaks, Calif.: Sage.

Milton, B., Attree, P., French, B., Povall, S., Whitehead, M., & Popay, J. (2012). The impact of community engagement on health and social outcomes: a systematic review. *Community Development Journal*, *47*(3), 316–334.

Miyamoto, J. M., & Eraker, S. A. (1985). Parameter Estimates for a QALY Utility Model. *Medical Decision Making*, *5*(2), 191–213.

Molewijk, B., Stiggelbout, A. M., Otten, W., Dupuis, H. M., & Kievit, J. (2008). First the facts, then the values?: Implicit normativity in evidence-based decision aids for shared decision-making. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, *102*(7), 415–420. doi:10.1016/j.zefq.2008.08.014

Möller, A., & Schaltegger, S. (2005). The sustainability balanced scorecard as a framework for eco-efficiency analysis. *Journal of Industrial Ecology*, *9*(4), 73–83.

Mook, L. (Ed.). (2013). *Accounting for Social Value*. Toronto: University of Toronto Press.

Mook, L., & Quarter, J. (2006). Accounting for the Social Economy: The Socioeconomic Impact Statement. *Annals of Public and Cooperative Economics*, *77*(2), 247–269.

Moore, M. H. (2000). Managing for value: Organizational strategy in for-profit, nonprofit, and governmental organizations. *Nonprofit and Voluntary Sector Quarterly*, *29*(1 Suppl.), 183–204.

Morrone, A., Tontoranelli, N., & Ranuzzi, G. (2009). *How Good is Trust?: Measuring Trust and its Role for the Progress of Societies*. OECD Statistics Working Papers.

Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification Strategies for Establishing Reliability and Validity in Qualitative Research. *International Journal of Qualitative Methods*, *1*(2), 13–22.

Mossberger, K., Tolbert, C. J., & McNeal, R. S. (2010). *Digital citizenship: The internet, society, and participation* (2.th ed.). Cambridge, Mass.: MIT Press.

Moynihan, D. P. (2008). *The dynamics of performance management: Constructing information and reform*. *Public management and change*. Washington, D.C.: Georgetown University Press.

Mruk, C. J. (2006). *Self-esteem research, theory, and practice: Toward a positive psychology of self-esteem* (3rd ed.). New York: Springer Pub.

Mulgan, G., Reeder, N., Aylott, M., & Bo'sher, L. (2010). *Social Impact Investment: the challenge and opportunity of Social Impact Bonds.*

Murphy, K. R., & Davidshofer, C. O. (2005). *Psychological testing: Principles and applications* (6th ed.). Upper Saddle River, N.J.: Pearson/Prentice Hall.

Musgrave, R. A. (1951). A Multiple Theory of Budget Determination. *FinanzArchiv*, *25*(1), 33–43.

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying Behaviors Among US Youth. *JAMA, 285*(16), 2094. doi:10.1001/jama.285.16.2094

Narayan, D., & Cassidy, M. F. (2001). A Dimensional Approach to Measuring Social Capital: Development and Validation of a Social Capital Inventory. *Current Sociology, 49*(2), 59–102.

Nas, T. F. (2016). *Cost-benefit analysis: Theory and application* (2nd ed.). Lanham: Lexington Books.

Nee, V., & Opper, S. (2010). Political Capital in a Market Economy. *Social Forces, 88*(5), 2105–2132. Retrieved from http://www.jstor.org/stable/40927540

Neely, A., Gregory, M., & Platts, K. (1995). Performance measurement system design: A literature review and research agenda. *International Journal of Operations & Production Management, 15*(4), 80–116. doi:10.1108/01443579510083622

Netzwerk: Soziales neu gestalten (Ed.). (2009). *Zukunft Quartier: Lebensräume zum Älterwerden*. Band 3: Soziale Wirkung und "Social Return". Gütersloh: Verlag Bertelsmann Stiftung.

Nicholls, A. (Ed.). (2006). *Social entrepreneurship: New models of sustainable social change*. Oxford: Univ. Press. Retrieved from http://www.gbv.de/dms/bowker/toc/9780199283880.pdf

Nicholls, A. (2009a). 'We do good things, don't we?': 'Blended Value Accounting' in social entrepreneurship. *Accounting, Organizations and Society, 34*(6-7), 755–769. doi:10.1016/j.aos.2009.04.008

Nicholls, A. (2009b). Capturing the Performance of the Socially Entrepreneurial Organization: An Organizational Legitimacy Approach. In J. A. Robinson, J. Mair, & K. Hockerts (Eds.), *International Perspectives on Social Entrepreneurship* (pp. 27–64). Basingstoke, Hampshire: Palgrave Macmillan.

Nicholls, A. (2010a). Institutionalizing social entrepreneurship in regulatory space: Reporting and disclosure by community interest companies. *Accounting, Organizations and Society, 35*(4), 394–415. doi:10.1016/j.aos.2009.08.001

Nicholls, A. (2010b). The institutionalization of social investment: The interplay of investment logics and investor rationalities. *Journal of Social Entrepreneurship, 1*(1), 70–100.

Nicholls, A., & Murdock, A. (Eds.). (2012a). *Social innovation: Blurring boundaries to reconfigure markets*. Houndmills, Basingstoke: Palgrave Macmillan.

Nicholls, A., & Murdock, A. (2012b). The nature of Social Innovation. In A. Nicholls & A. Murdock (Eds.), *Social innovation. Blurring boundaries to reconfigure markets* (pp. 1–32). Houndmills, Basingstoke: Palgrave Macmillan.

Nicholls, A., & Teasdale, S. (2017). Neoliberalism by stealth?: Exploring continuity and change within the UK social enterprise policy paradigm. *Policy & Politics, 45*(3), 323–341. doi:10.1332/030557316X14775864546490

Nicholls, J., Lawlor, E., Neitzert, E., & Goodspeed, T. (2009). *A guide to Social Return on Investment.*

Nord, E. (1994). The QALY - a measure of social value rather than individual utility? *Health Economics, 3,* 89–93.

Norris, P. (2013). *Democratic phoenix: Reinventing political activism*. Cambridge, New York: Cambridge University Press.

Norstrand, J. A., Glicksman, A., Lubben, J., & Kleban, M. (2012). The Role of the Social Environment on Physical and Mental Health of Older Adults. *Journal of Housing For the Elderly*, *26*(1-3), 290–307. doi:10.1080/02763893.2012.651382

Nussbaum, M. (1993). Non-Relative Virtues: An Aristotelian Approach. In M. Nussbaum & A. Sen (Eds.), *The Quality of Life* (pp. 242–269). Oxford University Press.

Nussbaum, M. (2003). Capabilities as fundamental entitlements: Sen and social justice. *Feminist Economics*, *9*(2-3), 33–59. doi:10.1080/1354570022000077926

O'Brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity*, *41*(5), 673–690. doi:10.1007/s11135-006-9018-6

OECD. (2016). *Society at a Glance 2016*. Paris: OECD Publishing.

Olsen, S. & Galimidi, B. (2008). *Catalog of approaches to impact measurement.*

Onyx, J., & Bullen, P. (2000). Measuring Social Capital in Five Communities. *The Journal of Applied Behavioral Science*, *36*(1), 23–42.

Ormel, J., Lindenberg, S., Steverink, N., & Verbrugge, L. (1999). Subjective Well-Being and Social Production Functions. *Social Indicators Research*, *46*(1), 61–90. doi:10.1023/A:1006907811502

Orpinas, P. (1993). *Skills training and social influences for violence prevention in middle schools: A curriculum evaluation* (Doctoral Dissertation). University of Texas Health Science Center at Houston, School of Public Health, Houston, Texas.

Orpinas, P., & Frankowski, R. (2001). The Aggression Scale: A Self-Report Measure of Aggressive Behavior for Young Adolescents. *The Journal of Early Adolescence*, *21*(1), 50–67. doi:10.1177/0272431601021001003

Osmani, S. R. (2016). *The Capability Approach and Human Development: Some Reflections: UNDP Human Development Report - Think Piece*. Retrieved from http://hdr.undp.org/sites/default/files/osmani_template.pdf

Paldam, M. (2000). Social Capital: One or Many? Definition and Measurement. *Journal of Economic Surveys*, *14*(5), 629–653. doi:10.1111/1467-6419.00127

Pan, W. (2001). Akaike's Information Criterion in Generalized Estimating Equations. *Biometrics*, *57*(1), 120–125. doi:10.1111/j.0006-341X.2001.00120.x

Paton, R. (2003). *Managing and measuring social enterprises*. London, Thousand Oaks: Sage Publications.

Patras, J., & Klest, S. K. (2012). Development of a collective efficacy measure for use in social service organizations. *Journal of Social Work*, *13*(1), 96–106. doi:10.1177/1468017311412034

Patton, M. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, Calif.: Sage.

