The plausibility of policy

The Plausibility of Policy;
Case studies from the social domain

De plausibiliteit van beleid;
Case studies uit het sociale domein

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op

donderdag 20 november 2014 om 9:30 uur

Vasco Lub
geboren te Raamsdonksveer

ERASMUS UNIVERSITEIT ROTTERDAM

Promotiecommissie:

Promotor(en):   **Prof.dr. G.B.M. Engbersen**

Overige leden: **Prof.dr. J.P.L. Burgers**
    Prof.dr. R.J. Van der Veen
    Dr. R. Kleinhans

Preface

Nowadays, there is hardly a sphere of public or private life that is not affected by some sort of government action or agenda. Since my career as a social researcher began, I have been studying a wide variety of policy areas from a wide variety of angles. I feel content that I can now present this dissertation as an academic conversion of some of my professional efforts.

Throughout my years working for different institutes and commissioners, I have learned that social policy is an ambivalent phenomenon. On the one hand, the question 'Does it work?' sometimes receives little attention. On the other hand, there is an increasing interest and growing debate on the effectiveness of measures (and the appropriate methodology on how to determine that effectiveness). I feel fortunate to have been able to research this nascent and fascinating field of study in recent years.

In this book I present some alternative approaches to assess whether social policy measures are plausible. One of my main lines of argument is to confront policy assumptions with existing scientific data. The effectiveness of a single social intervention might be hard to ascertain (given its complexity and context), but that does not mean that the validity of underlying assumptions cannot be put to the test. Indirectly such an approach provides knowledge about the probability of success of a specific measure. Because when policy is based on false assumptions, it risks a lack of effectiveness. But when it is based on strong assumptions its success rate is more likely. The fields of sociology, psychology and many other social science disciplines have yielded valuable insights into areas that might affect the success rate of social policy measures, and we would be foolish to leave this vast reservoir of scientific knowledge untouched.

In this thesis I also discuss the need for studying policy from multiple order perspectives to avoid blind spots (first, second and third). These specific three order perspectives also reflect my own personal growth as a researcher. As my career progressed, I have advanced from studying instrumental perspectives on policy to methodological and (critical) reflexive perspectives. My professional growth – and thus this dissertation - would not have been possible, however, without the support and inspiration of others. Special thanks go out to Godfried Engbersen, my doctoral advisor. It is such a pleasure to work with you. Your intellectual inspiration and your never failing positive attitude and humour have helped me more than once to overcome certain obstacles. Then there are people in my professional circle (past and present) who have contributed to this dissertation. Many thanks go out to Matthijs Uyterlinde, Peter Rensen, Erik Snel and Radboud Engbersen. I have learned a lot from you.

Furthermore, the analyses presented in this book would not have been possible without commissioned research projects by Movisie, the Netherlands Centre for Social Development, and funding of the Dutch Ministry of Social Care and Welfare. I owe these institutes gratitude. Also a word of gratitude to Boom Lemma / Eleven International Publishing (Joris Bekkers) for publishing my work.

My parents, despite their efforts to discourage me from becoming a sociologist ('Can you make a living like that?'), have always been a source of guidance and inspiration for me. My father Willem, from who I inherited my interest in science and thirst for knowledge, and my mother Adrienne who always believed in my abilities ever since she convinced me not to drop math in high school. I could not have done this without you. Finally, Larae, your trust and faith in me is endless. I love you and hope in the future I have many more papers for you to revise.

Vasco Lub
Rotterdam, June 2014

Contents

1   Introduction

1.1.   Introduction

In 2011, the Dutch Court of Audit issued a report that showed that the Dutch government lacked evidence as to whether 6 billion euros in subsidies had been spent effectively[1]. The report showed that in the period 2005-2010, only 81 of the 633 government subsidies were subjected to any form of evaluation (which is about 10%). Yet, as the court concluded, only a few of these evaluations was conducted in a 'sound' way. Most of the evaluations did not make clear whether the grants were spent as intended, if the objective of the subsidy was even achieved, and, more importantly, whether a positive result could be attributed to a specific measure or program.

One would expect that in times of fiscal crisis and budget cuts, the Dutch finance minister (at the time Jan Kees de Jager) would draw critical lessons from the report. Indeed, the report was in fact titled 'Learning from grant evaluations'. But the reaction of Minister De Jager was remarkably tepid. He, for example, did not agree with the court's recommendation to establish substantive quality criteria for grant evaluation in national legislation. He also did not support the introduction of a decision rule under which subsidies would be discontinued in case of a negative evaluation. Not surprisingly, the Court of Audit itself argued that only 'substantiated' insights on effectiveness are to be used in the design of government policies.

It is the 'substantiation' or plausibility of public policy which is the underlying theme of this dissertation. The plausibility of policy can be defined as the likelihood to which policy measures achieve their intended effects, and the degree to which claims about those intended effects are substantiated. In the social domain, in particular, many policies and programs are based on intuitive notions of 'what works' (cf. Rensen & Van der Kooij, 2009). Leeuw (2006: 270) warns for the favouritism of policy makers and politicians for particular theories, whose assumptions are not always in accordance with empirical knowledge. In addition, there is generally little attention for the undesirable and unintended consequences of policies (Engbersen, 2009). The neglect for substantiated policy is not just a matter of money. Social policy must contend with the reality that it often has a 'dark side' (Squires, 1990). This is visually illustrated in the book *Redirect* (2011) by the American psychologist Timothy Wilson. Wilson identifies the painful consequences of what he calls the 'failure of common sense',

---

[1] See *Learning from grant evaluations*, 2011. The Dutch Court of Audit checks if the government of the Netherlands spends public funds and conducts policy as intended.

where policymakers, social professionals and experts of all sorts rely on common sense in solving of social problems. Wilson convincingly shows that many of these presumed self-evident and feel-good motivated interventions cause prolonged stress, increase crime, lead to more drug use, make people unhappy, and, in some cases, even hasten their death.

1.2.   What works in social policy?

How can we assess the plausibility of social policies? Despite the disinterest of some politicians, the social domain has seen a growing debate on the effectiveness of measures, even though this may not as yet manifest itself in many 'sound' evaluations. Increasingly, the realization is taking shape that the legitimacy of social policy has to go beyond the intuition that it 'should' work. The (mental) healthcare sectors have already witnessed an increased emphasis on *evidence-based* practices (Roberts & Yaeger, 2006; Pope & Mays, 2006; Roe & Lysaker, 2012). Government bodies and funding agencies in these sectors increasingly demand systematic procedures and 'evidence' for the effectiveness of its methods. This paradigm has spilled over into broader social policy areas such as social work in disadvantaged areas and community based interventions (see e.g. Burton et al., 2004). But that particular transition has been far from smooth. Unlike methods in individual and mental healthcare, social programs typically have no sharp demarcation in time, intensity or target group and are implemented in a rich context of unforeseen and unknown geographical and situational variables. This makes it difficult (if not impossible) to assess their impact with research methods that centralize a monocausal effect of 'independent' interventions for well-defined target groups, like *randomized controlled trials* (which is considered the golden standard of evidence in health research[2]).

There are, however, alternative research approaches of which this thesis presents some examples. First, there is the mixed methods approach. In mixed methods research, the results of different lines of research (quantitative, qualitative, document analysis etc.) are linked. This means that conclusions are constructed by careful comparison of the results (triangulation), which in turn enables a 'higher conceptual order' of the

---

[2] A Randomized Controlled Trial (RCT) is a type of experimental research which seeks to answer whether a particular treatment (or intervention) is effective. The treatment test is conducted on a test group and compared with a control group, a similar group of people with the same problem, but that is not treated or treated with a different treatment. To ensure that no difference occurs in the classification of the groups that might affect the probability of success of the treatment, the allocation of research participants to different groups must also be done randomly. This is the meaning of the word 'randomized' in the RCT. Otherwise, it could be that the researcher who selects is affected by unconscious factors that might influence the outcomes.

phenomenon under investigation (see Teddlie & Tashakkori, 2009: 286). Mixed methods research can thus generate knowledge about the workings of policy that had remained hidden if only quantitative or qualitative research was used. To put it another way, it provides a richer picture of the functioning and impact of an intervention than a choice between either quantitative or qualitative methods. The great advantage of such approaches is that both techniques simultaneously and in conjunction benefit from each other's input, rather than a situation where a quantitative or qualitative research is set up solely for additional clarification after particular results of one or the other research method.

Second, there is a theory-driven approach. Despite the multitude of social projects and their unclear demarcations, most social interventions are rooted in some basic principles and assumptions on how and why they should work. The validity of these broader principles can be evaluated and reviewed on the basis of secondary scientific insights. The defining principle of any theory-driven approach is that - instead of measuring outcomes directly – underlying assumptions of social interventions are made explicit and are scrutinised for their validity (Chen, 1990; Shadish, 1987; Lipsey, 1993; Weiss, 1995; Pawson, 2006). Each social intervention is based explicitly or implicitly on certain assumptions about what effect it should yield and what causal mechanisms should bring about the intended change. Put together, these assumed causal pathways constitute a 'theory of change', which can be confronted with insights from the scientific literature (Weiss, 1995). Although literature reviews by no means yield unequivocal policy verdicts (Boaz & Pawson, 2005), by examining the extent to which these policy theories are supported by scientific insights, the plausibility of specific methods can be predicted more reliably.

Thirdly, there are qualitative methods of inquiry. Focusing on the natural behaviour of people and their perceptions of the social world (Yin, 1994; Denzin et al., 2005), qualitative research is increasingly used for evaluation purposes in social policy and health care. Common qualitative methods are interviews, observations and focus-groups. The health sector in particular has seen a surge in approaches and writings on *evidence-based* procedures and evaluation research that involve or require inclusion of qualitative methods (see e.g. Pope & Mays, 2006). Some of the reasons for this are an increased understanding and acknowledgement of the limits of experiment or questionnaire-based quantitative research on 'what works', a growing demand for ethical considerations in evaluations, insight into clients' and patients' well-being, and the need for a more thorough understanding of how experimentally determined *evidence based* interventions connect to people's emotions, experiences and habits. However, how to determine the validity (or 'truth value', Lincoln & Guba, 1985) of such investigations remains a difficult question, and a subject I will elaborate on in this thesis.

As for the context of this thesis, I am aware that the term 'social domain' is an ambivalent term. It can have different meanings to readers depending on their professional background or country of origin. Indeed, there is no single definition of the term 'social domain' or of 'social policy'. Moreover, especially nowadays, Western policies in the realms of healthcare, education, community organisation, urban regeneration and social work often blur into each other. I therefore commit to a loose definition of Dean (2006: 11) who purports that the domain of social policy involves the realm of wellbeing, the social relations necessary for wellbeing, and the systems by which that wellbeing may be promoted, may they be government or civil society induced.

1.3.   Presented case studies

This thesis will provide an analysis of the plausibility of policy based on four separate case studies from the social domain (see Table 1). Some case studies have been published as articles in academic journals; some represent summaries of previously published books and reports[3]. Although representing separate topics, the case studies provide - each in their own way and at different explanatory levels - insight into what works in social policy.

Case study 1 takes a mixed method approach to assess whether state-promoted civic engagement and promoting the participation of vulnerable groups (e.g. people with mental disabilities, psychiatric patients) is an appropriate government endeavour, and, in the process, exposes some of the 'dark sides' of social policy (real and potential) that are often overlooked by politicians and administrators. Case studies 2 and 3 take a theory driven approach to evaluate the plausibility of dominant policy assumptions in subsequently anti-radicalisation/polarisation and quality of life issues in deprived neighbourhoods. These case studies in effect take an opposite approach to the multi-layered mixed methods research of case study 1 by focussing on basic policy

---

[3] See: Lub, V., A. Sprinkhuizen & R. Cromwijk (2010). *Trendrapport WMO 2010. Het spel op het maatschappelijk middenveld*. Utrecht: MOVISIE.; Lub, V., N. de Groot & J. Schaafsma (2011). *Polarisatie en radicalisering: de onderbouwing van sociale interventies getoetst*. Utrecht: MOVISIE / Universiteit van Tilburg.; Lub, V. & Uyterlinde M. (2012). Evaluating State-Promoted Civic Engagement and Participation of Vulnerable Groups. The Paradoxical Policies of the Social Support Act in the Netherlands. *Journal of Social Policy*, 41, 2, 373-390.; Lub, V. (2013). *Schoon, heel en werkzaam? Een wetenschappelijke beoordeling van sociale interventies op het terrein van buurtleefbaarheid*. Den Haag: Boom Lemma.; Lub, V. (2014). *Kwalitatief evalueren in het sociale domein. Mogelijkheden en beperkingen*. Den Haag: Boom Lemma; Lub, V. (2013). Polarisation, Radicalisation and Social Policy. Evaluating the Theories of Change. *Evidence & Policy* 9, 2, 165-183.

assumptions (i.e. ideal types of theories of change) and confronting them with the scientific literature. With these two studies I aim to show that strong explanations about the plausibility of policy can be provided if they focus on theories that do not posit too much in them, like Ockham's famous reductionist principle of the 'razor'. Finally, case study 4 reflects on the question of validity in qualitative evaluation. Qualitative research is sometimes seen as too subjective and biased (allowing too many degrees of freedom in causal inferences and construction of conclusions) which of course has its problems when applied to evaluation purposes in social policy. Yet one can equally argue that the increased importance given to qualitative information in policy evaluation simply requires a more precise conceptualization of validity criteria that can be linked to evaluations purposes, and in the case study I explore the possibilities to this end.

Table 1. Presented case studies from the social domain on the plausibility of policy

Case study	Topic
1	A study on the paradoxical policies of state-promoted civic engagement and 'participation' of vulnerable groups (published in *Journal of Social Policy*, 41, 2, 2012)
2	A study on the plausibility of social policies against polarisation / radicalisation issues (published in *Evidence & Policy*, 9, 2, 2013)
3	A study on the plausibility of place-based social interventions against neighbourhood disorder (based on the book *Schoon, heel en werkzaam?,* 2013)
4	A theoretical and methodological discussion on validity in qualitative evaluation (based on the book *Kwalitatief evalueren*, 2014)

Below I describe the scope and methodology of the respective case studies in more detail.

Case study 1: a mixed methods approach
Case study 1 reports on the effects of the Social Support Act (Wmo) in the Netherlands, implemented in 2007. The Dutch Act resembles the current European political interpretation of citizenship: stressing self-responsibility in one's personal life, fighting against elements of the welfare state that are believed to be invasive and implying a shared responsibility of government and civil society for the care of socially vulnerable groups. Data were collected on the basis of two surveys evaluating the effects of the Wmo for voluntary organisations and professional non-profit institutions in social care and welfare (2007: $N = 383$ and 2009–10: $N = 389$). In addition, in-depth interviews with stakeholders and qualitative case studies were carried out. The study yields several paradoxical policy outcomes. Contrary to the objectives of the Social Support Act, a 'revitalisation' of the Dutch civil society – in terms of a greater contribution to social goals and policies – remains problematic, whilst professional entities thrive under the new governmental élan. Other paradoxical outcomes stem from a too-dogmatic approach to the social participation of people with severe mental disabilities. Instigating the socialisation of these groups through mandatory measures can in practice increase their isolation. To reduce unintended effects, I argue that the Social Support Act should take into account the divergent capacities of vulnerable groups and prioritise the psychological safety of clients over political and administrative policy objectives.

Case study 2: a theory driven approach
Case study 2 evaluates the validity of 'theories of change' of anti-polarisation and anti-radicalisation interventions. Assumptions of four dominant social policies in the Netherlands are confronted with the literature: social ecological interventions (1), peer mediation (2), intergroup contact interventions (3), self-esteem enhancement (4). The most defining underlying principle of a particular form of intervention was extracted along with the postulated relationships between the assumed 'working' components. For this purpose, generalisations (ideal types) of the various forms of interventions and their underlying assumptions were constructed (Weber, 1971 [1951]; see also Ritzer, 1992). The programme theories were reconstructed by analysis of written documents, such as programme descriptions of professional and civic organisations, government position papers and strategy outlines of funding agencies. To review the theories of change, the international literature was systematically searched. The scientific resources and impact assessments had to be related to one of the four themes and the assumed underlying causal assumptions. The literature yielded over 150 relevant international studies and impact assessments. None of the investigated policy approaches is strongly supported by the literature, though some theories of change display a moderate scientific validity. In addition, epistemological issues are discussed. Notions of 'what works and why', do not equate to straightforward application in anti-polarisation and anti-radicalisation methods. Also, the potential of the four policy

approaches depends in great measure on their conceptualisation of 'polarisation' or 'radicalisation'. A focus on interventions could be dismissed as naive, when no account is offered of wider sociopolitical factors fuelling radicalisation or ethnic tensions. I present the interaction between these two notions as a relevant topic for future research.

Case study 3: a second theory driven approach
Case study 3 evaluates the validity of 'theories of change' of social policies against neighbourhood disorder. In addition to physical measures and stricter law enforcement, deprived neighbourhoods in the Netherlands display many social and community-based interventions to tackle disorder and improve quality of life. The study evaluates the plausibility of such interventions by confronting them with insights from the scientific literature. Four policy categories were assessed: promoting neighbourhood contacts (1), resident representation and participation (2), community-based behavioural codes programs (3), neighbourhood watch schemes (4), and neighbourhood sports (5). The program theories were reconstructed by coding-analysis of written documents, such as government policy papers, descriptions of quality of life programs from professional and civic organizations and strategy outlines of funding agencies. First, an inventory was made of all active Dutch social programs designed to tackle neighbourhood disorder (ranging into the dozens). To this end, both national and local government databases were searched as well as various other databases containing descriptions of neighbourhood programs. The respective programs were then lumped into more general themes and categories of which subsequently the most defining underlying principle of a particular policy category along with the postulated relationships between the assumed 'working' components was extracted. In other words, as in case study 2, generalizations (ideal types) of the various policy themes and their underlying theories were constructed, which then served as a directive for the literature assessment. The literature yielded circa 100 relevant international studies and impact assessments[4]. According to the literature, only neighbourhood watch schemes can boast a strong scientific basis. I discuss implications for urban policy and theory, such as the need for a shift from residential networks (*horizontal* relations) to building organizational capacity (*vertical* relations) and the interrelation between instrumental and symbolic disorder measures.