Pawson, R. (2002). Evidence-based Policy: In Search of a Method. *Evaluation*, *8*(2), 157–181. doi:10.1177/1358902002008002512

Pearce, D. (1998). Cost benefit analysis and environmental policy. *Oxford Review of Economic Policy*, *14*(4), 84–100. doi:10.1093/oxrep/14.4.84

Perkins, D. D., Florin, P., Rich, R. C., Wandersman, A., & Chavis, D. M. (1990). Participation and the social and physical environment of residential blocks: Crime and community context. *American Journal of Community Psychology*, *18*(1), 83–115. doi:10.1007/BF00922690

Peter, J. P. (1979). Reliability: A Review of Psychometric Basics and Recent Marketing Practices. *Journal of Marketing Research*, *16*(1), 6–17. doi:10.2307/3150868

Pigou, A. C. (1952). *The economics of welfare*. New York: Cosimo Classics.

Poister, T. H., Hall, J. L., & Aristigueta, M. P. (2015). *Managing and measuring performance in public and nonprofit organizations: An integrated approach* (2nd ed.). San Francisco: Jossey-Bass.

Pollak, R. A. (1979). Bergson-Samuelson Social Welfare Functions and the Theory of Social Choice. *The Quarterly Journal of Economics*, *93*(1), 73–90.

Pope, H. C., Miller, M. C., Wolfer, T. A., Mann, J. R., & McKeown, R. E. (2013). Psychometric Analysis of a Scale to Assess Norms of Reciprocity of Social Support in Community-Based and Congregation-Based Groups. *SAGE Open*, *3*(4). doi:10.1177/2158244013516771

Porter, M. E., & Kramer, M. R. (2011). Creating Shared Value: How to reinvent capitalism - and unleash a wave of innovation and growth. *Harvard Business Review*, (January-February), 2–17.

Porter, M. E., Stern, S., & Loria, R. A. (2013). *Social progress report 2013.*

Portes, A. (1998). Social Capital: Its Origins and Applications in Modern Sociology. *Annual Review of Sociology*, *24*(1), 1–24. doi:10.1146/annurev.soc.24.1.1

Powell, W. W., & Grodal, S. (2005). Networks of innovators. In J. Fagerberg, D. C. Mowery, & R. R. Nelson (Eds.), *The Oxford handbook of innovation* (pp. 56–85). Oxford, New York: Oxford University Press.

Prahalad, C. K., & Hart, S. L. (2002). The Fortune at the Bottom of the Pyramid. *Strategy + Business*, *26*(First Quarter), 2–14.

Putnam, H. (2004a). *The collapse of the fact/value dichotomy and other essays* (3rd print). Cambridge, MA, [etc.]: Harvard University Press.

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.

Putnam, R. D. (2004b). Commentary: 'Health by association': some comments. *International Journal of Epidemiology*, 1–4.

Putnam, R. D., & Goss, K. A. (2002). Introduction. In R. D. Putnam (Ed.), *Democracies in flux. The evolution of social capital in contemporary society* (pp. 1–19). Oxford, New York: Oxford University Press.

Putnam, R. D., Leonardi, R., & Nanetti, R. Y. (1993). *Making democracy work: Civic traditions in modern Italy*. Princeton, N.J.: Princeton University Press.

Raiser, M., Rousso, A., Steves, F., & Teksoz, U. (2007). Trust in Transition: Cross-Country and Firm Evidence. *Journal of Law, Economics, and Organization*, *24*(2), 407–433. doi:10.1093/jleo/ewm060

Rangan, V. K., Chase, L., & Karim, S. (2015). The truth about CSR. *Harvard Business Review*, (January-February).

Rawls, J. (2005, c1971). *A theory of justice*. Cambridge, Mass.: Belknap Press.

Renneboog, L., Ter Horst, J., & Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, *32*(9), 1723–1742. doi:10.1016/j.jbankfin.2007.12.039

Richmond, B. J., Mook, L., & Quarter, J. (2003). Social Accounting for Nonprofits: Two models. *Nonprofit Management and Leadership*, *13*(4), 308–324.

Ritchie, J., & Spencer, L. (2002). Qualitative Data Analysis for Applied Policy Research. In A. M. Huberman & M. B. Miles (Eds.), *The qualitative researcher's companion* (pp. 305–331). Thousand Oaks, Calif., London: Sage.

Rivlin, A. M. (1971). *Systematic thinking for social action. H. Rowan Gaither lectures in systems science*. Washington: Brookings Institution.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, *126*(1), 3–25. doi:10.1037//0033-2909.126.1.3

Robinson, V. M. J. (2001). Descriptive and normative research on organizational learning: Locating the contribution of Argyris and Schön. *International Journal of Educational Management*, *15*(2), 58–67. doi:10.1108/EUM0000000005395

Rogers, A., Vassilev, I., Sanders, C., Kirk, S., Chew-Graham, C., Kennedy, A., … (2011). Social networks, work and network-based resources for the management of long-term conditions: a framework and study protocol for developing self-care support. *Implementation Science*, *6*(1), 56. doi:10.1186/1748-5908-6-56

Rosenberg, M. (1956). Misanthropy and Political Ideology. *American Sociological Review*, *21*(6), 690–695. doi:10.2307/2088419

Rosenberg, M. (1957). Misanthropy and attitudes toward international affairs. *Journal of Conflict Resolution*, *1*(4), 340–345. doi:10.1177/002200275700100403

Rosenberg, M. (1989). *Society and the adolescent self-image* (1st ed.). *Wesleyan paperback*. Middletown, Conn: Wesleyan University Press.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2009). *Evaluation: A systematic approach* (7. ed.). Thousand Oaks, Calif.: Sage Publ.

Rotheroe, N., & Richards, A. (2007). Social Return on Investment and social enterprise: Transparent accountability for sustainable development. *Social Enterprise Journal*, *3*(1), 31–48.

Rothstein, B., & Stolle, D. (2008). The State and Social Capital: An Institutional Theory of Generalized Trust. *Comparative Politics*, *40*(4), 441–459. doi:10.2307/20434095

Rubin, D. B. (2001). Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation. *Health Services and Outcomes Research Methodology*, *2*(3/4), 169–188. doi:10.1023/A:1020363010465

Rubin, J. D. (1995). A model of Intertemporal Emission Trading, Banking, and Borrowing. *Journal of Environmental Economics and Management*, *31*, 269–286.

Ruff, K. & Olsen, S. (2016). *The Next Frontier in Social Impact Measurement Isn't Measurement at All*. Retrieved from https://ssir.org/articles/entry/the_next_frontier_in_social_impact_measurement_isnt_measurement_at_all

Sabatier, P. A. (1998). The advocacy coalition framework: Revisions and relevance for Europe. *Journal of European Public Policy*, *5*(1), 98–130. doi:10.1080/13501768880000051

Sabatier, P. A., & Jenkins-Smith, H. C. (1993). *Policy change and learning: An advocacy coalition approach*. *Theoretical lenses on public policy*. Boulder, Colorado: Westview Press.

Sackett, D. L., Rosenberg, W., Gray, J., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ*, *312*(7023), 71–72. doi:10.1136/bmj.312.7023.71

Salamon, L. M., & Anheier, H. K. (1992). In search of the non-profit sector II: The problem of classification. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, *3*(3), 267–309.

Salter, W. E. G. (1969). *Productivity and Technical Change*. *Department of Applied Economics Monographs*: Cambridge University Press.

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. *Science*, *277*(15 August), 918–924.

Savedoff, W. D., Levine, R., & Birdsall, N. (2006). *When will we ever learn?: Improving lives through impact evaluation*. Washington D.C.

Sawhill, J. C., & Williamson, D. (2001). Mission Impossible?: Measuring Success in Nonprofit Organizations. *Nonprofit Management and Leadership*, *11*(3), 371–386. doi:10.1002/nml.11309

Schick, A. (2012). Faustlos–ein Gewaltpräventionsprogramm bewegt: Effektive und nachhaltige Gewaltprävention vom Kindergarten bis zur Sekundarstufe. *forum kriminalprävention*, (2), 6–11.

Schick, A., & Cierpka, M. (2003). Faustlos: Evaluation eines Curriculums zur Förderung sozial-emotionaler Kompetenzen und zur Gewaltprävention in der Grundschule. *Kindheit und Entwicklung*, *12*(2), 100–110.

Schick, A., & Cierpka, M. (2005). Faustlos: Evaluation of a curriculum to prevent violence in elementary schools. *Applied and Preventive Psychology*, *11*(3), 157–165. doi:10.1016/j.appsy.2005.05.001

Schick, A., & Cierpka, M. (2006). Evaluation des Faustlos-Curriculums für den Kindergarten. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, *55*(6), 459–474.

Schick, A., & Cierpka, M. (2013). International evaluation studies of Second Step, a primary prevention programme: a review. *Emotional and Behavioural Difficulties*, *18*(3), 241–247. doi:10.1080/13632752.2013.819250

Schmid, G., Schütz, H., & Speckesser, S. (1999). Broadening the scope of benchmarking: Radar charts and employment systems. *Labour*, *13*(4), 879–899.

Schmidt, I., Meurer, M., Saling, P., Kicherer, A., Reuter, W., & Gensch, C.-O. (2004). SEEbalance: Managing Sustainability of Products and Processes with the Socio-Eco-Efficiency Analysis by BASF. *Greener Management International*, *45*(Spring), 80–94.

Schober, C., & Then, V. (2015). *Praxishandbuch Social Return on Investment: Wirkung sozialer Investitionen messen*. Stuttgart: Schäffer-Poeschel.

Schorr, L. B. (2016). Reconsidering Evidence: What It Means and How We Use It. *Stanford Social Innovation Review*. Retrieved from https://ssir.org/articles/entry/reconsidering_evidence_what_it_means_and_how_we_use_it

Schreyögg, G., & Sydow, J. (2011). Organizational Path Dependence: A Process View. *Organization Studies*, *32*(3), 321–335. doi:10.1177/0170840610397481

Schumacher, E. F. (1973). *Small is beautiful: A study of economics as if people mattered*: Blond&Briggs.

Schumpeter, J. A. (1954). *History of economic analysis*. London: Allen & Unwin.

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35–37). Windsor, UK: NFER-NELSON.

Scott, S., Knapp, M., Henderson, J., & Maughan, B. (2001). Financial cost of social exclusion: follow up study of antisocial children into adulthood. *BMJ*, *323*(7306), 191. doi:10.1136/bmj.323.7306.191

Second Step. (2017). *Middle-school-curriculum*. Retrieved from http://www.secondstep.org/middle-school-curriculum

Sen, A. (1970). *Collective choice and social welfare*. *Mathematical economics texts: Vol. 5*. San Francisco, Calif: Holden-Day.

Sen, A. (1993a). Christopher Bliss: Life-Style and the Standard of Living. In M. Nussbaum & A. Sen (Eds.), *The Quality of Life* (pp. 437–444). Oxford University Press.

Sen, A. (2005). Human Rights and Capabilities. *Journal of Human Development*, *6*(2), 151–166. doi:10.1080/14649880500120491

Sen, A. K. (1985). *Commodities and capabilities*. *Oxford India paperbacks*. Delhi, New York: Oxford University Press.

Sen, A. K. (1993b). Capability and Well-Being. In M. Nussbaum & A. Sen (Eds.), *The Quality of Life* (pp. 30–53). Oxford University Press.

Sen, A. K. (2001). *Development as freedom*. *Oxford University Press paperback*. Oxford: Univ. Press.

Slaby, R. G., & Guerra, N. G. (1988). Cognitive mediators of aggression in adolescent offenders: I. Assessment. *Developmental Psychology*, *24*(4), 580–588. doi:10.1037/0012-1649.24.4.580

Small, M. L., Harding, D. J., & Lamont, M. (2010). Introduction: Reconsidering Culture and Poverty. In M. Lamont, M. L. Small, & D. J. Harding (Eds.), *Reconsidering culture and poverty* (pp. 6–29). Thousand Oaks, Calif.: Sage.

Smith, J. A. (1995). Semi-structured interviewing and qualitative analysis. In J. A. Smith, R. Harre, & L. van Langenhove (Eds.), *Rethinking methods in psychology* (pp. 9–26). London: SAGE Publications Limited.

Smith, P. K., Madsen, K. C., & Moody, J. C. (1999). What causes the age decline in reports of being bullied at school?: Towards a developmental analysis of risks of being bullied. *Educational Research*, *41*(3), 267–285. doi:10.1080/0013188990410303

Smith, S. R. (2010). Nonprofits and Public Administration: Reconciling Performance Management and Citizen Engagement. *The American Review of Public Administration*, *40*(2), 129–152. doi:10.1177/0275074009358452

Social Impact Investment Taskforce. (2014). *Measuring Impact*. Subject Paper of the Impact Measurement Working Group. Retrieved from http://www.socialimpactinvestment.org/reports/Measuring%20Impact%20WG%20paper%20FINAL.pdf

Social Value UK. (2017). *Global Value Exchange*. Retrieved from http://www.socialvalueuk.org/resources/global-value-exchange/

Somers, A., Nicholls, A., & Nicholls, J. (forthcoming). *Measuring social impact: Theories and methods*: SAGE Publications Ltd.

Sørensen, E., & Torfing, J. (2003). Network Politics, Political Capital, and Democracy. *International Journal of Public Administration*, *26*(6), 609–634. doi:10.1081/PAD-120019238

Spicer, M. W. (2008). The History of Ideas and Normative Research in Public Administration: Some Personal Reflections. *Administrative Theory & Praxis*, *30*(1), 50–70.

Spiel, C., Wagner, P., & Strohmeier, D. (2012). Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy. *International Journal of Conflict and Violence*, *6*(2), 177–186.

Stanca, L. (2009). Measuring indirect reciprocity: Whose back do we scratch? *Journal of Economic Psychology*, *30*(2), 190–202. doi:10.1016/j.joep.2008.07.010

Stanley, J. W., & Weare, C. (2016). The Effects of Internet Use on Political Participation. *Administration & Society*, *36*(5), 503–527. doi:10.1177/0095399704268503

Stehr, N., Henning, C., & Weiler, B. (2010). *The moralization of the markets*. New Brunswick: Transaction Publishers.

Stiglitz, J., Sen, A. K., & Fitoussi, J.-P. (2009). *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Retrieved from http://www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf

Stolzenberg, R. M., & Relles, D. A. (1990). Theory Testing in a World of Constrained Research Design: The Significance of Heckman's Censored Sampling Bias Correction for Nonexperimental Research. *Sociological Methods & Research*, *18*(4), 395–415. doi:10.1177/0049124190018004001

Stone, W. (2001). *Measuring Social Capital: Towards a theoretically informed measruement framework for researching social capital in family and community life*. Research Paper No. 24; Australian Institute of Family Studies.

Strauss, A. L., & Corbin, J. M. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory: Techniques and Procedures for Developing Grounded Theory*: Sage Publications Ltd (Digital).

Stuart, E. A., Lee, B. K., & Leacy, F. P. (2013). Prognostic score-based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *Journal of Clinical Epidemiology*, *66*(8 Suppl), S84-S90.e1. doi:10.1016/j.jclinepi.2013.01.013

Stuhlman, M. W., & Pianta, R. C. (2002). Teachers' Narratives About Their Relationships With Children: Associations With Behavior in Classrooms. *School Psychology Review*, *31*(2), 148–163.

Suldo, S. M., Savage, J. A., & Mercer, S. H. (2014). Increasing Middle School Students' Life Satisfaction: Efficacy of a Positive Psychology Group Intervention. *Journal of Happiness Studies*, *15*(1), 19–42. doi:10.1007/s10902-013-9414-2

Sullivan, A. (2001). Cultural Capital and Educational Attainment. *Sociology*, *35*(4), 893–912. doi:10.1177/0038038501035004006

Sund, P., & Lysgaard, J. (2013). Reclaim "Education" in Environmental and Sustainability Education Research. *Sustainability*, *5*(4), 1598–1616. doi:10.3390/su5041598

Swartz, D. L. (2003). Pierre Bourdieu's political sociology and governance perspectives. In H. P. Bang (Ed.), *Governance as social and political communication* (pp. 140–160). Manchester: University Press.

Tay, L., Chan, D., & Diener, E. (2014). The Metrics of Societal Happiness. *Social Indicators Research*, *117*(2), 577–600. doi:10.1007/s11205-013-0356-1

Thacher, D. (2006). The Normative Case Study. *American Journal of Sociology*, *111*(6), 1631–1676. doi:10.1086/499913

Then, V., Münscher, R., Callegari, B., Stahlschmidt, S., Wedemeyer, L., Bungard, P., & Strifler, A. (2012). *Creating Impact in Southern Norway: A Social Return on Investment Report to the Competence Development Fund of Southern Norway*. CSI Report. Retrieved from https://www.csi.uni-heidelberg.de/downloads/Creating_Impact_in_Southern_Norway.pdf

Thomas, J. B., Clark, S. M., & Gioia, D. A. (1993). Strategic sensemaking and organzational performance: Linkages among scallning, interpretation, action and outcomes. *Academy of Management Journal*, *36*(2), 239–270. doi:10.2307/256522

Titmuss, R. M., Oakley, A., & Ashton, J. (1970). *The gift relationship: From human blood to social policy*. New York: New Press.

Tomlinson, J., Baird, M., Berg, P., & Cooper, R. (2017). Flexible careers across the life course: Advancing theory, research and practice. *Human Relations*, *71*(1), 4–22. doi:10.1177/0018726717733313

Tsai, W., & Ghoshal, S. (1998). Social Capital and Value Creation: The Role of Intrafirm Networks. *The Academy of Management Journal*, *41*(4), 464–476.