Case study 4: a discussion on qualitative evaluation
Case study 4 provides a theoretical and methodological discussion of the question of validity in qualitative evaluation. Since roughly the 1970s increasing criticism of the reliability and objectivity of qualitative research has resulted in a growing interest in establishing more rigorous criteria and methodological standards. Within the academic

---

[4] See for an overview of the consulted literature: Lub, 2013b

community the idea seems to be dominant that qualitative researchers must demonstrate in one way or another that their research results are valid. Elaborating on epistemological and theoretical conceptualizations by Guba and Lincoln (1994) and Creswell and Miller (2000) the case study explores aspects of validity of qualitative research with the explicit objective of connecting them with aspects of evaluation. I will argue that different purposes of qualitative evaluation in social policy can be linked with different scientific paradigms and perspectives, thus transcending unproductive paradigmatic divisions as well as providing a theoretical framework for researchers and reviewers of qualitative evaluations.

The final chapter consists of a discussion of the need for studying policy from multiple order perspectives. In doing so, I aim to shed light on three different yet related questions that shift from the micro to the macro level: how plausible are social policies (first order perspective)? How are we better to assess their plausibility (second order perspective)? What wider social, cultural, political, institutional and paradigmatic factors might influence how the plausibility of policy is assessed (third order perspective)? I conclude the chapter with themes for future research.

Following the argumentation and case studies presented above, the main research questions of the study are formulated as follows:

What is the likelihood to which policy measures in the social domain achieve their intended effects? To what degree can claims about those intended effects be substantiated by scientific insights? How can qualitative information be validated in this respect? How can we incorporate reflexive perspectives in studying social policy?

1.4. Overview of the book

Chapters 2, 3, 4 and 5 of this thesis correspond with the presented case studies. First I will make an analysis of the effects of the Social Support Act in the Netherlands, based on mixed methods research (Chapter 2). Then I will assess the validity of 'theories of change' of anti-polarisation and anti-radicalisation interventions by confronting them with the scientific literature (Chapter 3). In a similar manner I will subsequently assess the validity of 'theories of change' of social policies against neighbourhood disorder and quality of life issues (Chapter 4). In chapter 5 I will discuss the question of validity in qualitative evaluation, elaborating on epistemological and theoretical conceptualizations by Guba and Lincoln and Creswell and Miller. The final chapter consists of a discussion of the need for studying policy from multiple order perspectives. I conclude the chapter with themes for future research.

2 The paradoxical policies of the Social Support Act

2.1. Introduction[5]

Marga Lammers[6] is showing us around a Regional Sheltered Housing Institution (in Dutch: *Regionale Instelling Beschermd Wonen* or RIBW), at a location in the west of theNetherlands. Lammers works as a site manager at the institution, which provides care to clients who have a mental impairment or multiple disabilities. The site is wedged between a centre of urban growth and a suburban residential neighbourhood. Around 400 clients live and take part in leisure activities here. They include young people, adults and elderly people with mild to severe mental disabilities, behavioural disorders or psychiatric problems. Lammers describes the RIBW institution as a 'private site with an open character'. Residents of the neighbouring suburb have free access to the area. Visitors to the site gain the impression that they have entered a village for the mentally impaired: the asphalt roads are lined with group apartments with coloured drawings displayed at the windows, minibus taxis carrying disabled people drive up and down and everywhere there are clients tearing around in go-carts.

We interview Marga Lammers about the ideal of the 'socialisation' of vulnerable groups in society, that is people who either have a severe mental disability or psychiatric disorder or who live in extreme social isolation (e.g. socially isolated elderly people, drug addicts). As a result of the Social Support Act, which was introduced in 2007, these vulnerable groups in the Netherlands are increasingly required to 'integrate with the community' in order to participate 'amongst the people' (Verplanke & Duyvendak, 2010). For the RIBW institution Where Marga works, this implies severe changes in the near future. A large number of the clients will be housed outside of the care site, in apartments amongst ordinary residents. The housing blocks are already under construction. Lammers is concerned about the imminent move:

> Soon it will be much harder for us to monitor how people are coping. For
> instance, we have to insulate bedrooms and shut clients in due to excessive
> noise levels. We have a girl who suddenly starts screaming at five o'clock in

---

[5] This chapter is a reworked version of: Lub, V. & Uyterlinde M. (2012). Evaluating State-Promoted Civic Engagement and Participation of Vulnerable Groups. The Paradoxical Policies of the Social Support Act in the Netherlands. *Journal of Social Policy*, 41, 2, 373-390.
[6] The names Marga Lammers and Peter Dijkstra are fictitious for reasons of privacy.

the morning. What's more, they can't cross the street on their own. Most clients don't know what traffic is, let alone being able to watch out for it. In the new situation, they must always be accompanied by a professional attendant at set times if they want to go for a walk. Whereas at the moment, they are able to move around the care site relatively safely. [1]

With the help of concerned family members, it was only just possible to avoid the relocation of the most vulnerable group of clients – those with an extremely low mental capacity – to individual apartments in the district. The site manager has few illusions as to the involvement of local residents when the clients are moved to the neighbouring residential area:

> We are doing everything we can to inform residents about the arrival of mentally disabled people in their environment: distributing leaflets, organising information evenings and events. But only one or two people have shown an interest. [2]

2.2.   The Dutch Social Support Act and research questions

Promoting the social participation of citizens has been one of the Dutch government's central policy objectives for a number of years. All citizens need to 'participate' more, including the socially vulnerable. From a social standpoint, this objective gained new momentum as a result of the Social Support Act (in Dutch: *Wet maatschappelijke ondersteuning* or Wmo). It is for this reason that in the Netherlands the Wmo – introduced in 2007 – is also referred to as the 'Participation Act' (see MOgroep, 2007; VNG, 2008). It encompasses the former Disabled Act (WVG), the Social Welfare Act and parts of the Exceptional Medical Expenses Act (AWBZ). In this chapter, I discuss its two main objectives. On the one hand, the Wmo aims to promote the participation and active involvement of vulnerable groups within society – for instance via active (re-)housing programs to relocate mentally disabled people to residential areas, as in the case of Marga Lammers' client groups. On the other hand, the Wmo focuses on promoting active citizenship: the participation of able-bodied citizens and their associations in the development and implementation of local social policy. Examples include residents' associations who keep an eye on isolated elderly people, sports clubs that allow young people with a disability to take part in sports or grass roots organisations providing the municipality with advice by taking part in local Wmo meetings.

The Wmo reflects the current European political trend in the interpretation of citizenship: stressing self-responsibility in one's personal life, fighting against elements of the welfare state that are believed to be invasive and implying a shared responsibility of government and civil society in the care and welfare for vulnerable groups (see Van Ewijk, 2010). In short, the Wmo makes a strong appeal for the revitalisation of civil society in the Netherlands (Hortulanus, 2004; Tonkens, 2010). It is clear that the two participation objectives of the Wmo are closely interrelated: in order to allow vulnerable groups to take part in society, it is essential for organisations that operate in citizens' immediate living environment – such as associations, companies and voluntary organisations – to ensure that socially vulnerable groups have more opportunities to take part. Primary responsibility for implementing the Wmo lies with the municipalities. The process of decentralisation is designed to bring social policy closer to citizens. Dutch municipalities are to assess which resources, facilities and manpower are required to meet the needs for support of their citizens. Municipalities therefore act as coordinators of the Wmo: they need to mobilise their local civil society. Four years have passed since the introduction of the Wmo. At the local level, the Social Support Act is beginning to be implemented. To assess its impact, the following questions are examined:

1. To what extent is it possible to raise enthusiasm amongst able-bodied citizens for the objectives of the new Act?
2. How do civil society organisations value their expected greater contribution to social goals and policies?
3. What impact do the Wmo and its associated measures have on socially vulnerable groups?
4. How tolerant is the receiving community (in neighbourhoods, civic associations and in the workplace) when it comes to the inclusion of people with mental disabilities?

Based on two studies carried out since the introduction of the Act, I claim that the Wmo is perhaps too dogmatic in its approach to increase the participation of both able-bodied and vulnerable citizens. As a result, it may produce paradoxical policy outcomes. I distinguish three participation paradoxes. The first paradox manifests itself on an institutional level. It refers to the relationship between professional and non-professional organisations. The goal of the Wmo is to revitalise civil society, however our research shows that up until now the main parties to profit from the Act are professional care and welfare organisations. The second, more profound research outcome deals with the psychological vulnerability, participation and inclusion of vulnerable groups in society. The paradox that emerges relates to the socialisation

policy that is applied to people with a *severe* mental disability or behavioural disorder – as outlined in the introduction to this article. In theory, the participation of these groups increases when they are housed in residential areas, due to a supposed increase in social contact. In reality, this policy has an ironic impact, for to allow them to live 'among the community' means that their freedom of movement must in fact be restricted. The third paradox relates to the principle of needs-based management in the care for vulnerable groups, on which the Wmo is based. Many of the people who drop out of society – for instance mental healthcare clients, addicts and socially isolated people – do not or only barely see or understand their own problems, let alone being articulate enough to express their own care needs. As is the case with those with a severe mental disability, their mental state is not sufficient for this to be possible. This can mean that they are not provided with the right care, support or incentives to take part, causing their situation to deteriorate further.

2.3.   State-promoted civic engagement and participation

To expand the logic of the empirical findings, I briefly discuss some key issues regarding state-promoted civic engagement and participation of vulnerable groups. The premises of the Wmo and its objectives can be traced back to the 1980s. In Anglo-Saxon countries such as the United States, Canada, the United Kingdom and Australia, a neo-liberal politics emerged, which focused on reduced government provision, stimulated market-based strategies in public services and emphasised the self-responsibility of citizens (Marinetto, 2003; Fletcher-Morgan & Leyland, 2010). In the Netherlands, a similar process occurred in the 1990s, which resulted in significant cuts in the welfare state and a withdrawal of government out of many public spheres. The neo-liberal strategy of the Dutch government also comprised a more instrumental (or functional) approach to citizenship. Promoting active citizenship became a means to achieve a political policy end. Previously, the government occupied a limited role in shaping citizenship, but in recent years it increasingly behaves as a gatekeeper that determines which forms of behaviour can be labelled as 'good citizenship' and which cannot (Prior & Barnes, 2009). The new political discourse implied a concrete interpretation of 'social citizenship' (Roche, 2002) and a stronger appeal to the civil society to include and care for vulnerable groups. As a result, voluntary organizations in the Netherlands were encouraged to support people who found themselves at the fringes of society.

Some authors have raised questions whether active citizenship can be utilised for the instrumental logic of government policy (cf. Verhoeven & Ham, 2010). Ostrom (2000) points to the phenomenon of 'crowding out citizenship': if citizens experience

government interventions as too controlling or compelling, it will impede their intrinsic motivation to contribute in the service for the public good. Also, a process of 'institutional isomorphism' comes to mind whereby a transfer of rationalisation and bureaucratisation processes from the competitive market place to the state makes organisations increasingly similar in their objectives and structuration (DiMaggio & Powell, 1983). Deakin (1996) explored the particular concerns of voluntary organisations in the United Kingdom in the face of new responsibilities in community and social care policy. Bestowed upon them by the government, many voluntary organisations saw themselves confronted with new legislation and even contracting procedures, stemming from the New Public Management model (Boston et al., 1996). One of Deakin's findings was that voluntary sector providers were sceptical about the potential benefits of this approach to end-users and the change in intrinsic values it implied for the organisations themselves.

A second key issue is the active involvement and inclusion of vulnerable groups within society. In Western countries, mental disability-related public policy predominantly emphasises reducing the number of people experiencing exclusion fromthe spaces of the social majority as being the pre-eminent indicator of inclusion (Milner & Kelly, 2009). In many countries, the policy aimed at people with long-term limitations has been dominated by efforts to retain and, where possible, restore independent functioning and participation in society. Hall (2004) has even argued that reducing the number of people experiencing exclusion from mainstream society is the unifying principle of social policy in the UK. In the Netherlands, the Dutch term *vermaatschappelijking* ('socialisation') is gaining currency as a means of describing a whole series of changes needed to make realisation of these aims possible (Kwekkeboom & Koops, 2005), of which the Wmo is the most recent and far-reaching addition. In many cases, these socialisation policies result in random deinstitutionalisation, whereby care institutes take clients out of their own community and transfer them to unknown residential areas.

From an ideological standpoint, the socialisation of people with mental impairments is worth pursuing. Promoting participation in society seems a prerequisite to social inclusion. However, there are some caveats that question absolute interpretations of this view. Milner and Kelly (2009) challenge the assumption that the path to social inclusion is unidirectional, that is involving people with disabilities making a journey to mainstream contexts, without any expectation that non-disabled people need to make the return journey. They argue that community participation for people with mental disabilities almost invariably involves a migration away from places where they feel known and secure to spaces in which they occupy positions of inferior knowledge, expertise or social capital. According to Milner and Kelly, services that have inverted the conventional route to participation by inviting the community to

engage with people with disabilities within the spaces they feel safe are worth contemplating. Moreover, the ways in which location may affect social participation and access to social support among people with disabilities are still unclear (McPhedran, 2011). The assumption that increased mainstream socialisation – irrespective of the locality or context – improves the overall well-being and life satisfaction of people with disability might be too one-sided. Verplanke and Duyvendak (2010) conclude that when Dutch psychiatric patients and mentally disabled people live on their own, they often receive very little care or support from the neighbourhood. Although some ex-mental healthcare clients are happy with the freedom that independent living provides them, they rarely manage to have a social life within the neighbourhood.

2.4.   Methodology and data

In 2007 – the year that saw the introduction of theWmo – we carried out the first study into the Wmo's impact in the welfare, care and civil society sectors. This study can be viewed as a benchmark for the involvement of professional and civil society organisations in theWmo, and how these organisations assess its impact. In this first study, the central research question was to what extent the services, policies, structure and the operational management changed for organisations as a result of the introduction of the Act, and what impact the Act had on staff, volunteers and clients of the organisations involved. Both quantitative and qualitative research was carried out in order to answer these questions. In 2007, a questionnaire (Wmo monitor) was distributed to implementing bodies ($N = 383$). An interview cycle ($N=25$) with representatives of sector organisations was also linked to the Wmo monitor.

A second study was conducted in 2009 and 2010. This study focused more directly on the involvement of the civil society with the Wmo – in terms of a stronger contribution to social goals and policies – and the consequences of Wmo-associated socialisation measures for vulnerable groups. Again, questionnaires were distributed in 2009 (Wmo monitor) and in-depth interviews ($N = 30$) were held in the course of 2009 and 2010. The survey was carried out amongst professional organisations in social care and welfare (e.g. general social work organisations, homecare facilities, mental healthcare institutions and welfare organisations) as well as civil society organisations such as voluntary organisations (e.g. charity work, social support associations, scout clubs), grassroots initiatives, citizens representative councils (e.g. neighbourhood platforms) and sports clubs ($N=389$, see Table 2). Qualitative data were gathered through on-site interviews with representatives of organisations and from case studies of people with mental disabilities, behavioural disorders or psychiatric problems. These

case studies were provided by a municipality official who collaborated in the research. For this article, we selected the case that most clearly illustrates the paradoxical outcomes of care avoiders, which will be discussed below in the section on paradox 3. In addition, the study involved analysing fifty municipal Wmo policy plans. This chapter mainly uses data from the second study into the effects of the Wmo.

Table 2. Distribution of organizations in the 2009 Wmo monitor sample (N=389)

Organization	Number	Percentage
Professional organizations:		
Housing associations	3	1,1
General social work	6	2,2
Home care	9	3,3
Informal care support centres	18	6,6
Elderly welfare foundations	18	6,6
Volunteer agencies	26	9,5
Nursing institutions/care homes	35	12,8
Mental healthcare	44	16,1
Welfare organizations	74	27,1
Other	40	14,7
Total:	**273**	**100**
Civil society organizations:		
Ideological organizations	2	1,7
Citizen representation councils	23	19,8
Voluntary organizations	69	59,5
Other	22	19
Total:	**116**	**100**

2.5. Results

Paradox 1: Civil society at the sidelines
The first study in 2007 revealed that the introduction of the Wmo had caused a great deal of commotion within the professional field. Professional care and welfare organisations extended their own professional networks and sought to coordinate with strategic partners and the municipality (Lub et al., 2008). However, the Act in 2007 largely passed over civil society organisations. An important sign in the response from the civil society organisations was that they did not recognise the Wmo as an Act that

had anything to do with them. The follow-up study in 2009 and 2010 shows that more than two years later, the involvement of civil society organisations was still as limited as it was when the Act was first introduced. Moreover, there are signs of a rift between civil society organisations and professional organisations in terms of their assessment of the Wmo. The second survey in particular reveals strong evidence of this. Table 3 illustrates the distance between the professional sector and civil society based on response percentages to a number of assertions presented in the Wmo monitor 2009.

Table 3. Responses to assertions in the Wmo monitor: differences between professional and civil society organizations.

	Percentage of professional organizations that agree	Percentage of civil society organizations that agree
Our organization plays an important role in the implementation of local Wmo policy	65%**	36%**
Wmo policy mainly excels on paper and in meeting rooms	51%*	57%*
We are taken seriously in terms of contributing ideas, shaping and implementing Wmo policy	55%**	38%**
The Wmo opens up new opportunities for us	52%**	38%**
It is not clear what is expected of our organization within the Wmo	27%**	45%**
The Wmo causes some target groups to wrongly fall out of the picture	45%*	53%*

*Difference between mean score of professional and c.s. organizations significant at 5% (ANOVA)
** Difference between mean score of professional and c.s. organizations significant at 1% (ANOVA)

It is clear that professional organisations feel that they are taken considerably more seriously in terms of contributing ideas about the Wmo and shaping and implementing Wmo policy than civil society organisations. Also when it comes to new opportunities perceived by organisations from civil society as a result of the Wmo, professional institutions are significantly more positive. The Wmo evidently creates more opportunities for professional organisations than for civil society organisations. Civil society organisations also generally have a more negative attitude towards the Wmo as

a system reform. They have a greater tendency than professionals to believe that the Act mainly excels on paper and in meeting rooms, whilst, in their view, the Wmo causes some target groups to wrongly fall out of the picture.

An initial explanation for the limited involvement of civil society is that municipalities do not have sufficient insight into the spectrum of civil organisations in their jurisdictions and are occasionally lax in involving their citizens in policy making. The case studies reveal that many municipalities are still trying to identify a proper way to meet their relatively new responsibilities as coordinators of the Wmo, and the civic partners that could support them in this task. This means that they are less able to clarify to civil parties what they expect from them. Although most municipalities state in their Wmo policy plan that they are making efforts to involve civil society, specific organisations are rarely mentioned. In other words, the intention is there, but not the ideas for realisation (see also Van Marissing & De Meere, 2009). All of this means that Wmo policy is still something that primarily affects the professional sector. Metz (2010) lucidly illustrated this on the basis of a local case study into Wmo policy in the Dutch municipality of Dordrecht: pressure from government agencies and professional institutions left no scope for input from civil society organisations.