Ttofi, M. M., Farrington, D. P., & Lösel, F. (2012). School bullying as a predictor of violence later in life: A systematic review and meta-analysis of prospective longitudinal studies. *Aggression and Violent Behavior*, *17*(5), 405–418. doi:10.1016/j.avb.2012.05.002

Tuan, M. T. (2008). *Measuring and/or Estimating Social Value Creation: Insights into eight integrated cost approaches*. Retrieved from http://www.gatesfoundation.org/learning/documents/wwl-report-measuring-estimating-social-value-creation.pdf

Turner, S. (2007). The Continued Relevance of Weber's Philosophy of Social Science. *Max Weber Studies*, *7*(1), 37–62.

United Nations. (2012). *Millenium Development Goals*. Retrieved from http://www.un.org/millenniumgoals/

United Nations. (2015). *Sustainable development goals*. Retrieved from http://www.un.org/sustainabledevelopment/sustainable-development-goals/

van den Hove, S. (2007). A rationale for science–policy interfaces. *Futures*, *39*(7), 807–826. doi:10.1016/j.futures.2006.12.004

van der Gaag, M., & Snijders, T. (2004). Proposals for the measurement of individual social capital. In H. D. Flap & B. Völker (Eds.), *Routledge advances in sociology: Vol. 9. Creation and returns of social capital. A new research program* (pp. 154–169). London, New York: Routledge.

van der Poel, M. G. M. (1993). Delineating personal support networks. *Social Networks*, *15*(1), 49–70. doi:10.1016/0378-8733(93)90021-C

van der Zee, H. T. M. (2002). *Measuring the value of information technology*. Hershey: Idea Group; Information Science.

van Deth, J. W., Montero, J. R., & Westholm, A. (Eds.). (2008). *Routledge research in comparative politics: Vol. 17. Citizenship and involvement in European democracies: A comparative analysis*. London: Routledge.

van Duijn, M. A. J., van Busschbach, J. T., & Snijders, T. A. B. (1999). Multilevel analysis of personal networks as dependent variables. *Social Networks*, *21*(2), 187–210. doi:10.1016/S0378-8733(99)00009-X

van Lange, P. A. M. (2015). Generalized Trust. *Current Directions in Psychological Science*, *24*(1), 71–76. doi:10.1177/0963721414552473

van Lier, P. A. C., Vuijk, P., & Crijnen, A. A. M. (2005). Understanding Mechanisms of Change in the Development of Antisocial Behavior: The Impact of a Universal Intervention. *Journal of Abnormal Child Psychology*, *33*(5), 521–535. doi:10.1007/s10802-005-6735-7

van Schoiack-Edstrom, L., Frey, K. S., & Beland, K. (2002). Changing Adolescents' Attitudes About Relational and Physical Aggression: An Early Evaluation of a School-Based Intervention. *School Psychology Review*, *31*(2), 201–216.

Vassilev, I., Rogers, A., Blickem, C., Brooks, H., Kapadia, D., Kennedy, A., … (2013). Social Networks, the 'Work' and Work Force of Chronic Illness Self-Management: A Survey Analysis of Personal Communities. *PLoS ONE*, *8*(4), e59723. doi:10.1371/journal.pone.0059723

Veenhoven, R. (2010). How Universal Is Happiness? In E. Diener, J. F. Helliwell, & D. Kahneman (Eds.), *Series in positive psychology. International differences in well-being* (pp. 328–350). Oxford, New York: Oxford University Press.

Verba, S., Nie, N. H., & Kim, J.-O. (1978). *Participation and political equality: A seven-nation comparison*. Cambridge: Cambridge University Press.

Vessels, G. G. (1998). *Character and community development: A school planning and teacher training handbook*. Westport, Conn: Praeger.

Vignoles, V. L., Regalia, C., Manzi, C., Golledge, J., & Scabini, E. (2006). Beyond self-esteem: Influence of multiple motives on identity construction. *Journal of Personality and Social Psychology*, *90*(2), 308–333.

Vissers, S., & Stolle, D. (2014). The Internet and new modes of political participation: Online versus offline participation. *Information, Communication & Society*, *17*(8), 937–955. doi:10.1080/1369118X.2013.867356

Waddock, S. A., & Post, E. J. (1991). Social Entrepreneurs and Catalytic Change. *Public Administration Review*, *41*(5), 393–401.

Weber, M. (1949). *The methodology of the social sciences*. Translated by Shils, Edward and Finch, Henry A. Glencoe [Ill.]: Free Press.

Weber, M. (1978). *Economy and society: An outline of interpretive sociology*. Berkeley: University of California Press.

Weber, M. (1988 (1920)). Die "Objektivität" sozialwissenschaftlicher und sozialpolitischer Erkenntnis. In J. Winckelmann (Ed.), *UTB für Wissenschaft. Uni-Taschenbücher: Vol. 1488. Gesammelte Aufsätze zur Wissenschaftslehre* (pp. 146–214). Tübingen: Mohr.

Weible, C. M. (2006). An Advocacy Coalition Framework Approach to Stakeholder Analysis: Understanding the Political Context of California Marine Protected Area Policy. *Journal of Public Administration Research and Theory*, *17*(1), 95–117.

Weible, C. M., Sabatier, P. A., & McQueen, K. (2009). Themes and Variations: Taking Stock of the Advocacy Coalition Framework. *The Policy Studies Journal*, *37*(1), 121–140.

Weick, K. E. (1991). The Nontraditional Quality of Organizational Learning. *Organization Science*, *2*(1), 116–124. doi:10.1287/orsc.2.1.116

Weinberger, D. A., & Schwartz, G. E. (1990). Distress and Restraint as Superordinate Dimensions of Self-Reported Adjustment: A Typological Perspective. *Journal of Personality*, *58*(2), 381–417. doi:10.1111/j.1467-6494.1990.tb00235.x

Weisbrod, B. A. (1975). Toward a theory of the voluntary nonprofit sector in a three-sector economy. In E. S. Phelps (Ed.), *Altruism, morality and economic theory*. New York: Russell Sage Foundation.

Weiss, C. H. (1972). *Evaluation research: Methods for assessing program effectiveness*. *Prentice-Hall methods of social science series*. Englewood Cliffs, N.J: Prentice-Hall.

Weiss, C. H. (1998). *Evaluation: Methods for studying programs and policies* (2nd ed.). Upper Saddle River, N.J: Prentice Hall.

Wellman, B. (1979). The Community Question: The Intimate Networks of East Yorkers. *American Journal of Sociology*, *84*(5), 1201–1231.

Wernet, C. A., Elman, C., & Pendleton, B. F. (2005). The Postmodern Individual: Structural Determinants of Attitudes. *Comparative Sociology*, *4*(3-4), 339–364.

Whittington, R. (2012). Big Strategy/Small Strategy. *Strategic Organization*, *10*(3), 263–268. doi:10.1177/1476127012452828

Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (2010). *Handbook of practical program evaluation* (3rd ed). *Essential texts for nonprofit and public leadership and management: Vol. 19*. San Francisco: Jossey-Bass.

Wieczorek, A. J., & Hekkert, M. P. (2012). Systemic instruments for systemic innovation problems: A framework for policy makers and innovation scholars. *Science and Public Policy*, *39*(6), 74–87.

Williamson, O. E. (1975). *Markets and hierarchies: analysis and antitrust implications: A study in the economics of internal organization*. New York, NY: The Free Press.

Williamson, O. E. (2000). The new institutional economics: Taking stock, looking ahead. *Journal of economic literature*, *38*(3), 595–613.

Wilson, S. J., & Lipsey, M. W. (2007). School-Based Interventions for Aggressive and Disruptive Behavior: Update of a Meta-Analysis. *American Journal of Preventive Medicine*, *33*(2), S130-S143. doi:10.1016/j.amepre.2007.04.011

Wolfs, C. A. G., Kessels, A., Dirksen, C. D., Severens, J. L., & Verhey, F. R. J. (2008). Integrated multidisciplinary diagnostic approach for dementia care: randomised controlled trial. *The British Journal of Psychiatry*, *192*(4), 300–305. doi:10.1192/bjp.bp.107.035204

Wood, D., Thornley, B., & Grace, K. (2013). Institutional impact investing: Practice and policy. *Journal of Sustainable Finance & Investment*, *3*(2), 75–94. doi:10.1080/20430795.2013.776256

Woodman, R. W. (1989). Organizational Change and Development: New Arenas for Inquiry and Action. *Journal of Management*, *15*(2), 205–228. doi:10.1177/014920638901500205

Wright, K., Hazelett, S., Jarjoura, D., & Allen, K. (2007). The AD-LIFE Trial. *Home Healthcare Nurse: The Journal for the Home Care and Hospice Professional*, *25*(5), 308–314. doi:10.1097/01.NHH.0000269964.34045.d5

Wu, C.-H., Tsai, Y.-M., & Chen, L. H. (2009). How do Positive Views Maintain Life Satisfaction? *Social Indicators Research*, *91*(2), 269–281. doi:10.1007/s11205-008-9282-z

Wysocki, A., Butler, M., Kane, R. L., & Shippee, T. (2013). *Future Research Needs: Long-Term Care for Older Adults* (Future Research Needs Paper No. 44). Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from www.effectivehealthcare.ahrq.gov/reports/final.cfm

Zahra, S. A., Gedajlovic, E., Neubaum, D. O., & Shulman, J. M. (2009). A typology of social entrepreneurs: Motives, search processes and ethical challenges. *Journal of Business Venturing*, *24*, 519–532.

Zajac, E. J., Kraatz, M. S., & Bresser, R. K. F. (2000). Modeling the Dynamics of Strategic Fit: A Normative Approach to Strategic Change. *Strategic Management Journal*, *21*(4), 429–453.

Zappala, G. & Lyons, M. (2009). *Recent appraoches to measuring social impact in the Third Sector: An overview*. Centre for Social Impact Background Paper No. 5.

Zeller, R. A., & Carmines, E. G. (2009). *Measurement in the social sciences: The link between theory and data*. Cambridge: Cambridge University Press.

Zheng, W. (2010). A Social Capital Perspective of Innovation from Individuals to Nations: Where is Empirical Literature Directing Us? *International Journal of Management Reviews*, *12*(2), 151–183. doi:10.1111/j.1468-2370.2008.00247.x

Zimmerman, S., Anderson, W. L., Brode, S., Jonas, D., Lux, L., Beeber, A. S., … (2013). Systematic Review: Effective Characteristics of Nursing Homes and Other Residential Long-Term Care Settings for People with Dementia. *Journal of the American Geriatrics Society*, *61*(8), 1399–1409. doi:10.1111/jgs.12372

Zimmerman, S., Sloane, P. D., & Eckert, J. K. (2001). *Assisted living: Needs, practices, and policies in residential care for the elderly*. Baltimore: Johns Hopkins University Press.

11 Appendix

Appendix 1 Scales community oriented housing

Table 27 Scales effects CO (part 1)

#	Cat.	Variable	Questions/scale	α	Scale categories
1	CC	Trust	(1) Most people can be trusted; (2) You can't be too careful when dealing with people; (3) Most people would take advantage of you, if they had the chance; (4) Most people would try to be fair; (5) Most people try to be helpful; (6) Most people only look after themselves.	0.85	1-5 ("Disagree strongly"-"Agree strongly")
2		Trust # NB	How many neighbours do you fully trust?	-	#
3	PC	Performing tasks	How much time do you invest in taking over tasks for the community of neighbours?	-	h
4		Problem solving	When there is a problem, neighbours solve it jointly.	-	1-5 ("Disagree strongly"-"Agree strongly")
5		Inhabitants' influence	The inhabitants have a big influence on what the facility is like and what happens in the facility.	-	As above
6		Self-efficacy	Generalised self-efficacy scale (10 questions)	0.82	As above
7	SC	NWS Fam&Fr	Please list all family members and friends you are in touch with regularly. How often are you in touch? Do you meet in person or talk on the phone, or both? How important is that relationship to you? How intimate/familiar is the relation?	-	NWS formula, see text
8		NWS NB	As above (on neighbours)	-	As above
9		# NB friends	How many neighbours do you consider to be "close contact" and/or "friends"?	-	#
10		Sense of belonging	(1) I feel as if I belong to the community of neighbours; (2) I am not sure whether I fit into the community of neighbours; (3) I have the feeling that neighbours hold me in high regard; (4) I don't have much to do with my neighbours.	0.77	1-5 ("Disagree strongly"-"Agree strongly")
11		Support NB instrumental	How much support do you receive in your daily life by your neighbours?	-	1-5 ("No support at all"-"Very high support")
12		Support NB compan.	How much do you interact and spend time with your neighbours?	-	As above
13		Support NB emotional	How much do you turn to neighbours to discuss to personal problems and worries?	-	As above

Remarks: NB=neighbour; NWS=network strength; compan.=companionship; CC=Cul. Cap; PC=Pol. Cap.; SC=Soc. Cap.

Table 28 Scales effects CO (part 2)

#	Cat.	Variable	Questions/scale	α	Scale categories
14		# NB inst	How many neighbours can you turn to for support?	-	#
15		# NB compan	As above	-	As above
16		# NB emot	As above	-	As above
17		Individual support h	How much time do you spend on supporting others?	-	h
18		Participation in events	How much time do you spend on participating in events?	-	As above
19	Sec.	Satisfaction	How satisfied are you with: (1) The facility's staff; (2) The opportunity to interact with your neighbours; (3) The opportunity to interact with people from outside the facility; (4) The activities of the facility that you can take part in; (5) The overall support that you are receiving in the facility; (6) Your living situation as a whole.	0.83	1-5 ("Disagree strongly"-"Agree strongly")
20		Life satisfaction	All in all, how satisfied are you with your life?	-	0-10 ("Not at all"-"Very satisfied")
21		HQoL	EQ-5D scale (5 dimensions; without EQ visual analogue scale VAS)	-	5 health states per dimension

Remarks: NB=neighbour; NWS=network strength; compan.=companionship; CC=Cul. Cap; PC=Pol. Cap.; SC=Soc. Cap.

Sources for scales effects (revised before application): see footnote.[71]

Table 29 Scales matching variables CO-AL

#	Cat.	Variable	Questions/scale	α	Scale categories
17	SC	Activity preferences PRE	(1) I like taking an active role when interacting with others (lead the conversation, entertain others, etc.) ; (2) I prefer to observe what is going on without taking action; (3) I am an active person and enjoy spending time with others (meeting others, participating in events, etc.); (4) I need a lot of time for myself and prefer being on my own.	0.75	1-5 ("Disagree strongly"-"Agree strongly")
22		Social contacts PRE	(1) Think about the number of social contacts that you typically had in your life: How would you categorise their number?; (2) How would you categorise their intensity (frequency and length of spending time); (3) How intimate/personal were they?; (4) How important were they to you?	0.76	1-5 ("Very low"-"Very high")

[71] 1: Rosenberg (1957); European Social Survey (2014b); 6: Schwarzer and Jerusalem (1995); 10: own.

Appendix 2 Regression results community oriented housing

See next pages in landscape format.

Table 30 Regression results CO (part 1)

Variable	**Trust LOG**			**Trust # people**			**Trust # NB**			**Own engagement h**			**Self-efficacy**			**NWS Fam&Friends**		
Regression / Year	Lin.	15-16		NEB	15-16		NEB	2016		NEB	15-16		Lin.	15-16		Lin.	15-16	
Controls	B	SE	Sig.	B	SE	Sig.	B	SE	Sig.	B	SE	Sig.	B	SE	Sig.	B	SE	Sig.
(Constant term)	1.188	0.0364	0.000	0.990	0.1939	0.000	-3.545	0.8874	0.000	-0.213	0.7500	0.776	3.031	0.2589	0.000	-0.110	0.0737	0.136
Year	-0.046	0.0115	0.000	0.188	0.0674	0.005				-0.128	0.1082	0.238	-0.111	0.0458	0.015	0.000	0.0016	0.767
[Model=CO]	0.063	0.0159	0.000	0.357	0.0937	0.000	0.620	0.2086	0.003	0.448	0.2111	0.034	0.064	0.0681	0.346	-0.074	0.0269	0.006
Gender																		
Age										-0.309	0.1407	0.028						
Members Household							0.731	0.3314	0.028				-0.235	0.1003	0.019			
Years of Residency																		
Income	0.021	0.0048	0.000															
Kids & Siblings (#)													0.034	0.0122	0.005	0.019	0.0057	0.001
Family Meeting Frequency	-0.006	0.0022	0.003													0.008	0.0035	0.024
Upbringing by single parent													-0.169	0.0911	0.064			
Region	-0.043	0.0150	0.004										-0.239	0.0691	0.001	0.073	0.0292	0.012
Locality	0.040	0.0154	0.010							-0.671	0.2141	0.002						
Dist. prev. residence High				0.276	0.1067	0.010												
Health-related QoL 2015																		
Disability High																		
Care Level 0/1/2													-0.228	0.0763	0.003			
Chronical Conditions High										-0.469	0.2030	0.021						
Activity Preference PRE	0.022	0.0083	0.009	0.146	0.0465	0.002	0.468	0.1006	0.000				0.140	0.0380	0.000			
Trust PRE							0.454	0.1910	0.017									
Self-efficacy PRE													0.182	0.0534	0.001			
Help individual PRE										0.441	0.1317	0.001						
Engagement community PRE																		
Social network PRE																0.104	0.0195	0.000
QIC reduced vs full model	33.61	76.15		347.95	361.17		336.09	356.99		788.44	825.12		207.04	246.82		48.59	114.78	

Remarks: LOG=logarithm; Lin.=linear; NEB=negative binomial; Tw.=Tweedie with log link, p=1.5; SE=standard error; NWS=Network strength; NB=neighbour; Inst. Sup.=institutional support; Emot. sup.=emotional support; Compan.=companionship; HQoL=health related quality of life; Dist. prev.=distance previous; "Education high" excluded in controls, since never sig.