However, the limited involvement of civil society organisation in the Wmo cannot be solely attributed to a wait-and-see approach by municipalities or to 'crowding out citizenship'. A second explanation relates to the fact that many voluntary organisations and associations feel that they have a limited involvement in the Act because they themselves do not (want to) label themselves as 'care organisations'. A significant obstacle is that part of civil society is not jumping at the prospect of being formally addressed about the results of social policy objectives, such as including vulnerable groups in their regular activities. For instance, can a scout club be expected to be able to adequately cope with children with a disability? And who is responsible if something happens to them? As a result, voluntary citizen associations are often reluctant when it comes to placing a greater focus on typical Wmo policy in their activities, such as the inclusion of groups of people who need support, or forming ties with the municipalities and professional institutions from the realm of care and welfare. This is further illustrated by the fact that only a limited proportion of civil society organisations in the Wmo monitor describes themselves as an important entity in the implementation of local Wmo policy (see Table 3).

Paradox 2: how socialisation can lead to isolation

A salient outcome of the study in 2009–10 is that professional organisations that are involved in supporting people with a mental disability, psychosocial care and addiction services (such as public mental healthcare organisations and care institutions, $N = 72$) have a significantly more negative attitude towards the impact of the Wmo than other professional respondents in the sample. For example, 61 per cent of these institutions agree with the assertion that the Wmo causes target groups to wrongly fall out of the picture, as opposed to 39 per cent of the other professional organisations (see Table 4). Table 3 further illustrates that results on assertions concerning the positioning and allocation of resources in Wmo policy point in the same direction.

Table 4. Responses to assertions in the Wmo monitor: differences between professional organizations providing psychosocial care (N=72) and other professional organizations in sample.

	Response percentages of professional organizations providing psychosocial care	Response percentages of other professional organizations in sample
The Wmo causes some target groups to wrongly fall out of the picture	61% agree**	39% agree**
The Wmo leads to a higher quality of life for vulnerable citizens	54% disagree*	36% disagree*
Our organization plays an important role in the implementation of Wmo policy	45% agree**	72% agree**
The municipality allocates enough resources in order to implement Wmo policy adequately	10% agree**	22% agree**

*Difference between mean score of psychosocial and other professional organizations in sample significant at 5% (ANOVA)
** Difference between mean score of psychosocial and other professional organizations in sample significant at 1% (ANOVA)

Just as in the introductory year of the Wmo, there is uncertainty in the field of mental health services about the loss of professional support for vulnerable groups, the associated socialisation measures and how the situation of clients will turn out under the new regime of the Wmo. Although socialisation policies in the Netherlands stem from long before the introduction of the Wmo, they did receive an extra boost with the advent of the Wmo. The fifth performance indicator of the Act explicitly states the

objective of promoting the social participation of people with disabilities and persons with chronic mental or psychosocial problems.

At the start of this article, I outlined the experiences of Marga Lammers who, as the Site Manager of a RIBW institution, experiences daily the implementation problems associated with the policy ambition to allow people with psychological problems to function independently in society. Staff at the institution have strong objections to the relocation of mentally disabled client groups to the neighbouring residential area. This deinstitutionalisation – that is relocating clients out of institutional care facilities − could have undesired consequences for clients' personal integrity and safety. Parents and family members are also concerned. The neighbourhood has problems with loitering teens and the question is how this will affect the safety of the clients. Family members are afraid clients will be harassed by young people or that people will laugh and stare. The bullying of disabled persons in the area by street youths is not a rare occurrence. For the RIBW institution, the socialisation process leads to a number of complications. The site formally belongs to the care institution, but anyone is free to access the grounds. Staff at the care institution are therefore faced with a dilemma:

> We can only approach residents to a limited extent. This will soon become even more difficult when our clients move into the neighbourhood. In their own living environment, residents might be even less tolerant of people with a disability. [3]

One of the attendants regularly accompanies clients to the riding stables and the vegetable garden on the care site. Some local residents let their dogs out in these areas, while some clients are terrified of dogs. According to the attendant:

> If I ask the dog owner to keep the dog on the lead, I usually get the response that 'it is a very friendly dog'. But clients cannot make this distinction and I can't get the owner to understand this. There is also regularly dog dirt lying around. People don't clean it up, but some clients will put it in their mouths. It is not simply the case that the clients need to get out into the neighbourhood more, the neighbourhood also comes to us. [4]

The rehousing of the clients not only places greater demands on the professional supervision of clients. It also has a paradoxical impact on their freedom of movement, personal integrity and safety. In their new living environment, homes need to be insulated to prevent excessive noise levels, clients can no longer cross the road without professional assistance and a vital question is whether the interaction in the public

space between residents and professionals and between residents and clients will go smoothly –which could lead to a further deterioration in the problems of clients. As becomes clear from the citations above [3, 4], the current interaction with residents on the protected care site does not hold out great prospects. Based on the information we gathered on individual cases and the results of the questionnaires as part of the Wmo study, we can conclude that care institutions are experiencing diminishing tolerance within society for 'people with difficulties'. Plans in the Netherlands to locate facilities for addicts or people with psychosocial problems in residential areas often lead to protest, as two regional managers of a care facility for drug addicts [5] and psychiatric care [6] explain:

> Citizens do not automatically accept their 'fellow-citizens' in their direct surroundings - especially when these fellow-citizens display odd behaviour, are prone to addiction or possess mental handicaps. It is very important to realise that at any time and in any area, there are opposing forces that can stir up the whole neighbourhood. [5]

> Some housing associations are reluctant to house psychiatric patients, because people don't want them in their neighbourhood. Our clients are very stigmatised. The housing association will say: 'hey, that's a guy from the psychiatric care facility, we don't want those nutcases'. Believe me, it's true, people just don't want them. [6]

Case studies of Verplanke and Duyvendak (2010) indeed show that Dutch clients with psychosocial problems who are housed individually are rarely received with open arms by local residents. This means that by living in an ordinary neighbourhood they run the risk of becoming socially isolated.

Paradox 3: the lack of demand for care from care avoiders
To encourage employment amongst socially vulnerable groups, many Dutch municipalities establish a link between the Wmo and the Work and Social Assistance Act (Wwb). A similar socialisation paradox can occur within this group as the one emerging in the context of the living environment of mental healthcare clients. As the problems encountered by this group often extend beyond a lack of employment skills, aiming to achieve social mobilisation often does not lead to participation in society. Worse still, instead of reducing the gap between this group and 'mainstream' society, this can actually serve to reinforce the position that these people occupy on the fringes of society, leading to a relatively high risk that they will slip through the net. It is not rare for them to actively seek to avoid care. In other words, they no longer accept

assistance or refuse to cooperate where assistance is provided. Care avoiders tend to 'withdraw' or 'give up' to a certain extent (see e.g. Schout, 2007; Linders, 2010). This brings me to the discussion of the third participation paradox: the lack of demand for care from care avoiders.

Peter Dijkstra (age twenty-four) meets the above profile. Since his youth, Peter has experienced problems with severe mood swings and he tries to suppress his emotions by smoking weed. As a result he is often tired and listless, and constantly depressed. His parents receive benefits, as do many of his family members. As a young person, he lived on the street for a number of years. He finally accepted help from the housing association and was assigned a flat. He is often mistrustful of others and becomes aggressive when he feels he is being placed under pressure. He also becomes stressed when he is required to do something. He did not finish school and was dismissed from the few jobs that he has held because he failed to turn up.

Under political pressure, social services of Peter's municipality have initiated a mass review of their client base. All social assistance clients are being screened to determine their ability to participate. Although Peter was rejected a few years ago due to psychological problems, he is also being called up. His diagnosis was 'provisional', which means that according to the municipality there is a chance that he may now once again be able to make a partial or full contribution towards society – for example, through employment. He is receiving support from an organisation that provides addiction services, however his case manager at social services does not consider it necessary to contact his care provider. Peter is signed up for a work−study programme to become a welder, which will enable him to obtain a welding diploma in one year, funded by the municipality. This will make him more attractive to employers. Following an initial interview with the company carrying out the project, Peter is admitted. The municipality invests €17,500 in Peter's training and education.

The first week everything goes well. Peter turns up in a presentable manner and pays attention during the classes. He has little contact with the teacher and fellow students, and during breaks he regularly stands outside smoking a joint. In the second week, Peter is required to carry out an assignment with another student. The two get into an argument and Peter becomes aggressive. The supervisor is used to 'difficult' students, but he is unable to get Peter under control. Eventually the police are called. Once at home, Peter decides to give up. He no longer leaves his house. On Monday, he fails to turn up again. After some investigation, the case manager contacts Peter's care provider at the addiction services. They visit his home on a number of occasions, but the curtains are shut and they can see no movement. They decide to push a letter in the letterbox, which they leave hanging out. A day later the letter has disappeared, indicating that Peter is at home. After five weeks, the training centre decides to abandon the course in consultation with the municipality. When his care provider

finally manages to contact Peter weeks later, it emerges that he became stressed as a result of the course and the people around him. He wasn't able to cope, he said, and he wanted to commit suicide. The care provider eventually arranges for him to be admitted in a psychiatric institution, because of his extremely filthy home and his continuous threats to take his own life if the municipality pressures him to do anything outside of his home.

Peter Dijkstra's case exposes a number of unintended effects. It shows that the principle of needs-based management, which forms the basis for current theory regarding participation, can turn out to be very problematic for some target groups. Many socially vulnerable people are unable or barely able to manage their own lives, let alone express their own need for care. The question arises whether government officials and social professionals *are* able to identify the support needs of such clients adequately (see also Van Rooyen, 2004). Peter's case shows that when professionals decide on their own what is good for them, this could backfire. The result is often a very difficult relationship with organisations that provide care, mistrust of care providers and sometimes even the complete avoidance of care and support (see also Schout, 2007; Linders, 2010). Encouraging employment amongst people with psychosocial problems can in fact cause clients to run into even greater difficulties. When Peter got stuck in the work−study programme, he completely withdrew into his own world.

2.6. Conclusions

The first participation paradox illustrates that, contrary to the objectives of the Wmo, a 'revitalisation' of the civil society remains problematic whilst professional entities thrive under the new governmental élan. Since the introduction of the Act, limited progress has been made in mobilising informal civil society in terms of a stronger contribution to social goals and policies. The Wmo expects a great deal from these mainstream civil society organizations (e.g. sports associations, voluntary clubs and religious organizations) that operate in the immediate living environment of citizens. Despite the appeal the Wmo makes for more social citizenship (inclusion of vulnerable groups, contributing to social policy making, linking care and welfare, etc.), they feel very little involvement in the new Act. As outlined earlier, the two participation objectives of the Wmo are closely interrelated: in order to allow vulnerable groups to take part in society, it is essential for civic organisations to ensure that socially vulnerable groups have more opportunities to take part. In this respect, our study exposes a number of barriers:

1. A wait-and-see approach by municipalities to actively involve civil society.
2. 'Crowding-out' citizenship by government bodies and professional institutions.
3. Restraint and inability of civic organisations to include and share responsibility for vulnerable groups.

Further research will have to clarify other possible dilemmas for volunteer organisations and citizen associations, and how these can be made more manageable. This may provide clues as to how the 'social return' of ordinary citizen associations can be increased, without having to sacrifice their core activities or be held unduly responsible for formal social policy objectives. The second and third participation paradox illustrate that a too-dogmatic approach to the social participation of vulnerable groups through mandatory measures can actually increase their isolation. The bottleneck lies in the assumption that increased socialisation into mainstream contexts, irrespective of the locality, improves the overall well-being and life satisfaction of people with mental impairments and disorders. The experiences of Marga Lammers and the case of Peter Dijkstra show that the causal relationship is not so clear-cut. It can even be counterproductive, sending clients further into a downward spiral.

      The assumption that the number of people who end up dropping out of society can be reduced by promoting their participation is in fact a circular argument: some people do not participate enough because they *cannot* participate enough. Kunneman (2000) states that modern society can be seen as a 'technopolis', in which inevitably people live who 'have nothing to offer'. They cannot keep up with the high pace of life and therefore lag behind. For them, society has become too complex to order their life by means of the resources available to them from society. Rauch and Dornette (2009) for instance show that when it comes to entering the labour market, simply applying the same rights to people with disabilities and to non-disabled persons does not help to resolve the problems faced by people with disabilities regarding their labour market integration. In light of this view, Bauman (2004) states that it is a natural characteristic of technologically developed and modern societies to produce 'waste': people who cannot keep up fall by the wayside. Of course, this does not imply that vulnerable groups should be marginalised. But it does challenges the notion of 'normalisation', the ideal that people with mental impairments should enjoy patterns and conditions of everyday living as close as possible to the norms and patterns of the mainstream society (Nirje, 1969; Yates et al., 2008). Normalisation *de facto* starts from the assumption that being 'different' is less desirable than being 'normal' and that disabled people should therefore strive to be something other than what they are (Morris, 1991).

2.7.   Discussion

It should be noted that the findings do not argue *against* promoting a certain level of participation of socially vulnerable people. Neither do they imply a renunciation of the Dutch government's encouragement of active citizenship. I do propose that the underlying assumptions of state-promoted participation should take a more nuanced and differentiated approach. The mental healthcare workers we interviewed emphasised that they do not a priori oppose socialisation. On the contrary, they would like nothing better than for ordinary citizens to learn to understand what it is like to have a mental disability or to live with a psychological disorder. These professionals are concerned that politicians and the public administration are failing to distinguish between target groups. Therefore, we advocate a differentiated approach to participation. First, a differentiated approach takes the *divergent capacities of various vulnerable groups* into account. Not all vulnerable groups are prone to paradoxical effects of participation policy. With certain adjustments, young people with a mild mental disability are usually able to live 'a normal life' in a residential area.And for elderly people in need of support, daily activities can often be helpful to take part in the community. But as we have seen in the case of the client groups of Marga Lammers, people with very limited mental capacities or *severe* psychological problems tend to fare better in a safe and specially created environment. Simply transferring these groups to residential areas is trying to cut too many corners at once. In this respect, the findings are at odds with recent Australian research of Muir et al. (2010), which suggests that meaningful participation in the community is possible, even for people with high levels of psychiatric disability (if a certain level of housing support is available). I purport that this is only possible for people with low levels of psychiatric disability, and, even then, the outcomes are not exclusively positive (see the case of Peter Dijkstra).

Second, a differentiated approach *prioritises the psychological safety* of clients over the objectives of participation policy. Dutch Wmo policy appears to be primarily the domain of politicians and civil servants who are attempting to tackle inactivity starting from a particular political discourse. Whether the associated measures are in the interest of the groups in question is often ignored. The potential of the 'inverted route to participation', as an alternative to the unidirectional approach of socialisation is challenged by some of our findings (see e.g. citation [4]). Social contact between ordinary citizens and people with mental disabilities might allow people without disabilities to 'see reflections of their shared humanity' (Milner & Kelly, 2009: 60), but on a more mundane level it can also lead to tense interactions in the public space. The question is whether Dutch policy-makers and implementing bodies sufficiently take the

negative dynamics of perceptions on mental disability into account. Instead of requiring mentally disabled persons to adapt to a potentially hostile social environment, targeting public opinions and policies that contribute to negative or ignorant perceptions of the mentally disabled might be a more fruitful route (Culham & Nind, 2003). Although expectations of increasing the understanding of the general public should not be too rosy (see e.g. citation [2]), more awareness can be created through education, for example encouraging children to fulfil school internships at institutions that provide psychosocial care, or providing government information campaigns that show how to deal with mentally impaired persons in the workplace or the neighbourhood.

In any case, the present study indicates that state-promoted civic engagement and participation of people with severe mental disabilities must extend beyond views that are too dogmatic. If not, the very policies that aim to improve the quality of life of vulnerable citizens are at risk of causing the opposite effect.

3 Polarisation, radicalisation and social policy

3.1. Introduction[7]

Throughout Western Europe, polarisation and radicalisation are frequent topics of discussion and study. Polarisation can be described as a sharpening of divisions between groups that share certain social, cultural or religious traits. Radicalisation refers to a process whereby citizens, starting from a serious discontent, develop thoughts or plans to disrupt the existing social or political order using violence or intimidation. In the Netherlands, government interest in radicalisation seems to have diminished over the last few years (in the absence of terrorist attacks). Polarisation, on the other hand, can rejoice in a growing interest (see, for example, RMO, 2009). According to some, there is a widening social and cultural gap between immigrant and native populations and between minority ethnic groups themselves (Van den Brink, 2006; SCP, 2009). Notten and Witte (2011: 61) argue that not radicalisation, but polarisation is the 'real problem' in the Netherlands.

After the murder of filmmaker Theo Van Gogh by the Muslim radical Mohammed B., riots by right-wing extremist youth and clashes between migrant youth in disadvantaged neighbourhoods, various methods have been designed to counter polarisation and radicalisation tendencies. Currently, the Dutch government funds dozens of social policy interventions in this area (Lub, 2009). These include:

- empowerment training for 'vulnerable young Muslims';
- personal and social support programmes for right-wing extremists;
- peer mediation in group conflicts;
- dialogue activities for young people of different ethnic origins, religions or subcultures.

Many of these interventions resemble local social policy measures in other European countries. Germany, Norway and Sweden have prolific experience in dealing with right-wing extremist youth groups and mainly target these issues through social support services and 'exit' programmes (Englund, 2002; Bjørgo et al, 2005; Rommelspacher, 2006). In France and the United Kingdom (UK), radicalisation of second-generation

---

[7] This chapter is a reworked version of: Lub, V. (2013). Polarisation, Radicalisation and Social Policy. Evaluating the Theories of Change. *Evidence & Policy* 9, 2, 165-183.

immigrant Muslims, as well as cultural clashes between underprivileged minority youth groups, have prompted governments to sponsor moderate religious alternatives for at-risk youth or provide programmes that aim to foster a sense of social cohesion (Bouzar, 2006; Brighton, 2007).