Table 31 Regression results CO (part 2)

Variable	**NWS NB**			**# NB=friends**			**Cohesion in facility**			**Inst. sup. # NB**			**Compan. # NB**			**Emot. sup. # NB**		
Regression / Year	Tw.	2016		NEB	2016		Lin.	15-16		NEB	2016		NEB	2016		NEB	2016	
Controls	B	SE	Sig.	B	SE	Sig.	B	SE	Sig.	B	SE	Sig.	B	SE	Sig.	B	SE	Sig.
(Constant term)	-2.941	0.4232	0.000	-0.712	0.3920	0.069	3.190	0.1776	0.000	-0.549	0.6002	0.360	-0.809	0.6079	0.183	-1.435	0.8325	0.085
Year							0.054	0.0493	0.275									
[Model=CO]	0.364	0.1263	0.004	0.607	0.1381	0.000	0.229	0.0911	0.012	0.652	0.1723	0.000	0.609	0.1337	0.000	0.503	0.2468	0.042
Gender	-0.376	0.1754	0.032													-1.096	0.3325	0.001
Age	-0.242	0.0722	0.001															
Members Household													0.570	0.1623	0.000			
Years of Residency													-0.057	0.0146	0.000	-0.048	0.0242	0.048
Income													0.185	0.0526	0.000	0.365	0.1060	0.001
Kids & Siblings (#)																		
Family Meeting Frequency																		
Upbringing by single parent	-0.401	0.1375	0.004										-0.413	0.1688	0.014	-1.048	0.2381	0.000
Region										-0.688	0.1755	0.000	-0.352	0.1524	0.021	-0.558	0.2831	0.049
Locality																		
Dist. prev. residence High				0.357	0.1573	0.023										-0.740	0.2545	0.004
Health-related QoL 2015	0.714	0.2528	0.005							0.983	0.3886	0.011						
Disability High													-0.428	0.1501	0.004			
Care Level 0/1/2													0.383	0.1245	0.002	-0.751	0.1997	0.000
Chronical Conditions High																		
Activity Preference PRE	0.394	0.0705	0.000	0.232	0.0819	0.005	0.191	0.0461	0.000				0.256	0.0730	0.000			
Trust PRE																		
Self-efficacy PRE																		
Help individual PRE				0.237	0.0885	0.007				0.144	0.1017	0.157	0.347	0.0977	0.000			
Engagement community PRE	-0.091	0.0527	0.085	-0.108	0.0630	0.087							-0.169	0.0636	0.008	-0.112	0.0732	0.127
Social network PRE	0.185	0.0850	0.030													0.745	0.1666	0.000
QIC reduced vs full model	130.05	151.99		262.58	279.23		300.75	355.30		256.09	272.12		267.04	280.51		214.35	231.86	

Remarks: LOG=logarithm; Lin.=linear; NEB=negative binomial; Tw.=Tweedie with log link, p=1.5; SE=standard error; NWS=Network strength; NB=neighbour; Inst. Sup.=institutional support; Emot. sup.=emotional support; Compan.=companionship; HQoL=health related quality of life; Dist. prev.=distance previous; "Education high" excluded in controls, since never sig.

Table 32 Regression results CO (part 3)

Variable	**Own support h**			**Part. events h**			**Overall satis. LOG**			**HQoL**		
Regression / Year	NEB	2016		NEB	15-16		Lin.	15-16		Lin.	2016	
Controls	B	SE	Sig.	B	SE	Sig.	B	SE	Sig.	B	SE	Sig.
(Constant term)	-1.444	0.3595	0.000	2.545	0.3537	0.000	1.225	0.0455	0.000	0.298	0.0642	0.000
Year				-0.156	0.0882	0.077	-0.007	0.0096	0.485			
[Model=CO]	0.413	0.2550	0.105	0.387	0.1181	0.001	0.041	0.0157	0.008	-0.003	0.0245	0.908
Gender				-0.286	0.1465	0.051						
Age												
Members Household												
Years of Residency				-0.041	0.0102	0.000						
Income							0.021	0.0055	0.000			
Kids & Siblings (#)												
Family Meeting Frequency												
Upbringing by single parent							-0.043	0.0224	0.056			
Region	0.923	0.2684	0.001	0.472	0.1178	0.000						
Locality	-0.536	0.2628	0.042									
Dist. prev. residence High												
Health-related QoL 2015										0.654	0.0717	0.000
Disability High										-0.119	0.0329	0.000
Care Level 0/1/2										-0.098	0.0282	0.001
Chronical Conditions High												
Activity Preference PRE												
Trust PRE												
Self-efficacy PRE				-0.195	0.0686	0.005						
Help individual PRE							0.019	0.0094	0.046			
Engagement community PRE	0.315	0.0872	0.000									
Social network PRE												
QIC reduced vs full model	305.71	323.88		573.76	578.35		28.68	94.06		23.64	65.13	

Remarks: LOG=logarithm; Lin.=linear; NEB=negative binomial; Tw.=Tweedie with log link, p=1.5; SE=standard error; NWS=Network strength; NB=neighbour; Inst. Sup.=institutional support; Emot. sup.=emotional support; Compan.=companionship; HQoL=health related quality of life; Dist. prev.=distance previous; "Education high" excluded in controls, since never sig.

Appendix 3 Scales and interview guide school-based intervention

Table 33 Scales quantitative scoping RA (part 1)

#	Cat.	Variable	Questions/scale	α	Scale categories
1	CC	Endorsement of aggression factors (P=physical; V=verbal)	(1) Sometimes you have only two choices—getting hit or hitting the other youths first. (P) (2) Sometimes you have to fight other youths to get respect. (P) (3) When a youth is picking on another, it's not right for you to join. (V) (4) It's OK to hit someone if they do something mean to you. (P) (5) There are always other ways to solve an argument besides insulting a youth or getting put down yourself. (V) (6) It's best to avoid even listening to gossip or rumours. (V) (7) There are only two kinds of youths—the youths who fight and the youths who get beaten up. (P) (8) When a friend of yours is in a fight, it's alright to cheer for them. (P) (9) If you hear something bad about someone it's OK to pass it on. (V)	0.8	1-4 ("Disagree strongly"-"Agree strongly")
2		Aggression (P=physical; V=verbal)	(1) I said things about other kids to make other students laugh. (V) (2) I pushed or shoved other students. (P) (3) I called other students bad names. (V) (4) I threatened to hurt or to hit someone. (P) (5) I encouraged other students to fight. (P)	0.7	# of times per last 7 days
3		Cooperative behaviour	(1) I helped someone stay out of a fight. (2) I cooperated with others. (3) I told other students how I felt when they upset me. (4) I helped other students solve a problem or dispute. (5) I channelled situations in which I was angry into something constructive. (6) I tried to see another person's point of view.	0.73	1-4 ("Disagree strongly"-"Agree strongly")

Remarks: CC=Cul. Cap; PC=Pol. Cap.; SC=Soc. Cap.

Sources for scales (revised before application): see footnote.[72]

[72] 1: van Schoiack-Edstrom, et al. (2002)(derived from Slaby and Guerra (1988); Erdley and Asher (1998); Crick and Bigbee (1998); Crick and Grotpeter (1995)); 2: Dahlberg, et al. (2005, p. 168) (in relation to Orpinas (1993, p. 74); Orpinas and Frankowski (2001, 64f.)); 3: Dahlberg, et al. (2005, p. 168) (in relation to Orpinas (1993, p. 74); Orpinas and Frankowski (2001, 64f.)).

Table 34 Scales quantitative scoping RA (part 2)

#	Cat.	Variable	Questions/scale	α	Scale categories
4		Consideration for others	(1) Doing things to help other people is more important to me than almost anything else. (2) I often go out of my way to do things for other people. (3) I think about other people's feelings, before I do something they might not like. (4) I enjoy doing things for other people, even if I don't receive anything in return. (5) I make sure doing what I want will not cause problems for other people. (6) Before I do something, I think about how it will affect the people around me. (7) I try very hard not to hurt other people's feelings.	0.81	1-5 ("Disagree strongly"-"Agree strongly")
5		Self-esteem	(1) I feel I am a person of worth, at least on an equal par with others. (2) I feel that I have a number of good qualities. (3) All in all, I am inclined to feel like a failure. (4) I am able to do things as well as most other people. (5) I feel I do not have much to be proud of. (6) I take a positive attitude towards myself. (7) On the whole, I am satisfied with myself.	0.82	As above
7/8	PC	Group cohesion/ Collective efficacy	(1) Students in this class are willing to help their fellow students. (2) This is a close-knit class. (3) If there is a problem in the class, students can get it solved together. (4) Students in this class generally don't get along with each other. (5) Students in this class do not share the same values. (6) Students from different social classes and races get along well.	0.85	As above
9	SC	Individual influence	I have almost no influence over what my class is like.	-	1-5 ("Disagree strongly"-"Agree strongly")

Remarks: CC=Cul. Cap; PC=Pol. Cap.; SC=Soc. Cap.