It remains unclear, however, whether underlying assumptions of social policy against polarisation and radicalisation are valid and whether the policy measures achieve their intended outcome. Evaluations of measures in the field of polarisation and radicalisation are scarce or hardly meet scientific standards. This leaves many questions about the effectiveness and legitimacy of the designed social interventions unanswered. How susceptible are Muslim youth to religious radicalism and what does their level of self-esteem have to do with this? How effective are peers in mediating in intergroup conflicts? Are right-wing extremists indeed 'softened up' by improving their personal problems and social living conditions? And does intergroup contact between hostile youth groups promote mutual understanding and cause a decrease of animosity, and if so, under what circumstances?

This chapter evaluates the validity of causal mechanisms implicit in social policy interventions in the field of polarisation and radicalisation, by confronting them with scientific evidence from the literature.1 For this purpose, a theory-driven evaluation perspective was used. The defining principle of any theory-driven approach is that assumptions of social interventions are made explicit and are scrutinised for their validity (Chen and Rossi, 1983; Chen, 1990; see also Shadish, 1987; Lipsey, 1993; Weiss, 1995). Each social intervention is based on certain assumptions about what effect it should yield and what causal mechanisms should bring about the intended change. Put together, these assumed causal pathways constitute a theory of change. By examining the extent to which these policy theories are supported by scientific evidence, the potential effectiveness of specific methods can be predicted more reliably. Many interventions against polarisation, for example, focus on intercultural dialogue techniques, bringing together youngsters from different ethnic origins or religions. The underlying principle, broadly rooted in the contact hypothesis of Allport (1954), is that communication will lead to reduced prejudice. But is this assumption plausible? At first glance, it seems self-evident that bridging contact leads to more tolerance towards 'the other'. Yet it is equally conceivable that it yields a negative effect. Perhaps negative stereotypes between groups are confirmed by the contact, through which mutual tensions could increase instead of decrease, a process described in conflict theories by Campbell (1965) and Sherif (1966).

3.2. Conceptualisations and research questions

Earlier surveys show that almost all Dutch interventions described in the field of polarisation and radicalisation can be linked to four themes, which overlap with social policy measures in other European countries (Lub et al, 2011). The present study evaluates the validity of their theories of change. The four types of intervention are:

- *Social ecological interventions:* These interventions aim to prevent or counter extremist behaviour of young people by offering support within their social ecological context (family, peers, school, work). They apply a coherent method of social support with a focus on problematic family relations, associations with negative peers ('wrong friends'), performance at school, adequacy of housing and perspective on work. The approach resembles multi-systemic therapy, designed for juvenile offenders (Borduin, 1995; Henggeler et al, 1996), and mainly targets right-wing extremist youth.
- *Peer mediation:* The premise of peer mediation is that young people are better equipped to remove tensions and mediate in conflicts between hostile youth groups than (adult) professionals.
- *Intergroup contact interventions:* These interventions are designed to increase tolerance between young people of different ethnic origins, religions or subcultures. Their purpose is to promote mutual understanding and remove (potential) animosity and to challenge stereotypes or prejudice. Intergroup contact interventions often include activities such as role plays or exercises in debate techniques.
- *Self-esteem enhancement:* This is geared towards 'empowering' supposed mentally vulnerable youths, perceived as susceptible to radicalism. These types of interventions usually provide individual counselling programmes or empowerment training for groups of adolescents from minority ethnic groups.

It is important to note that a particular 'theory of change' of a given intervention does not refer to a social theory in the general scientific sense. Instead, it comprises a set of causal assumptions and mechanisms (pathways), which are implicitly bound up in a given policy intervention or social programme (Weiss, 1995). However, the intervention theories of the respective policy approaches central to this article do display similarities with conventional social scientific theory. The social ecological approach and peer mediation can be traced to the social bond theory derived from criminology (see Hirschi, 1969). This emphasises an absence of social attachments among juvenile delinquents, whereby attachment to families, commitment to social

norms and institutions (school, employment) and involvement in mainstream activities are believed to be important starting points for behavioural change (see Hirschi, 1969: 16). Underlying assumptions of intergroup contact interventions overlap with the contact hypothesis, first articulated by Allport (1954). Allport's theory states that interpersonal contact is an effective way to reduce prejudice or hostility between members of different (sub)cultures or ethnicities, and such ideas form the basis of introductory activities for hostile youth. Finally, the similarities between self-esteem enhancement and social identity theory are apparent. Social identity theory states that a significant portion of an individual's self-concept derives from perceived membership of a relevant social group (see Tajfel & Turner, 1986). In the context of radicalisation issues, this notion strongly resembles dominant government assumptions that immigrant (Muslim) youth often lack a sense of social belonging in the wider society (alienation) and as a result struggle with their identity and are more vulnerable to radical thoughts or influence by radical people.

Note that the different social policy approaches to addressing polarisation and radicalisation overlap in some conceptual areas, but draw dividing lines in others (see Table 5). The social ecological approach and peer mediation, for example, have similar theoretical roots but differ in their conception of polarisation/radicalisation. Social ecological interventions perceive these problems predominantly as deviant behaviour by youths, while peer mediation is more geared towards group conflict. Intergroup contact can also be categorised in the framework of group conflicts, but peer mediation and intergroup contact approaches do not share their theoretical underpinnings in the scientific sense (social bond theory versus contact hypothesis).

Table 5. Conceptualisations of social policy approaches to addressing polarisation and radicalisation

Policy approach	Overlap with scientific theory	Conception of polarisation/radicalisation	Assumed key to success
Social ecological	Social Bond Theory	Individual pathology	Social support
Peer mediation	Social Bond Theory	Group conflict	Peer influence
Intergroup contact	Contact Hypothesis	Group conflict	Mutual understanding
Self-esteem enhancement	Social Identity Theory	Alienation	Empowerment

As is clear from Table 5, scientific theories and policy theories can overlap and they are indeed often discussed interchangeably. From an analytical standpoint, however, they can be quite different. Often scientific theories obtain a different meaning when they are translated into policy interventions. In other cases, the boundaries between scientific theories and policy theories are not clear to begin with and an intervention theory does not necessarily have to fall back on a scientific underpinning at all. Considering the policy approach of the study, this article is more concerned with intervention theories than with scientific theory. The divergence in overlap with scientific theory, conceptions of polarisation/radicalisation and assumed keys to success is also why the four themes have specifically been studied as four distinct approaches.

In the last few decades, the fields of social psychology and the humanities have yielded valuable insights into identity formation, cognitive processes related to intolerance, stereotyping of groups and mechanisms that trigger violent and antisocial behaviour by youths. This enables the consultation of a wide variety of literature to establish whether implicit policy theories on what works against polarisation/ radicalisation, as well as how and why they work, have any plausible basis. Indirectly, such an investigation also sheds light on to what extent more general social theoretical notions (such as the contact hypothesis and social bond theory) are applicable in the context of polarisation and radicalisation. The research questions read as follows:

- To what extent are underlying causal mechanisms implicit in social policy in the field of polarisation and radicalisation supported by scientific evidence?
- What particular assumptions can be considered valid or invalid according to the literature?
- What implications does the knowledge gained have for future polarisation/ radicalisation policy?
- What are the epistemological potentials or shortcomings of a theory-driven perspective in this area?

3.3. Methodology

Within the field of social programme evaluation, theory-driven evaluation has become increasingly associated with qualitative action research (see, for example, Weiss, 1995; Green & McAllister, 2002). In such qualitative-oriented, theory-based approaches, the plausibility of underlying theories is usually gauged on the basis of experiences of practitioners. The goal is to determine the nature and meaning of the intervention to

uncover information on how and why a particular outcome should be achieved or not achieved. The present study deviates from the qualitative theory-based approaches in a number of ways.

First, the treatment models of the different polarisation/radicalisation interventions were not specified at length. Instead, the most defining underlying principle of a particular form of intervention was extracted along with the postulated relationships between the assumed 'working' components. For this purpose, generalisations (ideal types) of the various forms of interventions and their underlying assumptions were constructed (Weber, 1971 [1951]; see also Ritzer, 1992). Second, instead of interviewing practitioners, the programme theories were reconstructed by analysis of written documents, such as programme descriptions of professional and civic organisations, government position papers and strategy outlines of funding agencies. Finally, and most importantly, the 'action research' component of qualitative theory-based approaches, which views stakeholders as relevant sources of information on 'what works, how and why', was omitted. By contrast, existing, peer-reviewed scientific literature as the standard of comparison was focused upon. Quantitative impact studies (meta-analyses, experimental research and longitudinal studies), but also rigorous (peer-reviewed) qualitative research, served as the quintessential touchstone.

To review the theories of change, the international literature was systematically searched. The scientific resources and impact assessments had to be related to one of the four themes and the assumed underlying causal assumptions. No restrictions were made in terms of publication date and size of the study. The literature yielded over 150 relevant international studies and impact assessments, mostly from Europe and North America. Whether the identified studies are representative of all studies conducted in this area is not known. Not all research useful to the study may have been published. We assume, however, that the literature we collected contains the most relevant insights, partly because during the research, academic experts were consulted on the various topics.

3.4.   Evaluating the theories of change

The social ecological approach
Social ecological interventions comprise methods that aim to prevent or counter extremist behaviour of young people by offering help within their social ecological context (family, peers, school, work). These interventions apply a coherent method of social support with a focus on problematic family relations, associations with negative peers ('wrong friends'), performance at school, adequacy of housing and perspective on work. The approach resembles multi-systemic therapy, designed for juvenile

offenders (Borduin et al., 1995; Henggeler et al., 1996). The underlying assumed working mechanism is that by enhancing the social ties of radical youths and their relationship to their broader social environment, personal problems are minimised and feelings of social deprivation are reduced. This should in turn reduce the (risk of) radical behaviour. In Germany, the Netherlands and Scandinavia, this approach has been applied widely to right-wing extremist youth. The ideal-typical assumptions and causal pathway of the social ecological approach are provided in Figure 1.

Figure 1. Ideal typical assumptions and causal pathways Social Ecological Approach

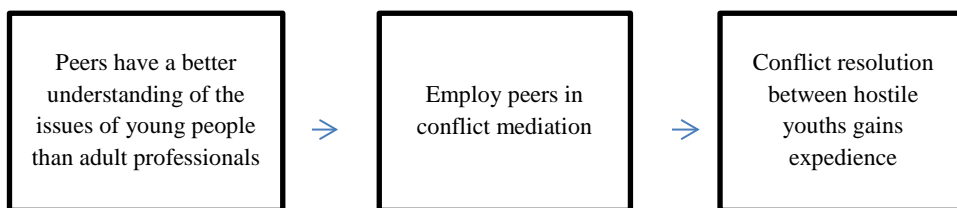| Enhance and improve social ties | → | Solve personal problems | → | Reduces feelings of social deprivation | → | Reduction of radical thoughts and behaviour |
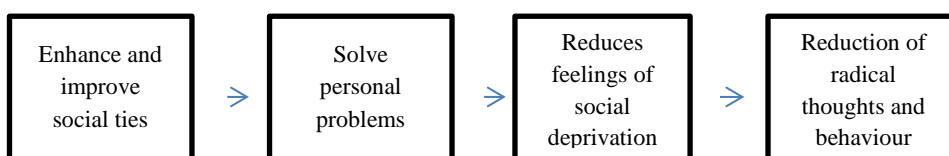
Effect studies show that a social ecological approach is effective against socially problematic behaviour of young people and generally yields better results than alternative treatment programmes (Curtis et al., 2004; Borduin et al., 2009). At-risk young people who undergo such an intervention show a decrease in antisocial behaviour, violence and criminal activities. The efficacy of a social ecological approach is explained by the fact that problematic behaviour of young people is contained by improving and strengthening family ties, thus increasing the problem-solving ability of the family. This makes the prevalence of other negative impact factors (e.g., criminal friends) less likely.

Case studies from Germany, the Netherlands, Norway and Sweden indicate that social ecological interventions also increase the chance of a 'normal life' of extremist youth – in much the same way as it does with young criminal offenders (Englund, 2002; Bjørgo et al., 2005; Rommelspacher, 2006; Demant et al., 2009). By applying a consistent treatment, radical and at-risk youth can improve their personal living conditions and strengthen their social ties outside their radical circle. This, in turn, can lead to a steady decline in extremist behaviour and withdrawal out of radical movements. It is difficult to say whether this applies to both right-wing 'white' extremist youth and young immigrant Muslim radicals. Most of the reported successful cases apply to right-wing extremists. However, it is certain that the approach works mainly with young hangers-on and less on hard-core radicals. Empirical results furthermore do not clarify whether a normal life of radicalised individuals also leads to an *ideological* change. It is not assured that when a radical young person improves their social living conditions and changes their lifestyle, this automatically leads to a renunciation of radical ideas.

Peer mediation

The premise of peer mediation is that young people are employed to remove tensions and mediate in conflicts between hostile youth groups. The underlying assumption of peer mediation is that peer mediators are a natural part of the social environment of young people, therefore have a better understanding of the issues of other young people, and consequently exert a stronger influence on them than adult professionals (see, for example, CCV, 2010, and SMN, 2008, in the Netherlands; and the Peer Mediation Network, in the UK). In many European countries, this line of thinking speaks to the imagination of both administrators and policy makers. The ideal-typical assumptions and causal pathway of peer mediation are provided in Figure 2.

Figure 2. Ideal typical assumptions and causal pathways Peer Mediation

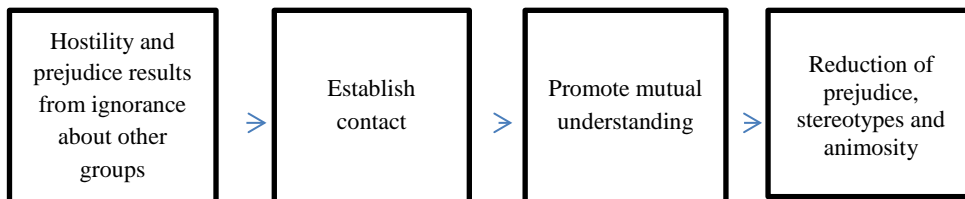| Peers have a better understanding of the issues of young people than adult professionals	→	Employ peers in conflict mediation	→	Conflict resolution between hostile youths gains expedience

An investigation into the plausibility of peer mediation produced a diffuse picture. According to available scientific effect studies on school-based peer mediation (mostly from the United States), mediation by peers – in most cases – dilutes polarised relationships and conflicts between young people (for a systematic review, see Jones, 2004). The causal link, however, between the employment of young people and the positive outcome of peer mediation remains to be established. As far as we can derive from the literature, there are no studies available that systematically compare similar mediation methods with different mediators (peers versus adults). It is therefore unclear whether the benefits of peer mediation are *caused* by the fact that the mediation is performed by peers. Other, less imaginative, factors might be responsible for the positive results, such as the close attention that is given to the causes and consequences of the conflict itself. As yet, there is no scientific basis for the theory that when at-risk youth are supervised by someone from their 'own group', this yields a more favourable result than when they are mediated by someone who is supposedly socially or mentally farther removed from them, such as an adult professional. It also appears that adolescents who are specifically trained as mediators benefit most from this approach. Some comparative and longitudinal studies show that they significantly improve their conflict resolution and social skills due to the training and professional support they receive. Yet a positive impact of peer mediation on the mediated youths is much

smaller and in many cases even absent (see, for example, Jones, 1997; Lane-Garon, 2000; Bickmore, 2002).

Intergroup contact
To reduce polarisation, interventions have been designed that aim to improve contact and increase tolerance between young people of different ethnic origins, religions or subcultures. Their purpose is to promote mutual understanding, remove (potential) animosity and challenge stereotypes or prejudice. In the Netherlands, intergroup contact interventions often include activities such as role plays or exercises in debate techniques. The central assumption underlying this approach is that hostility towards other people can result from ignorance about 'the other'. The ideal-typical assumptions and causal pathway of intergroup contact are provided in Figure 3.

Figure 3. Ideal typical assumptions and causal pathways Intergroup Contact

| Hostility and prejudice results from ignorance about other groups		Establish contact		Promote mutual understanding		Reduction of prejudice, stereotypes and animosity

Meta-analyses by Pettigrew and Tropp (2006, 2008) indeed show that, on average, prejudices and stereotypes about other groups are significantly reduced by intergroup contact. But there are several caveats. The impact of the positive effect is generally small. Reported effect sizes of the association between intergroup contact and reduction of prejudice in the meta-analyses of Pettigrew and Tropp range between -.204 and -.225 (Pearson correlation *r*). In social scientific research, such correlation values generally are considered to be small (see Pettigrew & Tropp, 2006: 756, 757). Second, there is no real evidence for a long-term impact, that is, it is uncertain whether positive results of the contact (e.g., improved relations or reduced stereotypes) extend beyond the immediate aftermath of the intervention. Third, positive effects of intergroup contact cannot automatically be generalised to interethnic contact between adolescents, especially in the school environment. A review by Lindo (2008) shows that many empirical studies in this area report mixed results in prejudice reduction between minority and majority groups (see, for example, Cook, 1985; Eshel & Dicker, 1995; McClenehan et al., 1996; Molina & Wittig, 2006). Some studies even report no significant results or a *negative* effect of the contact (see, for example, Bullock 1976;

Vornberg & Grant, 1976; Wagner et al., 1989; Bradnum et al., 1993). A final point of reservation is that the majority of the studies that do report benefits of interethnic youth contact are unable to pinpoint the causal direction of the determined effect. Few investigations in this area have a design based on which a judgement can be made about causation. In other words, it remains largely unclear whether the interethnic contact between young people causes a more positive image about other ethnic groups, or whether a positive image was already present, which in turn established the interethnic contact.

Self-esteem enhancement
Many anti-radicalisation interventions are geared towards enhancing the self-esteem of mentally vulnerable adolescents. These types of interventions usually provide individual counselling programmes or empowerment training for groups of adolescents from minority ethnic groups. In the Netherlands, self-esteem enhancement is particularly used for preventing radicalisation among Muslim youth. The dominant assumption of Dutch government officials and policy makers is that immigrant Muslim youth often struggle with their identity. As a result they are assumed to be more vulnerable to radical thoughts or influence by radical people (see, for example, Dutch Home Office, 2008; SMN, 2008). The ideal-typical assumptions and causal pathway of self-esteem enhancement are provided in Figure 4.