Sources for scales (revised before application): see footnote.[73]

[73] 4: Dahlberg, et al. (2005, 181f.) (in relation to Orpinas (1993); modified in Bosworth, et al. (1996, p. 73)); 5: Dahlberg, et al. (2005, 134ff.) (in relation to Weinberger and Schwartz (1990)); 7/8: Dahlberg, et al. (2005, p. 147) (in relation to Rosenberg (1989); Gray-Little, et al. (1997)); Dahlberg, et al. (2005, 324f., 46, 343) (in relation to Sampson, et al. (1997); Barnes, et al. (1997) and Vessels (1998) and Perkins, et al. (1990)).

Table 35 Semi-structured interview guide for pupils RA

#	Subject area	Prompts	Components
1	Expectations, satisfaction	(1) Can you recall your participation well (think of who conducted the programme, what you did, how you felt etc.)? (2) Have you been looking forward to participating in the programme? (3) Has it been clear to you what the programme was about or were you surprised when participating? (4) Have you felt comfortable during the programme? (5) Overall, how have you liked participating in the programme?	
2	Impressions, assessment of key pillars of programme	(1) Was there anything specific that you liked or disliked about what you did? (2) Which of the components of the programme can you best recall? Which ones made an impression on you? And which ones did you find unimportant?	
3	Personal change	Do you feel your participation has changed anything for you? Did it change how you think or how you interact with others?	General prompt
		What would you do in the following situations? Do you think your reaction has changed as compared to before participating in the programme? (1) Someone has started shoving you around, just for fun. (2) Someone has insulted you, because they had a dif-ferent opinion on a topic than you and a dispute between the two of you has started. Etc.	Aggression (scenarios)
		Do you feel there has been a change in the following areas in the recent weeks? Could you give an example? (1) Have you been in conflict with someone of late (verbally or physically)? (2) Have you supported someone or shared your feelings with someone of late? (3) Have you been mindful of others of late (done something special for others, thought about how they felt etc.)? (4) How confident in yourself have you felt of late (judged by your ability, mood, general motivation or how you've seen yourself as compared to others)? (5) Whether and how you can have an influence in this class?	Aggressive bahaviour, supportive behaviour, caring for others, self-esteem, individual influence
4	Changes in the class room/group	Have you noticed any changes of how pupils interact with each other in these areas: (1) How pupils work together; (2) How they respect each other; (3) How they act (together) when a problem comes up that needs to be solved.	Joint problem solving and interaction
5	Concluding assessment	(1) In your opinion, have you learned anything during the programme? (2) Would you assess what you have learned as important and useful to you? And do you think that it will remain useful to you in the future? (3) Would you recommend participating in the programme to others?	

Appendix 4 Additional analysis political web-platform

Note: The descriptive analysis provides a more detailed picture of some of the characteristics of the use of AW as presented as a background for the presentation of results. The issues addressed span from specifics on the intensity and frequency of using AW, to the parties which users perceived as most accessible.

Length, intensity and purpose of using the platform

The mean length of using AW was 2.76 years for citizens and thus at about three years (SD=2.34; N=717). About 4 percent of respondents had only started using AW in 2016, while 2.5 percent had used it for ten or more years. Politicians had been using AW for almost 5.5 years (SD=3.192; N=237). About 8 percent of the politicians had only started using AW in 2016, whereas about 10 percent had been using it for 10 years or more.

As shown in Tables XW and XW almost 60 percent of citizens reported they used AW "once or several times per month", another third even "once or several times per week" (see Table 36).[74] For half of them the time spent on AW amounted to "less than half an hour per week", for another 37 percent to at least that but to no more than 1 hour per week (see Table 37). The use by politicians was considerably less intense and amounting to "less than half an hour per week" for 95 percent of MPs and a frequency of use of "once to several times per month" for 23 percent and of "less than once a month" for 76 percent.

[74] Here and in other tables: Percentage figures may not total to 100%, due to rounding.

Table 36 Frequency of use AW

	Citizens	%	Politicians	%
< 1 time per month	66	9%	177	76%
1 - several times per month	425	59%	54	23%
1 - several times per week	224	31%	1	0%
Total	715	100%	232	100%

Table 37 Intensity of use AW

	Citizens	%	Politicians	%
< 0.5 h per week	375	52%	215	95%
0.5-1 h per week	264	37%	10	4%
> 1 h per week	84	12%	2	1%
Total	723	100%	227	100%

Respondents were furthermore asked to indicate the main purpose for which they used AW (multiple answers possible). Table 38 shows that the large majority of citizens (96 percent) use the platform for "gathering information", followed by "posing questions to politicians" (60 percent). For politicians it is the other way around. They mainly use AW for communication purposes (87 percent) and only in the second instance for gathering information (37 percent). Commenting on political statements is of lower importance, yet still important for 43 percent of the citizens and 27 percent of the politicians. 28 percent of the citizens furthermore use the platform for discussing issues with other users.

Table 38 Main purpose of use AW

	MAX	Citizens	%(of MAX)	MAX	Politicians	%(of MAX)
Gathering information	745	716	96%	255	94	37%
Asking/ answering questions	745	444	60%	255	222	87%
Commenting on political statements	745	318	43%	255	68	27%
Discussing with other users	745	209	28%	N.A.		

Alternative modes of communication

The survey also served to gauge how the use of AW compared to other means of communication between citizens and politicians and how the use of alternative means had changed through the use of AW. Table 39 indicates whether the use of alternative means had "become higher" after using AW, "they have never been used at all", "remained stable", or "decreased." It shows that potential effects differ strongly between citizens and politicians.

Table 39 Alternative modes of communication AW

	Citizens					Politicians				
	Total	Higher	Never used	Stable	Lower	Total	Higher	Never used	Stable	Lower
Personal	667	35	200	313	119	239	15	4	214	6
	100%	*5%*	*30%*	*47%*	*18%*	*100%*	*6%*	*2%*	*90%*	*3%*
Letter	619	69	261	222	67	233	9	7	182	35
	100%	*11%*	*42%*	*36%*	*11%*	*100%*	*4%*	*3%*	*78%*	*15%*
E-Mail	676	47	119	190	320	234	56	4	170	4
	100%	*7%*	*18%*	*28%*	*47%*	*100%*	*24%*	*2%*	*73%*	*2%*
Website	515	37	100	175	203	167	31	6	125	5
	100%	*7%*	*19%*	*34%*	*39%*	*100%*	*19%*	*4%*	*75%*	*3%*

For citizens electronic modes of communication are most important. Letters and personal contact are comparatively less important, yet still matter to about 60-70 percent of CS respondents (only 30 and 42 percent in category "never used"). Shifts in communication also

primarily occur at the electronic level. After having started to use AW, many respondents reported a decrease in communication through other websites (decrease reported by 39 percent) and e-mail (decrease reported by 47 percent). With regard to correspondence or personal contact in contrast, the share of people for whom it has increased and those for whom it has decreased (almost) level each other out. This indicates that AW has become a preferred mode of communication for a substantial share of CS respondents. This is backed up by the fact that about 35 percent of respondents (260 responses, N=605) said that AW was their "main means of political communication" to a separate question.

For reasons of completeness citizens had been asked which alternative websites they used for communicating and interacting with politicians. These are the most frequent ones found in the text boxes of the survey: Campact, Avaaz, Change.org, politicians' own websites, Facebook, Bundestag.de. It is to be remarked that the first three are online spaces for petitions and that the German section of Change.org is led by one of the initiators of AW.

For politicians in turn none of the shifts detected for citizens have occurred. All but one politician reported that AW was not their main means of political communication. As can further be seen in Table 39, other means of political communication are equally important (very low percentages of "never used" for any of the four means). Also, the other means' use before and after starting to use AW is remarkably stable (no changes in the other means of communication for about 70-90 percent of politicians, depending on the means). There is some increase in website and e-mail communication (19 and 24 percent respectively). This is unlikely an effect of AW though, but probably marks a generally intensified use of electronic communication by politicians, provoked by contemporary communication habits or through positions being taken up by first-time elected MPs.

Thus, while AW is one of many communication channels for politicians, it gains importance after initial use for citizens relative to others, in particular relative to other electronic means of communication.

Frequency of communication

Respondents were also asked how their frequency of contacts to politicians during a year had changed after having started to use AW. Responses were recorded on one scale each for the situation "before" and "after", with categories ranging from 0-3 (where 0=never, 1=once per year, 2=several times per year, 3=at least once a month). The median shifted from 1 to 2 when comparing the two situations. The sign test was applied to test for the significance of the shift. This was done because: the comparison was made between two observations on the same sample (related sample); and the data were not normally distributed, and the applied scale was ordinal and not an interval one. The difference turned out to be highly significant ($z=-14.434$, $p=0.000$, $N=681$). It is particularly remarkable that the share of people, who had "never been in touch" with politicians in any way, neither personally nor by any other means of communication, decreased from about 39 in the situation "before" people started using AW to 17 percent in the situation "after." For comparison, of the people surveyed in ESS 2014 almost 85 percent (N=3039) reported they had never been in touch with a politician during the previous year.