Figure 4. Ideal typical assumptions and causal pathways Self-esteem enhancement
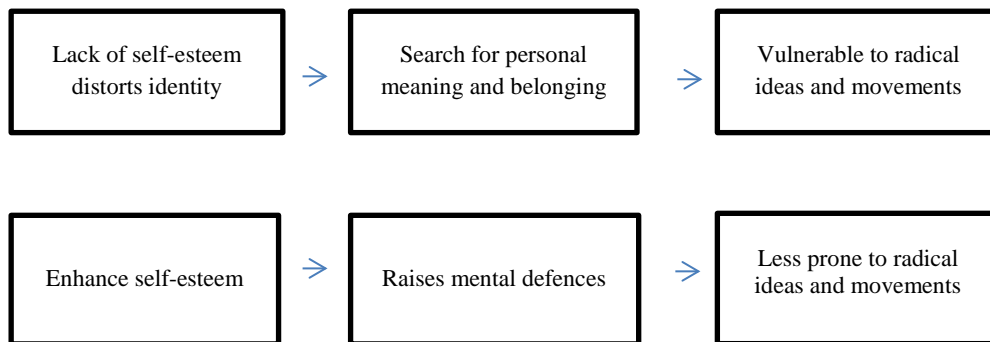
Several authors highlight a link between radicalisation and the formation of identity and self-image, a matter that is supposedly especially apparent during adolescence (see, for example, Vertovec & Rogers, 1999; Phalet, 2003; Heitmeyer & Sitzer, 2008). Some argue that developing a positive self-image is important for developing self-esteem (e.g., Harter, 2003) and, in the case of immigrant youth, a sense of social belonging and

meaning (e.g., Heitmeyer et al., 1997). The assumption of Western-European governments is therefore that young people are vulnerable in this respect, and in their quest for personal meaning have a greater chance of coming into contact with radical movements because these movements offer an ideology that responds to their life questions.

The theory, however, shows problematic premises. First, the indefinable nature of the target group stands out. Which Muslim youth can be labelled as mentally 'vulnerable'? And why is mental vulnerability (and its antonym: mental resilience) primarily associated with Islamic radicalism? Among native 'white' youth, generally no greater receptiveness to right-wing extremism is assumed in the case of a disruption of the self-image or low self-esteem. The more fundamental question of whether young people with low self-esteem – Muslim or not Muslim – are more susceptible to radicalisation or more prone to join radical movements is difficult to answer. The scientific research is inconclusive about the supposed benefits of self-esteem enhancement to socially desirable behaviour. In fact, there is substantial evidence that interventions that unilaterally aim to boost self-esteem encourage narcissism, aggressiveness or antisocial behaviour (see, for example, Baumeister et al., 1996; Heatherton & Vohs, 2000; Stucke & Sporer, 2002; Donnellan et al., 2005). Moreover, several studies report that young people with relatively high levels of self-esteem show greater in-group bias (Crocker et al., 1987; Verkuyten, 1996; Aberson et al., 2000). Young people with a higher self-confidence tend to value members of their own ethnic group more positively than members of the out-group. Boosting confidence levels of (Muslim) youth, therefore, cannot be considered a self-evident mechanism against extremism, radical thoughts or influence by radical people.

3.5. Conclusions

Table 6 provides an overview of the scientific basis of the evaluated intervention theories. It illustrates that none of the investigated policy approaches is supported by a strong scientific foundation.

Table 6. The scientific basis of social policy theories in the field of polarisation and radicalisation

Theory of change	Scientific basis				
	Weak	Moderate	Strong	*Absence of Evidence*	*Evidence of Absence*
Social ecological approach		X			
Peer mediation	X			X	
Intergroup contact		X			
Self-esteem enhancement	X				X

Social ecological approaches seem to be effective in reducing extremist behaviour in radical youth, but do not necessarily lead to an ideological change. Therefore, the scientific basis of the theory is moderate. Mediation by peers seems to exert a positive influence in conflict resolution in most cases, but the causal link between the employment of young people and positive outcomes remains unclear. Despite the many studies reporting positive effects of peer mediation, the theory in itself has a weak scientific basis. However, in this case, 'absence of evidence' is not 'evidence of absence'. As yet, there are no experimental studies available that systematically compare similar mediation methods with different mediators (peers versus adults). Perhaps such research designs will indeed confirm that peers are the beneficial causal factor. But for now such an assumption is impossible to verify conclusively. Intergroup contact on average reduces prejudices about other groups, but effect sizes are generally small and there is no evidence for a long-term impact. Positive outcomes of intergroup contact also cannot automatically be generalised to interethnic contact between adolescents. The theory of change underlying intergroup contact must therefore be judged as moderate.

Finally, the theory that enhancing the self-esteem of at-risk youth will prevent radicalisation has a weak scientific basis. It is doubtful that self-esteem enhancement

makes (Muslim) youth more resilient against radical tendencies. The scientific evidence is ambiguous about whether increasing confidence levels results in socially desirable behaviour or improved social relations. In fact, most studies report negative behavioural effects as a result of one-sided mental empowerment methods. Boosting levels of self-confidence can even contribute to greater in-group bias.

The evaluation of the various theories of change yields several policy implications. Regarding the social ecological approach, the question is justified as to whether radicals who withdraw from terrorist groups and renounce violence, should also be persuaded to give up their radical beliefs. This also depends on the goal of the intervention (is the objective solely a behavioural change or also a mental change?). Still, the radical ideas of young people are not a negligible factor. Exemplary are the so-called 'London bombers', responsible for the terrorist suicide attacks in the London Underground in 2005. The four British young men of Pakistani and Jamaican descent studied or worked and were members of the local cricket club. One of them was even married and had recently become a father. Yet they were sufficiently influenced by radical messages to commit a terrorist attack. In terms of effectuating mental change, a purely social ecological approach in the treatment of radicalising youth might therefore fall short.

Encouraging responsible behaviour by family members of radical young people is a crucial success factor in establishing a mental change through social ecological approaches. Dutch case studies show that parents especially can have a negative influence on the thinking of young people prone to radical behaviour, such as when they ventilate stereotypes about other ethnic groups or propagate a message of intolerance (see, for example, Cadat & Engbersen, 2006; Pels & Vollebergh, 2006; Van den Brink, 2006). This applies both to young radical Muslims and extremist rightwing youth. In some immigrant families, parents – either consciously or subconsciously – convey an anti-Western sentiment or justify similar behaviour from their children (see, for example, Werdmölder, 2005; Van San, 2006). The parents of native Dutch, right-wing extremist youth often share the xenophobic ideas of their children (Cadat & Engbersen, 2006). The intention of the Dutch government to focus more on the role of parents in order to prevent the radicalisation of young people therefore seems legitimate (see Dutch Home Office Polarization and Radicalization Action Plan 2007–11).

In peer mediation, it is recommended to be cautious when it comes to conflict resolution in cultural clashes in disadvantaged neighbourhoods. Adult mediators should not be replaced by peers entirely. Still little is known about whether peers are indeed the causal factor in successful school conflict resolutions, let alone if they can have a positive impact in (ethnic) clashes outside the school environment, such as the neighbourhood. The safety of mediating young people cannot be guaranteed in all

conflicts, and it is conceivable that peer mediators will not always be able to make a constructive contribution in serious ethnic group collisions. Case studies show that conflict gradations can determine the outcome of peer mediation significantly (see, for example, Cohen, 1999). In the later stages of a conflict in which both sides make a hardening stand, the likelihood that informal mediation offers solace diminishes. If peer mediation is considered in contexts outside the school environment, it is therefore more likely to contribute in conflicts that are in their early stages and have not yet escalated.

In intergroup contact, reducing feelings of uncertainty about the contact with members of groups that are perceived as different, turns out to be an important precondition for contact situations (see, for example, Blascovich et al., 2000; Mendes et al., 2002; Shelton & Richeson, 2005, 2006). When this condition is not met during introductory activities, it could potentially result in further alienation. This might explain some of the mixed results of interethnic contact reported by some studies. Naturally, this finding is at odds with organising random familiarisation sessions for young people in order to reduce (potential) prejudice, and signifies the importance of thorough professional supervision. The present evaluation of the empirical research on intergroup contact also sheds light on other conditions that can optimise its effect. It is equally desirable in intergroup contact to:

- work on the basis of concrete examples and scenarios of stereotyping (Gurin et al, 1999; Finlay & Stephan, 2000);
- implement role plays in which young people participate in 'perspective taking' (Stephan & Finlay, 1999);
- ensure an open discussion environment whereby moralisation on behalf of the trainers is avoided and xenophobic youth are allowed to express themselves (Cameron et al., 2006; Bekhuis et al., 2009);
- integrate cooperative learning techniques in which students of different ethnic backgrounds work together to solve a problem (Cooper & Slavin, 2004; Roseth et al., 2008).

Finally, the assessment of self-esteem interventions implies that for young people already in the process of radicalisation, it seems useful to buffer their self-image through individual programmes rather than to increase it directly through 'empowerment' sessions. This 'buffering' entails personal values of young people being reinforced, which may attenuate aggressive reactions after the undermining of the self-image. Dutch trajectories through which radicalised Muslim youth learn to act not only religiously, but also morally, seem promising in this respect (see Gielen, 2009). Experimental research suggests that such methods remove the sharp edges of

dogmatic views and may increase tolerance towards dissenters (see Cohen et al., 2000; Thomaes et al., 2009).

3.6. Discussion

This chapter provides insight into what plausibly works and does not work against polarisation and radicalisation issues, without costly and time-consuming evaluations needing to be conducted of separate interventions and programmes. But what works in theory does not always correspond with what works in practice. With regard to polarisation and radicalisation, in particular, a few significant epistemological challenges should be considered.

First, one can have a detailed notion of what factors influence radical or xenophobic behaviour of young people, but this does not mean that those insights can be applied automatically in newly designed policy interventions. The studies on which the conclusions of this article are based were often very detailed and differentiated many subtle variations in behaviour and psychological dispositions. This complicates the objective of applying the particular intervention that is consistent with the 'diagnosis' of the (radical) behaviour, that is, if that diagnosis can even be made accurately. For example, it remains unclear which Muslim youth can or should be labelled as 'radical'. Some Dutch cases painfully illustrate the difficult assessment of the often subtle differences between religious orthodoxy and supposed radicalism of some Muslim youth (see Demant et al., 2008: 58–61). Also, it is one thing to conclude that, for example, the moral education of parents influences xenophobic ideas of atrisk youth; to effectively incorporate such notions in government policy is another matter. And most of the studies on intergroup contact have been conducted in the psychological laboratory. But how might we bring people together under the right conditions in everyday life (see also Wilson, 2011)? These are difficult questions and insight into the possibilities of social policy can thus paradoxically also lead to an understanding of its limitations.

Second, one can question whether it is possible to make statements about the effectiveness of existing approaches solely by examining the literature. The method provides a reference point of what works, but not a recipe. From the literature, insights emerge about effective elements based on empirical findings. Whether such potential effectiveness takes shape in practice depends on a variety of social, political and administrative constraints and contextual factors (see, for example, Pawson & Tilley, 1997; Spicker, 2011). However helpful, literature reviews by no means yield unequivocal policy verdicts (Boaz & Pawson, 2005; see also Chambers et al., 2012). The viewpoint of confronting causal assumptions with scientific evidence from the

literature can never paint a complete picture of the efforts required to address issues such as extremism and ethnic tensions. Moreover, no intervention occurs in a social vacuum. Practitioners in the field often have their own interpretation of what mechanisms ought to contribute to specified goals. The interpretations of these 'street-level bureaucrats' (Lipsky, 1980) can deviate from problem definitions and goal specifications put forth in official documents. Practitioners may narrow the goals of policy makers or elaborate upon them, or substitute entirely different goals. Such a transformation can produce a gap between the (evidence-based) design of a particular programme and its actual implementation.

Finally, although a theory-driven assessment evaluates causal assumptions of interventions and thus indirectly problematises the legitimacy of the interventions themselves, only underlying theories are tested that are considered relevant for policy analysis. This is because existing government-funded interventions constitute the starting point of inquiry. The risk, however, is that factors that may work against polarisation or radicalisation, but fall outside the analytical framework, are overlooked. While theory-driven evaluation enables a peek inside the famous 'black box' (Astbury & Leeuw, 2010), it does not peer outside of it. In sociological systems theory, this phenomenon is known as the problem of the 'blind spot' (Luhmann, 1997). When an observer chooses a certain perspective, they are inevitably confined by this perspective. Only when the observer takes a step back to look at their own point of view, can the consequences of their original choices be critically brought to light. When it comes to tackling polarisation and radicalisation issues, then, policy analysts and researchers should not only evaluate present measures but also be willing to ask radically different questions about what could contribute to solving the problems.

The most immediate question is at what point to take into account the social context. This depends in great measure on the conceptualisation of 'polarisation' or 'radicalisation'. If these issues are conceived as treatable pathologies (as the four approaches discussed in this article tend to do), then it makes sense to focus on treatment plans (ie, 'interventions'). However, such an approach could arguably be dismissed as naive, when no account is offered of sociopolitical factors or perceived injustices fuelling radicalisation or ethnic tensions. As mentioned, the 'London Bombers' are an exemplary case. It is essential to investigate further to what extent contextual variables are essential components of interventions seeking a lasting reduction of these problems. Should, for example, in the case of social ecological approaches for Muslim youth, Imams and mosques be seen as desirable 'social attachments' or as potential detractors from mainstream society? In what way do migration factors influence the efficacy of individual counselling programmes for minority ethnic youth? And how strongly do geopolitical developments that influence domestic relations between ethnic groups hinder or accommodate endeavours to

promote mutual understanding through intergroup contact? The interaction between government-induced interventions against polarisation and radicalisation and wider sociopolitical factors is a relevant topic for future research.

4. The softer side of urban regeneration

4.1. Introduction[8]

In addition to physical measures such as housing improvement and stricter law enforcement, deprived neighbourhoods throughout the Western world display many social interventions to tackle disorder and crime. In the United Kingdom, the United States and France, for instance, a lot of effort has been put into strengthening social ties and resident participation in deprived neighbourhoods (see e.g. Hall & Hickman, 2011; Sampson, 2011). Other internationally popular programs include redirecting at-risk youth away from loitering and petty crime by introducing them to constructive leisure activities, such as sports (see Carmichael, 2008). In the Netherlands, in particular, the government invested millions of euros in various community programs under the common name of the *Wijkaanpak*, comparable to the U.K's nationwide urban regeneration program New Deal for Communities. Dutch programs include resident participation and watch schemes, neighbourhood sports programs, dialogue techniques to improve resident relations and formulate desired codes of conduct, and many projects on social cohesion.

Although physical interventions and – to a lesser extent - law enforcement strategies usually are well monitored and evaluated (Stephens, 2011; Sherman, 2011), place-based *social* policies in urban areas are usually less thoroughly investigated. As far as the softer side of Dutch urban regeneration is concerned, it often seems like 'anything goes'. Part of the reason for their limited scrutiny, has to do with the specific properties of such policies. Place-based social programs typically have no sharp demarcation in time, intensity or target group and are implemented in a rich context of unforeseen and unknown geographical and situational variables. This makes it difficult to assess their impact with conventional research methods.

However, despite the multitude of urban projects and their unclear demarcations, place-based social interventions are rooted in some basic principles and assumptions on how they should contribute to quality of life. The validity of these broader principles can be evaluated and reviewed on the basis of scientific insights (see also Lub, 2013a). Using the Netherlands as a case study, this paper puts prevailing assumptions of social policies against neighbourhood disorder to the test, by

---

[8] This chapter represents a selection of findings from the book *Schoon, heel en werkzaam? Een wetenschappelijke beoordeling van sociale interventies op het terrein van buurtleefbaarheid.* Lub, V. (2013). Den Haag: Boom Lemma.

confronting them with the scientific literature. For this purpose, a theory-driven evaluation perspective was used. The defining principle of any theory-driven approach is that - instead of measuring outcomes empirically - assumptions of social interventions are made explicit and are scrutinised for their validity (Chen, 1990; Shadish, 1987; Lipsey, 1993; Weiss, 1995; Pawson, 2006). Each social intervention is based explicitly or implicitly on certain assumptions about what effect it should yield and what causal mechanisms should bring about the intended change. Put together, these assumed causal pathways constitute a 'theory of change', which can be confronted with insights from the scientific literature (Weiss, 1995).

4.2. Methodology

The program theories were reconstructed by coding-analysis of written documents, such as government policy papers, descriptions of quality of life programs from professional and civic organizations and strategy outlines of funding agencies. First, an inventory was made of all active Dutch social programs designed to tackle neighbourhood disorder (ranging into the dozens). To this end, both national and local government databases were searched as well as various other databases containing descriptions of neighbourhood programs. The respective programs were then lumped into more general themes and categories of which subsequently the most defining underlying principle of a particular policy category along with the postulated relationships between the assumed 'working' components was extracted. For instance, the inventory yielded several neighbourhood sports programs. Though different in scope and approach, many started from the same principles and displayed very similar assumptions on how they should contribute to quality of life issues. In other words, generalizations (ideal types) of the various policy themes and their underlying theories were constructed, which then served as a directive for the literature assessment.

   To review the theories of change, the international literature was systematically searched, using strings of key words. The scientific resources and impact assessments had to be related to one of the six themes and the assumed underlying (causal) assumptions. The search yielded various academic reviews (meta-studies and narrative reviews) as well as academic single case studies (experimental, longitudinal, surveys and qualitative investigations) and relevant policy studies and evaluations. No restrictions were made in terms of publication date and size of the study. The literature yielded circa 100 relevant international studies and impact assessments[9]. Whether the identified studies are representative of all studies conducted

---

[9] See for an overview of the consulted literature: Lub, 2013b

in this area is not known. Not all research useful to the study may have been published. We assume, however, that the literature we collected contains the most relevant insights, partly because during the research, academic experts were consulted on the various topics.

4.3.   Conceptualizations and research questions

The present study focuses on social programs that aim to improve the *quality of life of the public space* in deprived neighbourhoods, i.e. matters that relate to anti-social behaviour, a run-down physical environment and street crime and violence (Wilson & Kelling, 2003; Sampson, 2009; Skogan, 1990). Social programs that target other problematic neighbourhood indicators, such as health issues, poverty, long-term unemployment or the position of marginalised groups (such as migrants) are not part of the investigation. The policy approaches central to this paper thus cover the most dominant social programs in the Netherlands aimed at disorder. They include: promoting neighbourhood contacts (I), resident representation and participation (II) community-based behavioural codes programs (III), neighbourhood watch schemes (IV), and neighbourhood sports (V).

- *Neighbourhood contacts:* Interventions that aim to promote contacts between residents are a frequent appearance in deprived neighbourhoods, in the Netherlands and beyond. Examples include block parties, neighbourhood theatres, street parades, kitchen table dialogues and 'coffee mornings' in primary schools, community centres, parks or playgrounds. The most fundamental assumption is that such interventions strengthen the social fabric of a neighbourhood. A cohesive neighbourhood is supposed to increase social control, giving anti-social behaviour and physical degradation less chance. It is also assumed that the strengthening of social networks is a prerequisite for the willingness of residents to become active in their immediate living environment.
- *Resident representation and participation:* Like promoting social contacts, resident participation and representation in local government is a reoccurring theme in European and North-American urban regeneration policy (Baker & Evans, 2002; Clark & Southern, 2006). In the Netherlands, many community platforms are founded specifically to address quality of life issues. The active residents form an interest group that interlock with public bodies. Resident groups focus on transferring information as well as set priorities for local services.

- *Community-based behavioural codes programs:* These programs aim to collectively formulate and enforce agreements between residents about the desired behaviour in a neighbourhood or street. Residents meet in joint beat meetings and – assisted by community workers – formulate and finalize a set of rules by which residents need to abide. The premise of such programs (sometimes called 'street etiquette') is that public familiarity between residents is insufficient and that this in turn causes a negative spiral of social control, and subsequently more disorder.

- *Neighbourhood watch schemes:* Many urban areas in Europe and the U.S. employ neighbourhood watch schemes to actively counter disorder and petty crime. Different assumptions underpin the neighbourhood watch, such as direct intervention of deviant behaviour and an increase of information flow to the police (see Bennett et al. 2006: 438). In the Netherlands, the main purpose of the teams is to increase the level of informal social control.

- *Neighbourhood sports:* Throughout the western world, local community workers organise various types of public sports activities. Providing opportunities for sport in deprived neighbourhoods can have many objectives, such as improved health and well-being. But involvement in sport is also seen as an effective method for positively involving 'at-risk youth' to the neighbourhood. The idea is that sports fosters pro-social behaviour and mitigates antisocial and criminal tendencies by offering an alternative social environment where young people learn positive values and norms (cf. Schafer, 1969). In time this improved behaviour should also benefit the quality of life in the neighbourhood.

It is important to note that a particular 'theory of change' of a given intervention does not necessarily refer to a social theory in the general academic sense. In most cases, it comprises a set of causal assumptions and mechanisms (pathways), which are implicitly bound up in a given policy intervention or social programme (Weiss, 1995). However, some theories of the respective policy approaches central to this article are reminiscent of conventional urban theory. Underlying assumptions of the neighbourhood contact approach, neighbourhood watch schemes and active citizenship are rooted in the social disorganization theory, first articulated by members of the Chicago School (see e.g. Shaw & McKay, 1942). It argues that differences in crime and disorder between neighbourhoods can partly be related to social ecological characteristics, such as the level of social organization among residents of those neighbourhoods. Similar urban theoretical notions apply for resident representation and participation schemes. Sampson et al. (1997) hypothesized that collective efficacy, the willingness of residents to intervene on behalf of the common good, is linked to

reduced urban violence. Government appeals to act on behalf of the common good and interlocking resident and professional bodies by means of resident representation schemes thus seem obvious approaches to tackle disorder. Finally, there are similarities between neighbourhood sports and the criminological social bond theory. The social bond theory emphasises an absence of social attachments among at risk youth and juvenile delinquents, whereby commitment to positive social norms and institutions (school, employment) and involvement in mainstream activities are believed to be important starting points for behavioural change (see Hirschi, 1969: 16; and for the link with sport Schafer, 1969).

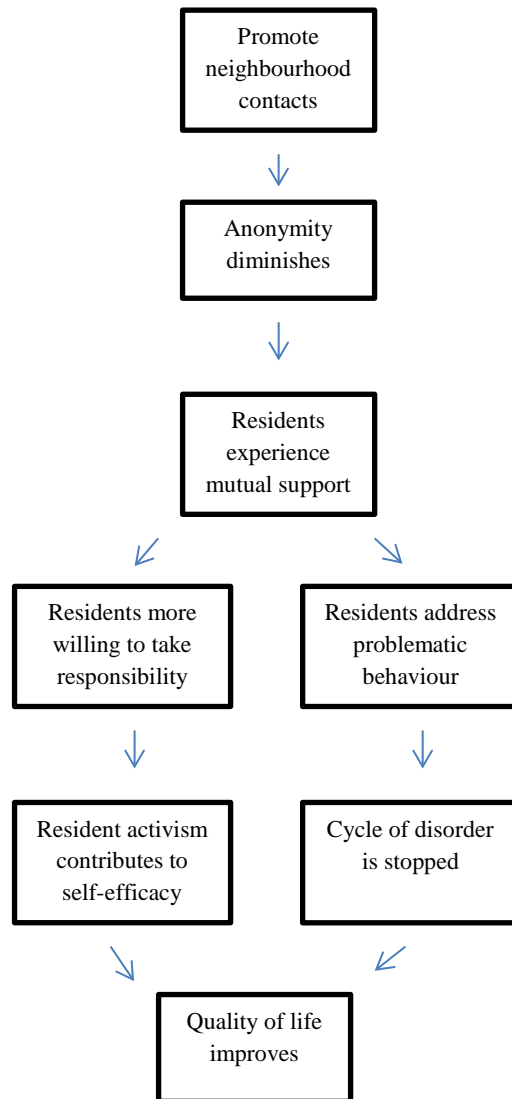In sum, the research questions of this chapter read as follows:

- How plausible are underlying causal mechanisms implicit in neighbourhood-based social policies, given scientific knowledge?
- What particular assumptions can be considered plausible or implausible according to the scientific literature?
- What implications does the knowledge gained have for future urban policy and theory?

4.4.   Results

Neighbourhood contacts
Interventions that aim to promote contacts between residents are a frequent appearance in deprived neighbourhoods. Examples include block parties, neighbourhood theatres, street parades, kitchen table dialogues and coffee mornings in primary schools, community centres, parks or playgrounds. The most fundamental assumption is that such interventions strengthen the social fabric of a neighbourhood. The policy assumption is that a cohesive neighbourhood increases social control, giving anti-social behaviour and physical degradation less chance. Second, it is often assumed that the strengthening of social networks is a prerequisite for the willingness of residents to become active in their neighbourhood: resident activism as a spin-off of social cohesion. The ideal-typical assumptions and causal pathways of the neighbourhood contacts approach are provided in Figure 5.

Figure 5. Ideal typical assumptions and causal pathways of the neighbourhood contacts approach

```
                    ┌─────────────────┐
                    │     Promote     │
                    │  neighbourhood  │
                    │    contacts     │
                    └─────────────────┘
                             ↓
                    ┌─────────────────┐
                    │    Anonymity    │
                    │   diminishes    │
                    └─────────────────┘
                             ↓
                    ┌─────────────────┐
                    │    Residents    │
                    │    experience   │
                    │  mutual support │
                    └─────────────────┘
                      ↙             ↘
         ┌─────────────────┐   ┌─────────────────┐
         │  Residents more │   │ Residents address│
         │  willing to take│   │    problematic   │
         │  responsibility │   │    behaviour     │
         └─────────────────┘   └─────────────────┘
                  ↓                     ↓
         ┌─────────────────┐   ┌─────────────────┐
         │ Resident activism│  │ Cycle of disorder│
         │  contributes to │   │   is stopped     │
         │  self-efficacy  │   │                  │
         └─────────────────┘   └─────────────────┘
                  ↘                   ↙
                    ┌─────────────────┐
                    │  Quality of life│
                    │    improves     │
                    └─────────────────┘
```

Despite its popularity, there is no scientific evidence that decisively shows that more frequent or closer contacts between residents in disadvantaged neighbourhoods produce more social control and thereby exert a positive influence on the neighbourhood quality of life. The literature reveals that these issues certainly are interrelated (see e.g. Pratt & Cullen, 2005), but a firm *causal* relationship has never

been established, and several studies show contradictory results (Morenoff et al., 2001, Greenberg et al., 1985; Browning et al., 2004).

A strong social infrastructure might be more important than social cohesion. It is well documented that the number of public institutions, voluntary organizations and the number of inhabitants who are represented in neighbourhood associations, contribute to the social control capacity in the neighbourhood, in a way that exceeds the traditional viewpoint of social cohesion, i.e. resident contacts. Silver and Miller (2004: 569), for instance, conclude that social bonds in disadvantaged neighbourhoods hardly affect the willingness of residents to exert informal social control over youth. Satisfaction with police on the other hand, does lead to significantly higher levels of informal social control. Surveys by Morenoff et al. (2001) support these findings. Their research in Chicago shows that the density of local community organizations and the involvement of residents in active residents groups (voluntary associations and organizations), significantly contributes to a higher level of collective self-efficacy. Browning et al. (2004) even found that close ties between residents undermine the positive effect of collective self-efficacy in the neighbourhood, with denser ties leading to more crime. Qualitative research of Small (2009) uncovers that neighbourhood organizations are equipped to lead a successful process of collective self-efficacy by using their organizational capacity and strategic networking, which in turn sets professional or collective action in motion. In sum, communities can exhibit intense private ties (e.g., among friends, kin, neighbours), and perhaps even shared expectations for control, yet still lack the institutional capacity that is required to achieve social control (Hunter, 1985).

According to Sampson (2011: 219), such findings indicate that the effectiveness of social control in the neighbourhood depends primarily on organizational and institutional conditions and the number of contacts between residents groups and professional agencies within and outside the neighbourhood, and that this dynamic need not reflect the number of contacts *among* all residents. The literature thus raises questions about the validity of the numerous undirected resident contacts initiatives aimed at improving quality of life.

The literature also suggests that the strengthening of social networks does not constitute a strict requirement for resident activism. Research from the Unites States shows that the presence of active resident groups and backing of representatives of local public institutions, such as police, housing services and community workers (i.e. a strong social infrastructure), prove more important factors for people to address others on deviant behaviour than social cohesion itself (Silver & Miller, 2004; Skogan, 2006). Surveys and qualitative studies in the Netherlands show similar findings and indicate that personal motives, individual competencies (such as social and conflict resolution skills) and the nature of local disorder and crime problems are more

influential in the readiness for active civic engagement and social control than social cohesion (Lelieveldt & Van der Kolk, 2005; Van Marissing, 2008; Kleinhans & Bolt, 2011).

Resident representation and participation
Like promoting social contacts, resident participation and representation in local government is a reoccurring theme in European and North-American urban regeneration policy (Baker & Evans, 2002; Clark & Southern, 2006). Dutch disadvantaged neighbourhoods are no exception and often display a range of different community platforms. Many are founded specifically to address quality of life issues, such as dilapidated housing blocks, litter, vandalism, (youth) disturbance and petty crime. The active residents form an interest group that interlock with public bodies. Resident groups that deal with problems of disorder first and foremost focus on transferring information. The civil boards provide the police and city services with information about signs of decay in the neighbourhood or illegal activities. The policy assumption is that resident's input enables a more effective approach by law enforcement and other services in tackling everyday disorder problems. The input of residents is therefore also supposed to help set priorities for neighbourhood services. The ideal typical assumptions and causal pathways of resident representation and participation are provided in Figure 6.

Figure 6. Ideal typical assumptions and causal pathways of resident representation and participation
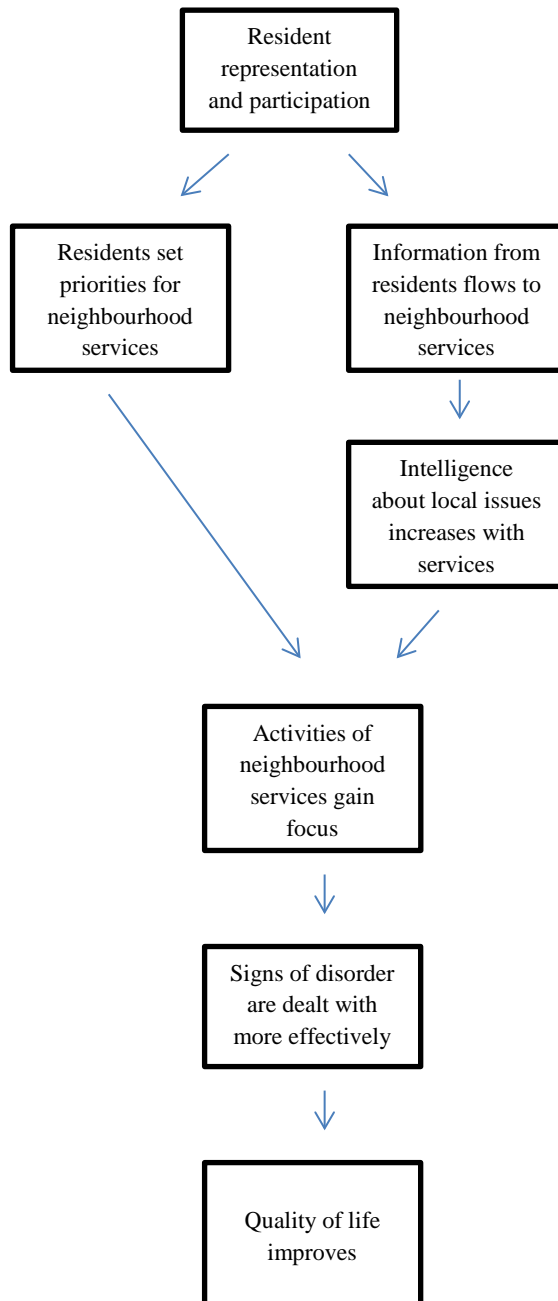
The plausibility of resident participation in neighbourhood safety policy is difficult to determine. Rigorous research in this area is scarce. From what is scientifically known, however, the contribution of resident platforms to resolving quality of life issues seems limited. Quasi experimental evaluations by Wycoff and Skogan (1993) and Skogan and Hartnett (1997) of community policing programs and community schemes cooperating with local services, show mixed results. Sometimes the interaction with government institutions improved the perception of neighbourhood safety with residents, but effect sizes are small. Moreover, actual improvements of objective levels of crime and decay were rarely recorded. Qualitative studies from the Netherlands indicate that often conflicting logics between the resident panels and professional bodies arise (Lub et al., 2009; Boutellier and Van Marissing, 2011). Residents and professionals often have different interpretations of local quality of life issues and rarely speak the same language. This may be a possible explanation for the scant results of the community policing programs in the United States. Another possible determinant for conflicting logics might stem from diverging interpretations of neighbourhood problems between residents themselves. Both quantitative and qualitative studies in Dutch and Flemish neighbourhoods show that people - according to their demographic and biographical background – often view safety issues and signs of disorder in the public space very differently (De Leeuw & Van Swaaningen, 2011; Müller, 2011; Damen, 2010; Hardyns et al., 2010). This is corroborated by Skogan's assessment of the community policing program CAPS[10] in Chicago. His research showed that active residents estimate indicators of physical and social disorder and forms of petty crime considerably more problematic than their non-active neighbours (see Skogan, 2006: 157). CAPS-participants were more concerned about neighbourhood problems than residents who did not participate in the beat meetings. Although a group of citizens may contribute to a more effective policy for safety in a neighbourhood, such differences in viewpoints make an effective contribution of citizens through consultative bodies to local safety difficult, as they are often unable to speak with a single voice.
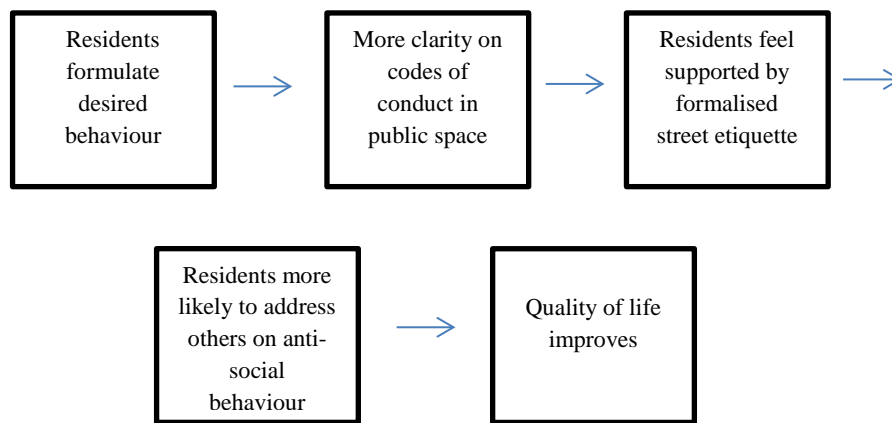
Community-based behavioural codes programs
A unique Dutch measure against neighbourhood disorder is community-based behavioural codes programs. These programs aim to collectively formulate and enforce agreements between residents about the desired behaviour in a neighbourhood or street. Residents meet in joint beat meetings and – assisted by community workers – formulate and finalize a set of rules by which residents need to abide, such as 'In this neighbourhood we greet each other' or 'We don't make noise after 22:00'. The premise

---

[10] Chicago Alternative Police Strategy

of such programs (sometimes called 'street etiquette') is that public familiarity between residents is insufficient and that this in turn causes a negative spiral of social control. Disadvantaged neighbourhoods in particular are assumed to be characterized by a high degree of anonymity and indifference in this respect. Such attitudes would then increase norm-breaking behaviour and ultimately lead to tensions, aggression or even violence. The ideal typical assumptions and causal pathways of community-based behavioural codes programs are provided in Figure 7.