The same questions were posed to politicians. However, no enhancing effect could be detected. This is not surprising, since communication with citizens is of course at the heart of any MP's tasks. Around 38 percent of the politicians reported that they answered citizens' questions "several times per week", another 36 percent even "daily" (N=220).

Policy levels and parties

As illustrated in Table 40, when asked about the policy level participants were interested in (multiple answers possible), citizens gave priority to the national level (96 percent), but the EU and the federal state levels followed suit at close distance (89 and 86 percent respectively). The gap to the local level was greater, but it was still of high significance to 70 percent of the respondents. Politicians, when asked about the main policy level they were active at, focussed on the federal state and the local level (85 and 61 percent). The focus on the national (45 percent) and the EU level (29 percent) was lower. This gives an indication that federal state MPs made up for the largest share within the group of PS respondents, national MPs the second largest and EU MPs the smallest.

Table 40 Main policy levels AW

	MAX	Citizens	%(of MAX)	MAX	Politicians	%(of MAX)
Local level	745	525	70%	255	155	61%
Fed. state level	745	642	86%	255	218	85%
National level	745	713	96%	255	115	45%
EU level	745	661	89%	255	74	29%

In order to furthermore find out about the primary addressees of citizens' communication, they have been asked to indicate with which parties they had been in touch with most often. Table 41 shows that only 16% of the users of AW had never been in touch with any politician at all and that SPD (Social Democratic Party) is leading the list of frequent contacts with 61 percent, followed at slightly lower levels by CDU (Christian Democratic Party) and Die Grünen (Green Party; 56 and 52 percent each). Then there is a marked and

gradual decrease from Die Linke (The Left Party; 37 percent) to AFD (Alternative für Deutschland; 9 percent).[75]

Table 41 Contact with parties AW

	None	SPD	CDU	Die Grünen	Die Linke	FDP	Piraten	AFD
Yes	120	456	420	385	275	155	104	66
%	16	61	56	52	37	21	14	9
Total	745	745	745	745	745	745	745	745

A further question was concerned with the accessibility of parties as perceived by citizens. It was rated on a scale from 0-10 ("not at all accessible"-"very accessible"). The results are reported by means of density plots that illustrate slopes of the distribution of answers. Parties have been split into two groups for the illustration due to strong differences in the number of people who chose to respond. While at least 300 respondents had judged the accessibility of group 1 formed by SPD (336 respondents), CDU (334), Die Grünen (340) und Die Linke (300), considerably fewer had done so for group 2, namely FDP (218), Piratenpartei (168), and AFD (127). Please note that the number of responses is low for all parties: less than half of the respondents to CS chose to answer these questions at all. The displayed results are therefore to be interpreted with caution.

[75] Other parties were sporadic in their occurrence and have therefore not been included in the analysis.

Figure 13 Parties' accessibility AW (frequent responses)

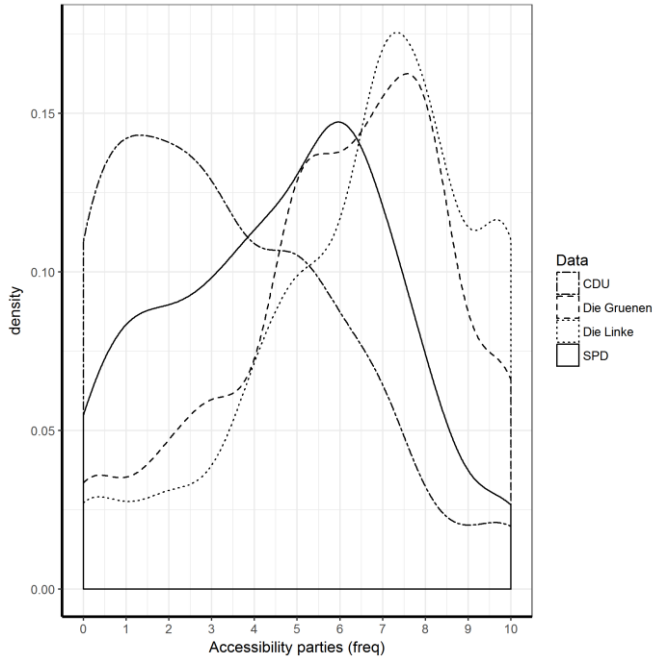
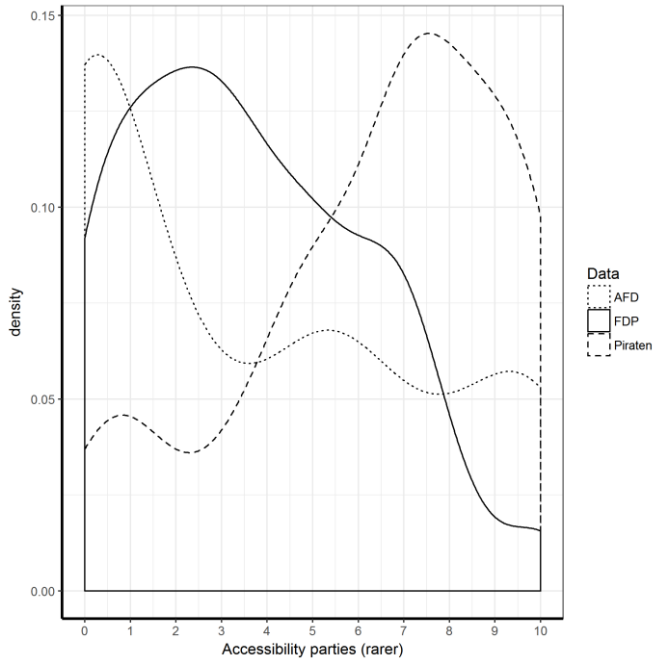


Figure 14 Parties' accessibility AW (rarer responses)



As can be seen from visual inspection of Figure 13 and Figure 14, Piratenpartei, Die Linke, and Die Grünen (in this order) are seen as most accessible, followed by SPD with a middle to high tendency in terms of accessibility, FDP and CDU with a lower tendency, and AFD at the lowest level.

Appendix 5 Scales political web-platform

Table 42 ESS questions AW

#	Cat.	Variable	Question	Scale ESS	Scale CS	Unified
1	PC	Voting	Did you vote in the last national election?	Binary	"Never"-"seldom"-"sometimes"-"mostly"-"always"	Binary ("always"=1)
2		Political orientation	Which party did you vote for in the last national election? (indirect party mandate, "Zweitstimme")	List of parties	List of "political orientations" (CS), list of parties (PS)	List of "political orientations"
3	CC	Interest	On an average weekday, how much time do you spend watching news or programmes about politics and current affairs?	"No time at all"-"over three h" in 0.5 h increments	Same as ESS	Unchanged
4		Tolerance (1)	Gay men and lesbians should be free to live their own life as they wish.	1-5 ("Agree strongly"-"Disagree strongly")	Same as ESS	Unchanged
5		Tolerance (2)	To which extent do you think the government should allow people of a different race and ethnic group from most people in Germany to come and live here?	1-4 ("Allow many"-"Allow some"-"Allow few"-"Allow none")	1-5 ("Agree strongly"-"Disagree strongly")	1-4 (Recoded, CS/PS medium category split evenly)
6		Solidarity	The governement should take measures to reduce differences in income levels.	1-5 ("Agree strongly"-"Disagree strongly")	Same as ESS	Unchanged
7		Community	Some say European unification should go further. Some say it has already gone too far. What number on the scale best describes your position?	0-10 ("Gone too far"-"Should go further")	1-5 ("Agree strongly"-"Disagree strongly")	1-5 (CS recoded, ESS stratified)
8		Trust (1)	Would you say that most people can be trusted, or that you can't be too careful in dealing with people?	0-10 ("Can't be too careful"-"Most can be trusted")	0-10 ("No trust at all"-"Complete trust")	0-10 ("Not at all"-"fully")
9		Trust (2)	How much do you personally trust the legal system?	0-10 ("No trust at all"-"Complete trust")	Same as ESS	Unchanged
10		Trust (3)	How much do you personally trust political parties?	As above	Same as ESS	Unchanged
11	Sat.	Satis. (1)	How satisfied are you with the way the government is doing ist job?	0-10 ("Extremely dissatisfied"-"Extremely satisfied")	Same as ESS	Unchanged
12		Satis. (2)	How satisfied are you with the way democracy works in Germany?	As above	Same as ESS	Unchanged
13		Satis. (3)	How satisfied are you with your life as a whole nowadays?	As above	Same as ESS	Unchanged

Remarks: PC=Pol. Cap.; CC=Cul. Cap; Sat.=Satisfaction.