Figure 7. Ideal typical assumptions and causal pathways of behavioural codes programs
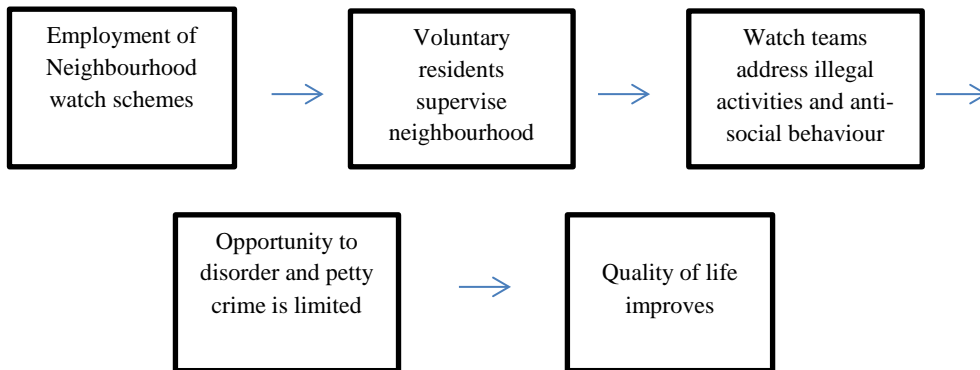

According to the literature, it is unlikely that community-based behavioural codes programs contribute to the quality of life in disadvantaged neighbourhoods in a significant way. The picture that emerges from various Dutch studies of this policy approach is of a small group of residents that strongly focuses on their immediate living environment, and that is attracted to a specific type of 'hygienic' social norm (De Meere et al., 2004; Lub, 2005; Uitermark & Duyvendak, 2006; Diekstra, 2006; Bongers & Langeveld, 2012). The projects that are intended to establish collective norms in this way thus only reach a limited number of people, usually the residents that were already actively involved in the neighbourhood before the project. For these residents, the street agreements are likely to have an additional bonding effect. But it is unlikely that in this way a wider circle of people is reached who are willing to commit to the 'established' rules. Analysis of the social psychological and sociological literature corroborates these findings. The main problem is that social norms are always 'group-specific' (Chekroun, 2008), meaning that for people to conform to them, they have to experience some collective identity. Yet the literature implies that given the

demographic, social and cultural diversity of urban neighbourhoods and the geographical expansion of social networks, a street or neighbourhood hardly constitutes a meaningful collective sphere wherein group norms can be enforced effectively (see e.g. Webber, 1963; Fischer, 1977; Pescosolido & Rubin, 2000). As a result, behaviour that is perceived as 'uncivilized', does not have the same meaning for different social groups. The discursive process between local residents in the establishment of the norms can even cause adverse effects, and a few cases have been recorded of frictions between residents or tensions between residents and local agencies as a result of the programs (see Bongers & Langeveld, 2012).

Neighbourhood Watch schemes
Many urban areas employ neighbourhood watch schemes to exert social control and actively counter disorder and street crime. In the Netherlands, the first active voluntary watch teams date from the late 1980's (Van Noije and Wittebrood, 2008). They derived from similar schemes in the United States and the United Kingdom (U.K.). Bennett *et al*. (2006) estimate that in the U.S. more than 40 percent of urban residents live in a neighbourhood where some form of formal resident surveillance takes place. In the U.K., this applies to more than a quarter of the population. The concrete focus of the teams may vary by city or location, but usually it involves burglary prevention, reduction of drug related disturbance, preventing vandalism and addressing general anti-social behaviour. Different theories underpin the neighbourhood watch (see Bennett 1990; Bennett et al., 2006: 438). The most frequently suggested mechanism by which it is supposed to reduce crime and disorder is by residents looking out for suspicious activities and reporting these to the police. Neighbourhood watch might also reduce the opportunities for deviant behaviour by an increased level of informal social control (cf. Greenberg et al., 1985; Rosenbaum, 1987), which is the prevailing assumption behind the neighbourhood watch in the Netherlands. The direct supervision of residents should enforce 'desired' standards of behaviour. The ideal typical assumptions and causal pathways of the neighbourhood watch schemes are provided in Figure 8.

Figure 8. Ideal typical assumptions and causal pathways of neighbourhood watch schemes

Employment of Neighbourhood watch schemes	→	Voluntary residents supervise neighbourhood	→	Watch teams address illegal activities and anti-social behaviour	→

| Opportunity to disorder and petty crime is limited	→	Quality of life improves

Does neighbourhood watch constitute a plausible contribution to the quality of life in troubled neighbourhoods? From a results standpoint, the signs seem hopeful. A meta-analysis by Bennett et al. (2006) of the efficacy of the Neighbourhood Watch shows that more than half of the 43 evaluations yielded positive results, that is to say, a smaller increase in crime relative to the control area. The statistical calculations of Bennett et al. also show that the average effect size – i.e. the relative impact of the watches on neighbourhood safety - is significant. On average, of the assessments with a positive result, the crime rate in the control areas (where no watch took place) increased by 36 percent relative to the experimental areas (where a neighbourhood watch did take place). This is a very robust effect size and, moreover, a significant difference. Although the meta-analysis exclusively focused on outcomes for street crime and also shows that the size of the effect may vary - in a few cases neighbourhood watch schemes yielded no or negative results - we can at least conclude that neighbourhood watch schemes can establish significant improvements in the quality of life in deprived neighbourhoods. However, under what conditions they work, if social control plays a major role in their effectiveness, and why they sometimes seem to have a negative impact on the perceived safety of residents (see e.g. Skogan, 1990; McConville & Shepherd, 1992) remains unclear. As far as can be derived from the literature, there are no studies available that rely on systematic observations of neighbourhood watches in the field that might shed light on these matters.

Neighbourhood sports
Finally, organizing neighbourhood sports programs is one of the most popular interventions in community development. Throughout the western world, local community workers organise football tournaments, martial arts classes or other types

of public sports activities. Position papers of local and national government bodies in the Netherlands resonate a strong conviction about the social effect of sport, a trend that can also be observed internationally (see Carmichael, 2008). Although sports can have many objectives, here the idea is that sports in itself fosters pro-social behaviour. In this view, sport mitigates antisocial and (potentially) criminal behaviour of at-risk youth by offering an alternative social environment where young people learn positive values and norms (cf. Schafer, 1969). In time, the quality of life in neighbourhoods should also benefit from this, because with the acquired skills learned in the sport context, youngsters would also be more likely to behave appropriately in everyday life. The ideal typical assumptions and causal pathways of neighbourhood sports are provided in Figure 9.

Figure 9. Ideal typical assumptions and causal pathways of neighbourhood sports
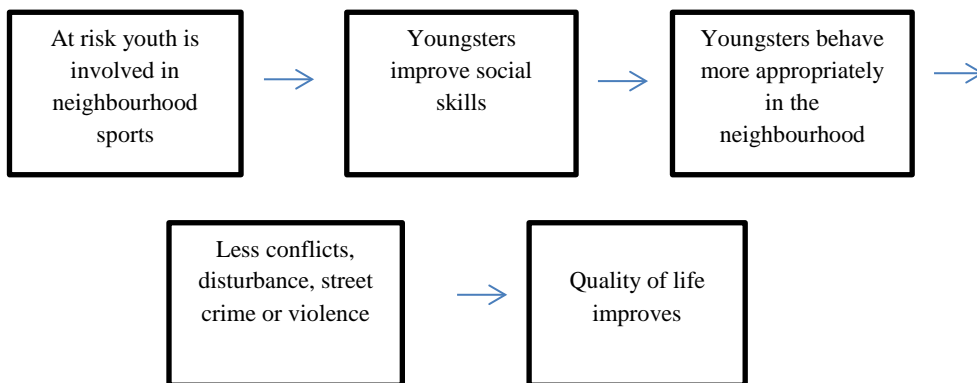
How do these ideals fit in reality? Scientific research shows that sport has a much smaller impact on the behaviour and development of children and adolescents than is generally assumed. A large survey by Steptoe and Butler (1996) showed that the effect sizes of the relationship between sport and (pro) social behaviour of young people, as well as the explained variance of the statistical models, are minor. This means that sport plays a limited role in the behavioural development of boys and girls. The modest behavioural impact of sport is further illustrated by a longitudinal study of Hartmann and Massoglia (2007), who found no direct link between performing sports in youth (regardless of the duration and intensity) and general deviant behaviour in (early) adulthood.

Neighbourhood sports can, however, serve as a pedagogical starting point for at-risk youth. In an experimental study of the promotion of moral development of elementary school children through physical education, Romance et al. (1986) found

significant differences in moral reasoning between the experimental and control group. Children who participated in the educational program geared towards moral reasoning *in conjunction* with physical education showed a higher degree of moral reasoning than kids who received the moral program without sports. In a similar study Gibbons *et al.* (1995) showed that treatment group participants of a moral school curriculum had significantly higher post-test scores on all measures of moral reasoning as compared to students in the control group. Daniels and Leaper (2006) showed that the effects of sport on child behaviour should be viewed in its social context; though the effect of sports itself is modest, it does have a mediating effect on the emotional well-being of adolescents, by offering a setting were peer acceptance can take place.

Sport can thus be of value to at-risk youth in disadvantaged neighbourhoods, but only if it is utilized as a pedagogical vehicle. The alleged intrinsic power of sport itself seems limited. Given its weak overall effects, it is unlikely that through sports alone youngsters will commit to more positive values and norms which they will subsequently apply in their everyday behaviour in the neighbourhood. Scandinavian and American research even suggest that disciplines with a high masculine culture and content such as martial arts and boxing, amplify violent behaviour of young boys (Endresen & Olweus, 2005; Kreager, 2007), though it is unclear whether proper pedagogical supervision can overrule these effects.

4.5. Conclusions

What picture emerges when we assess the scientific plausibility of prevalent social policy assumptions in the Netherlands against neighbourhood disorder? Table 7 provides an overview. Note that the study does not go so far as to falsify or confirm the theories of change, but instead labels their scientific basis as strong, moderate or weak. A strong scientific basis means that causal relationships between the assumed working components are largely confirmed by scientific insights, with considerable effect sizes. A moderate scientific basis can either refer to research results only partly confirming the assumed working components or to a lack of scientific insights required to assess the theory (absence of evidence instead of evidence of absence). When a theory of change has a weak scientific basis, this means that policy assumptions about what works and why can be confidently dismissed on the basis of scientific insights.

Of the evaluated theories, one can currently boast a relatively strong scientific basis: the assumption that neighbourhood watch schemes contribute positively to neighbourhood quality of life. This type of intervention is therefore promising, albeit that the understanding of the processes and conditions surrounding the functioning of the neighbourhood watch is still underdeveloped. For policy assumptions about the

contribution of resident representation to resolve disorder issues, the scientific basis can be labelled as 'moderate'. The empirical evidence in this area is not immediately dismissive, but it is scarce and results that are available, are modest. Three of the examined theories have a weak scientific basis. Given the current body of knowledge, policy assumptions in the field of contact stimulation, the collective shaping of behavioural codes of conduct and sports programs will probably contribute little to the quality of life in disadvantaged neighbourhoods, though sport can be used as a vehicle or 'gateway' for alternative behavioural change programs for at-risk youth.

Table 7. The scientific basis of place-based social interventions against neighbourhood disorder

Policy approach	Theory of change (assumed working components)	Scientific basis		
		Weak	Moderate	Strong
Neighbourhood contacts	Cohesion increases social control and activism	X		
Resident representation and participation	Residents can provide appropriate input for local agencies		X	
Behavioural codes programs	Residents can develop shared perspective on desired behaviour	X		
Neighbourhood watch schemes	Intra-communal social control in hotspots			X
Neighbourhood sports	Sport fosters pro-social behaviour	X		

The assessment of the various policy theories yields various implications for both urban policy and theory. In the case of promoting neighbourhood contacts, the study relativizes the importance given to initiatives aimed at strengthening resident networks to tackle disorder. The social disorganization theory of the Chicago School should perhaps not be understood in terms of relations *between residents,* as it is often perceived by policy makers and administrators, but to the degree of *organizational capacity* that some of the residents and professionals can collectively bring about in troubled districts. As shown, the extent to which social control takes place in

disadvantaged areas greatly depends on the strength of the professional infrastructure and the quality of the contacts between community groups and neighbourhood organizations. These institutional dynamics need not necessarily reflect the number of contacts among residents (cf. Sampson, 2011). In the context of quality of life policy it is therefore desirable to shift from *horizontal* to *vertical* relationships (cf. Bursik & Grasmick, 1993). The priority must lie on promoting, equipping and guiding resident activism through functional and energetic professional bodies, not on the unfocused encouragement of resident contact. A strengthening of these vertical ties (extra-local resources) will make it more likely that residents will deploy the necessary political and administrative capital, as well as utilize police ties more effectively. That outside help will in turn increase the likelihood that local disorder problems can be solved more efficiently.

Subsequently, the study raises doubts about the importance that is often attributed to *informal* social control or 'natural surveillance' in order to tackle signs of disorder. The question is: are the causal arrows of the policy assumptions facing the right direction? Do deprived neighbourhoods display a high level of disorder and crime because of low levels of social control? Or do deprived neighbourhoods display a low level of social control because they are disorderly and crime ridden? Given the complexity of neighbourhood processes, such questions might never be answered definitively, and the relation might be reciprocal. But particularly in Europe, policy makers often use the first line of reasoning as a justification for countless social cohesion and dialogue projects. The fact that the neighbourhood watch is the only intervention that can boast a rather strong scientific basis and the insight that trust in local institutions leads to significantly higher levels of informal social control, are signs that social control is perhaps better enforced through more formal surveillance methods and a process of professional backing.

Similar problems can be observed in the instigation of the behavioural codes programs and sports. Though social identity theory states that a significant portion of an individual's behaviour derives form perceived membership of a relevant social group, this notion does not automatically imply that residents in deprived neighbourhoods can create this sense of collective membership (as is the premise of behavioural codes programs). Given the demographic, social and cultural diversity of urban neighbourhoods and the geographical expansion of social networks, a street or neighbourhood hardly constitutes a meaningful collective sphere wherein group norms can be enforced effectively. And although effect studies show that social ecological interventions that involve mainstream institutions are effective against socially problematic behaviour of young people and generally yield better results than alternative treatments (Curtis et al., 2004; Borduin et al., 2009), sports seems to be too weak a mechanism to be placed in this category. Moreover, given the negative effects

of 'power' sports, administrators should perhaps be cautious with promoting martial art lessons for youngsters in disadvantaged neighbourhoods, as is popular in the Netherlands and other European countries.

4.6. Discussion

The present study promps a debate on *instrumental policy* (policy as a means to an end) versus *symbolic policy* (policy as communicative strategy). This paper was mainly concerned with instrumental policy, and from that perspective, the literature questions the plausibility of several social measures against neighbourhood disorder and crime. But perhaps this is not the 'real' goal of the softer side of urban regeneration policy. Perhaps the real motivation behind such policy measures is their symbolic function, which might serve the interests of administrators, or indirectly lead to positive results or 'by-catches' (see Edelman, 1964). By offering sports programs and building 'Cruyff courts' (named after the most famous Dutch soccer player Johan Cruyff), for example, administrators can show that they find the development of disadvantaged youth important, perhaps thereby empowering these youth indirectly. And through the instigation of resident committees, governments emphasize that the community's voice counts in tackling disorder. That in turn may lead to greater public confidence in local government. Sieber (1981) labels this process *placation* which refers to a policy approach that is aimed at the reassurance of citizens, and is predominantly rhetorical in nature.

When applied to quality of life issues, however, this symbolic motive of policy has some significant caveats. Symbolic measures often overshadow much useful, but less visible, work of local community workers and law enforcement officers. The problem is that the activities of these street-level bureaucrats (Lipsky, 1980) working in human service organisations are not very mediagenic. They are less visual than, say, a block party or a 'Cruyff court'. Naturally, the actions of street professionals are also not always *evidence based*. But their daily operations in problem areas seem more relevant than many of the interventions described in this paper. Moreover, government bodies often insufficiently realize that citizens themselves have little desire for the symbolic function of quality of life policy. When residents participate in a neighbourhood committee they want their input to be taken seriously. This requires flexibility of local government bodies and a professional network that is mandated to translate residents' input into effective action. If this precondition is not met, it is more likely that such schemes will lead to a *decrease* in resident activism. Finally, even well-intentioned programs can have negative effects. The international literature yields

numerous examples of social initiatives that appealed symbolically, but instrumentally showed adverse results (see e.g. Petrosino et al., 2002).

Urban social policy will probably always balance between the forces of symbolism and instrumentalism. Future sociological and socio-historical research could, however, shed light on when the scales tip more often in favour of one or the other, and how this influences the design of urban policy.

5.   Validity in qualitative evaluation

5.1.   Introduction[11]

Since the days of pioneers such as the anthropologist Franz Boas and members of the Chicago School of Urban Sociology, a vast literature has developed on the procedures and underlying philosophies of qualitative research. Focusing on the natural behaviour of people and their perceptions of the social world (Yin, 1994; Denzin et al., 2005), one relatively recent development is the increasing use of qualitative methods and information for evaluation purposes in social policy and health care. The health sector in particular has seen a surge in approaches and writings on *evidence-based* procedures and evaluation research that involve or require inclusion of qualitative methods (see e.g. Pope & Mays, 2006). Some of the reasons for this are an increased understanding and acknowledgement of the limits of experiment or questionnaire-based quantitative research on 'what works', a growing demand for ethical considerations in evaluations, insight into clients' and patients' well-being, and the need for a more thorough understanding of how experimentally determined *evidence based* interventions connect to people's emotions, experiences and habits.

How to determine the validity (or 'truth value', Lincoln & Guba, 1985) of such investigations is a difficult question. Although validity in qualitative research has been widely reflected upon in the methodological literature (and is still often subject of fierce debate), the link with evaluation research is underexplored. A conceptual analysis of this interrelation is nonetheless opportune. The increased importance given to qualitative information in the *evidence based* paradigm requires a more precise conceptualization of validity criteria that goes beyond just academic reflection. After all, one can argue that policy verdicts that are based on qualitative evaluations must be legitimized by valid research, just as quantitative effect research is subject to validity standards.

In this chapter I will explore aspects of validity of qualitative research with the explicit objective of connecting them with aspects of evaluation. In doing so I will argue that different purposes of qualitative evaluation in social policy and health care can be linked with different scientific paradigms and perspectives, thus transcending unproductive paradigmatic divisions as well as providing a theoretical framework for researchers and reviewers of qualitative evaluations. It is, however, with a more general discussion of qualitative inquiry and validity that my exploration must begin.

---

[11] This chapter is based on content in the book *Kwalitatief evalueren. Mogelijkheden en beperkingen*, Lub, V. (2013), Den Haag: Boom Lemma.

5.2.   Qualitative research and the question of validity

Since roughly the 1970s increasing criticism of the reliability and objectivity of qualitative research has resulted in a growing interest in establishing more rigorous criteria and methodological standards. This attention has somewhat shifted from standards for the implementation of the study by the researcher to verification strategies for evaluating the credibility of qualitative findings by external reviewers (Morse et al., 2002). Validity is a key concept in this discussion. In the positivistic, rational tradition of science methodology, 'validity' can be defined as the degree to which the indicators or variables of a research concept are made measurable, accurately represent that concept. Does, for example, a response scale that measures interactions with members of other ethnic groups indeed refer to intercultural tolerance? Obviously this rational definition of validity does not work well in qualitative naturalistic research - which does not focus on variables on interval or ratio level. As a result, in the qualitative methodological literature 'validity' has been labelled with all sorts of alternative terms such as such authenticity, adequacy, plausibility and neutrality (see e.g., Lincoln & Guba, 1985, Maxwell, 1996; Meriam, 1998). Nevertheless, within the academic community the idea seems to be dominant that qualitative researchers must demonstrate in one way or another that their research results are valid. Several authors have therefore sought to develop specific research procedures and criteria aimed at increasing the validity of qualitative outcomes.

Probably the most influential is the work of Guba and Lincoln (see Guba & Lincoln, 1981; Lincoln & Guba, 1985). Guba and Lincoln were one of the first to develop specific criteria for qualitative research. They started from the premise that although all research must possess high 'truth value', the properties of knowledge within the 'rational' (or quantitative) paradigm is different from the properties of knowledge within the 'naturalistic' (or qualitative) paradigm (cf. Morse, 2002). According to Guba and Lincoln, each paradigm requires specific criteria to determine the veracity of the research. Within the rational paradigm criteria can be formulated in terms of internal validity, external validity, reliability and objectivity. Within the naturalistic paradigm, one is better to speak of criteria such as 'credibility', 'fittingness' and 'confirmability'. Later Lincoln and Guba (1985) redefined these concepts to 'credibility', 'transferability' and 'dependability', though this seemed more of a cosmetic make-over than a conceptual one. Guba and Lincoln subsequently formulated several procedures aimed to increase the credibility of qualitative research.

The most widely cited are *negative case selection*, *peer debriefing*, *prolonged engagement and observation* in the field, *audit trails* and *member checks*. Negative case selection is the process of data analysis through which the interpretation of the

data is stretched by consciously seeking out and explaining outliers ('negative cases') in the data. Peer debriefing is a form of external evaluation of the qualitative research process. Lincoln and Guba (1985: 308) describe the role of the peer reviewer as the 'devil's advocate'. It is a person who asks difficult questions about the procedures, meanings, interpretations and conclusions of the investigation. Prolonged engagement implies that the investigator performs the study for a considerable period. That is to say, a period long enough to adequately represent the subject under investigation. An audit trail - also called decision trail - means that researchers document the research process and the choices during that process meticulously and chronologically, for example through logs and memos. This documentation trail allows external evaluators to check the following questions: can the findings be supported by the data? Are the conclusions logical? Can methodological choices be justified? Member checking involves systematic feedback obtained from informants or participants on the collected data, set categories, interpretations and conclusions of the study. Its aim is to minimize the risk of misinterpretations by the researchers.

Many qualitative researchers still regard these criteria as methodological standards. In the wake of Guba and Lincoln many authors supplemented or perfected their criteria, or suggested alternative terminology for similar procedures. Around of the turn of the last century Morse et al. (2002: 15) concluded that this had resulted in a 'plethora' of terms and criteria that often brought more confusion than clarity in establishing the validity of qualitative research. Today, still, methodological textbooks on this point shows a lot of overlap and most criteria are directly obtained from the themes first conceptualized by Guba and Lincoln.

5.3. Critique on validity standards in qualitative research

Despite efforts to advance the debate on validity, some authors reject the desirability of predetermined criteria for qualitative research altogether. Sandelowski and Barroso (2002), for example, distance themselves from the search for general criteria for qualitative research because in their view the epistemological range of qualitative methods is too broad to be represented by a uniform set of criteria. Instead, they argue for a more rhetorical approach in which the quality of each project must be determined separately for every study. Sandelowski and Barroso (2002: 8) write: 'The only site for evaluating research studies - whether they are qualitative or quantitative - is the report itself.' In the same vein, Rolfe (2006) points out that qualitative research cannot fall back on a single scientific paradigm. Any attempt to reach consensus on qualitative criteria, according to Rolfe, therefore has little chance. There simply is no common understanding of the field of qualitative theory or methodology which can collectively

be described as 'qualitative research' (unlike quantitative research, perhaps, that despite the diversity in applications is based on similar mathematical laws). Rolfe argues his case by showing contradictions and paradoxes of common validity checks. Member checking and peer debriefing, for instance, are problematic because if it is assumed that there is no universal truth, but only different and additionally constructed truths to which every individual provides his or her own meaning (in effect the premise of much qualitative research), then we cannot expect that the respondents or external evaluators of qualitative studies will come to corresponding categories and conclusions (cf. Sandelowski, 1993: 3).

Hammersley (2007) is also critical of the attempt to formulate uniform criteria of qualitative research. He points out that there are several qualitative approaches that explicitly reject the idea that the production of knowledge should be the only immediate goal of research, and instead insist on political 'action'. Proponents of this approach believe that qualitative research is a part of the education and social advancement of people, and that this function is rendered useless when education is separated from research (see e.g. Elliot, 1988). Related approaches call for a political function of qualitative research by requiring that they should be focused on bringing change of one kind or another: for example by challenging capitalism, racism, homophobia or social disadvantage. Hammersley emphasises that it is important to point out that these approaches produce alternative considerations in assessing the quality of research, in addition to traditional epistemological considerations. Such alternative criteria should be much more formulated in terms of education, politics, ethics, aesthetics or even economics (e.g. does the study offer value for money?).

Like Rolfe and Sandelowski, Hammersley ultimately rejects the idea that a final set of universal criteria can be formulated. The obstacles to this not only originate out of political 'action' objectives, but also out of differences in value assumptions. He illustrates this with the example of the growing research on the impact of gender differences in educational achievement of children (see Hammersley, 2007: 294-295). To accept this as a relevant research topic, argues Hammersley, it is vital that one believes in the equality of the sexes (which may not be shared by certain religious groups or socio-biologists). One also has to share the assumption that certain disparities in the classroom affect educational performance, defined in terms of exam success. However, there are people who see gender differences as a predominantly social construct; and there are those who deny that school exams provide a sound indication of educational performance. What Hammersley shows with this example is that research in the social domain is framed by a series of value-assumptions which can produce serious differences. The fewer underlying assumptions of a particular research field are shared, the more difficult it is to defend the relevance of the research and the more difficult it is to reach consensus on the validity criteria of that research.

Hammersley nevertheless believes that certain criteria, in the form of 'guidelines', can play a role for a more rigorous assessment of qualitative research, though he does not clarify what these guidelines should be.

> My conclusion is that guidelines for qualitative research are desirable [..].
> However, the barriers to our being able to produce any set of common guidelines, even among qualitative researchers, are formidable. At the same time, we should not simply accept at face value methodological pluralism, reinforcing it by treating each qualitative approach as having its own unique set of quality criteria. Dialogue on this issue across different approaches, and indeed across the qualitative - quantitative divide is essential for the future of social and educational research (Hammersley, 2007: 301).

Finally, according to some, the debate on validity criteria has little attention for the ethics of qualitative research. One of the defining characteristics of qualitative methods is that they - more than quantitative methods - provide a participatory function to the researcher (Pope & Mays, 2006). Qualitative research requires that the researcher talks to people and observes them up close, and captures their behaviours and experiences accurately. The social interaction with the respondent thus requires tact and sensitivity of the researcher. Davies and Dodd (2002) argue that because of this, the quest for greater rigor cannot be separated from the interaction with the research subject and the ethics that the researcher should take into account (see e.g. Grol, 2001 on 'building bridges' between professional pride, payer profit and patient satisfaction). In their eyes, qualitative research should certainly be transparent and accountable, but not at the expense of the interests of the respondent and its context. Davies and Dodd therefore argue that the validity of the research should also be formulated in terms of attentiveness, empathy, carefulness, sensitivity, respect, reflection, conscientiousness, engagement, awareness and openness on the part of the investigator(s).

5.4.   A model for validity in qualitative evaluation

One's stance on the question of validity in qualitative research, then, primarily depends on which scientific paradigm is supported, leading some authors to reject the desirability of predetermined criteria for qualitative research altogether. Yet one could equally argue that *different* paradigms require *different* criteria and this line of reasoning also has implications for thinking about validity standards in qualitative evaluation research (to which I will come in a moment). Creswell and Miller (2000) argue that general discussions about validity in qualitative research provide little

guidance as to why one procedure might be selected for use by researchers over other procedures. They suggest that this choice is essentially governed by two perspectives: the lens researchers use to validate their studies and the researchers' paradigm assumptions. In order to advance this idea Creswell and Miller constructed a two-dimensional framework that can help researchers identify appropriate validity procedures (see Table 7). In the framework, three traditionally competing paradigms are central, derived from Guba and Lincoln (1994), that can shape ones epistemological position towards qualitative research: post-positivism, constructivism and the critical paradigm (see Creswell & Miller, 2000: 125-126). The post-positivist researcher assumes that qualitative research - as well as quantitative research - must consist of rigorous methods and systematic forms of research. Within this paradigm one in fact is looking for the qualitative equivalent of the rigid methodological protocols in the quantitative research community (see e.g. Maxwell, 1996). The constructivist researcher assumes more pluralistic, interpretive and contextualized perspectives of reality (i.e. sensitive to time, place and situation). The procedures within this paradigm hence look for an alternative vocabulary for validity labels, for example 'transferability' instead of 'external validity'. The third paradigm assumption involves the critical perspective. This perspective emerged as a critique of alleged structural inequalities of modern society and power structures, and was embraced among qualitative researchers who committed to the empowerment of marginalized groups, for instance through action research. The implication for validity checking within this paradigm is that the validity of the study should constantly be criticized and 'negotiated' with stakeholders and participants, and that researchers should be reflexive and transparent about the kind of knowledge they disclose.

Based on the three paradigm assumptions, Creswell and Miller identify nine different types of validity procedures (see Table 8). Besides the paradigm assumptions, the procedures are arranged to different perspectives - Creswell and Miller call these 'lenses' – by which the validity of qualitative research can be assessed (see vertical axis of the table). These lenses constitute the researchers' own perspective, that of the participants in the research or that of external reviewers or readers.

Table 8. Validity procedures within qualitative lens and paradigm assumptions (from: Creswell and Miller, 2000: 126)

Paradigm assumption/Lens	Postpositivist Paradigm	Constructivist Paradigm	Critical Paradigm
Lens of the researcher	Triangulation	Disconfirming evidence	Researcher reflexivity
Lens of Study Participants	Member checking	Prolonged engagement in the field	Collaboration
Lens of People External to the Study (Reviewers, Readers)	Audit trail	Thick description	Peer debriefing

Member checking, audit trail, prolonged engagement, peer debriefing and disconfirming evidence (negative case selection) are criteria discussed earlier from the work of Guba and Lincoln. Triangulation is a validity procedure where researchers base their categories and/or conclusions on different sources of information (see e.g. Oliver, Aicken & Arai, 2013). The researcher might look, for example, whether certain conclusions derived from interviews are consistent with findings from document analysis and observations. The more the categories and conclusions are confirmed by different data sources, the more valid the results. Reflexivity of the researcher refers to the extent to which researchers make their personal values and beliefs explicit in the research report, in such a way that is clear to what extent they might have influenced the results. Creswell and Miller point out that it is especially important that researchers describe this early in the research so that readers and external reviewers evaluate their personal position - and the possible impact on later outcomes - properly. This can be done in the form of a methodological paragraph or comments throughout the report. Thick description involves the detailed description of the setting, the participants and the themes of the study. The purpose of thick description is that it creates 'probability', that is, a statement of affairs that takes readers as much as possible into the studied world and its main characters. Detail is the key word here. Researchers should describe, for instance, interactions with informants, personal experiences, or provide a detailed description of the emotions of the respondents. Collaboration is a criterion that is particularly associated with the critical paradigm, meaning that participants should be involved in the study as co-researchers, or in less formal relationships.

Creswell and Millers' model advances the debate on validity in qualitative research in several ways. It elegantly unites different worldviews or paradigms within qualitative research with key perspectives by which the validity of qualitative research

can be assessed: that of the researcher, the respondent and the external reader. It further explicates the criteria that are essential for each respective paradigm and/or perspective.

The framework of Creswell and Miller provides a model for validity in qualitative evaluation. In principle, qualitative evaluation can have three different purposes. It can focus on the *instrumental effectiveness* of the policy itself (does it work? what are its main working components?), on the *meaning* of the policy or program for clients, target groups and practitioners (how do clients and practitioners experience it? How do practitioners shape it?), and it can follow an *emancipatory approach* in which the research itself can take either two perspectives, but simultaneously and deliberately aims to empower or educate those involved in the program (see e.g. the many forms of participatory action research). Given their properties and focal points, these evaluation purposes can be linked with the paradigm assumptions Creswell and Miller distinguish. Instrumental effectiveness corresponds to post-positivism. Within the postpositivist worldview a particular social program or policy is primarily seen as a separate entity – as an 'instrument' - causing a certain effect in a linear process, whose independent effect can be evaluated accordingly. Postpositivists also tend to believe there is a single reality, whereas constructivists believe that there are multiple, constructed realities (Lincoln & Guba, 1985). Evaluating the significance/meaning of the intervention for clients and target groups thus corresponds to constructivism, which aims to expose the multiple realities about the implementation and functioning of the policy or program constructed by those involved in the policy or program. Finally, the emancipatory function can be linked to the critical paradigm, which underlines the educational and social advancement of clients and target groups and cooperation between researchers and respondents involved in the evaluation. By linking these purposes and paradigms we can create a new model with relevant validity criteria, specifically for qualitative evaluation (see Table 9).

Table 9. Validity procedures of qualitative evaluation aligned to purposes, paradigms and perspectives

Purpose evaluation > *Perspective*	**Instrumental effectiveness policy / program** *(postpositivist paradigm)*	**Meaning policy/program for target group and practitioners** *(constructivist paradigm)*	**Empowerment clients / target group /practitioners** *(critical paradigm)*
Evaluator perspective	Triangulation (contrasting)	Disconfirming evidence (Fair dealing)	Researcher reflexivity
Evaluation participant perspective	Member checking	Prolonged engagement in the field	Collaboration
External reader/reviewer Perspective	Audit trail	Thick description	Peer debriefing

Each procedure in effect serves as a counterweight for inherent methodological weaknesses of the respective evaluation purposes. In case of a qualitative evaluation that primarily focuses on the instrumental effectiveness of a particular policy or program (does it work? what are its working components?), the criteria triangulation, member checking and conducting an audit trail are essential. These criteria are most appropriate to avoid or detect spurious (causal) inferences and possible biases, which in itself are significant potential distortions when assessing the instrumental effectiveness of a program or policy. If the goal is to uncover the meaning of the intervention for clients and target groups, then the research should acknowledge disconfirming evidence (or: negative case selection), there must be prolonged engagement in the field (not a snapshot study) and external readers should be able to identify the experiences of respondents adequately through thick description. These criteria counterbalance a too one-sided report of the experiences of particular individuals (disconfirming evidence) or circumstances (prolonged engagement) and allow for a thorough understanding of the experiences of respondents (thick description). If the evaluation has an emancipatory intent (empowerment), then reflexivity of the researcher in the study becomes more important. It should become clear how personal beliefs or dispositions might have influenced the investigation as most empowerment based evaluations (e.g. participatory action research) require a strong involvement of the researcher with his

research subjects and the theme under study (with the possible risk of 'going native'). From the perspective of the participants, empowerment evaluations must also employ collaboration, which means that participants should be involved in the evaluation as co-researchers, or in less formal relationships.

Let me illustrate the model with a hypothetical example. Suppose a qualitative case study is performed which aims to investigate the effectiveness and working components of a particular program. In the case study, interviews, observations and documentation analysis are conducted. Given its main purpose – evaluating the effectiveness of the program itself – it is essential that from the evaluator's perspective triangulation is performed (do findings from interviews, observations and document analysis overlap?), that from the participant perspective there is member checking (do participants endorse certain conclusions/interpretations made by the evaluators?) and an audit trail is conducted so that external reviewers can verify if presented findings can be supported by the data and causal inferences about the workings of the program are grounded. The same steps can be followed with the other evaluation purposes (meaning, empowerment) 'checking' procedures from the columns down and linking them with the perspectives from the rows.

Note that Creswell and Millers' criteria are completed with other relevant procedures. Triangulation can be enhanced by *contrasting* outcomes with findings from other types of research or previous research outcomes (see Onwuegbuzie & Leech, 2007). For instance via so-called multi-site studies whereby observations from different situations in different research locations are compared, or through the systematic comparison of results of qualitative research with insights from the (scientific) literature. The principle of *fair dealing* (Mays & Pope, 2000) is a logical addition to disconfirming evidence. Fair dealing ensures that many different perspectives are covered in the evaluation (not only that of the policy's or program's target group) so that the viewpoint of one particular group cannot be presented as an overall representation of the workings of the intervention. For example, if one wants to evaluate the effect of a particular support program for pregnant teenagers, this criterion ensures that not only the pregnant teenagers are interviewed but also the relevant groups around the teenagers such as social professionals, trainers, family members etc. After all, in such an evaluation one does not investigate 'pregnant teenagers' but the program designed for their education and support and to assess it adequately, several perspectives are needed.

5.5. Conclusions and discussion

The framework presented in this article constitutes an ideal type (cf. Weber, 1971 [1951]). In reality, objectives of qualitative inquiry in evaluation (instrumental, meaning, empowerment) can - and often do - overlap and some methodological criteria can be applied in several paradigms or purposes. It is important in this respect to note that the framework can serve as a checklist but that its main value is as a (theoretical) reference point of the key issues when taking the different purposes, paradigms ad perspectives of qualitative evaluation into account. One of its most important features is that it avoids 'taking sides' in a paradigmatic and epistemological sense. Instead it accommodates a more pragmatic approach. What criteria are preferred for 'good' qualitative research will always be dependent on one's scientific world view. It can be expected, therefore, that practitioners and policy makers will continue to make use of different types of qualitative evaluations – emphasizing different purposes and starting from different paradigms - to support their specific programs and policies. All the more reason not to assess the 'truth value' of such evaluations on one particular monolithic world view, but instead to let qualitative evaluation criteria correspond with the different functions that qualitative evaluations can have (instrumental, meaning, empowerment) and the paradigms and lenses through which it can be assessed. In this way the model not only combines flexibility with rigour, it also – to some extent - answers the call for 'common guidelines' (Hammersley, 2007) while at the same time respecting paradigmatic differences.

One can pose the question if the model should be supplemented with ethical criteria, such as those formulated by Davies and Dodd (2002). I propose not to. Not because ethical considerations are not important, but because the objective of upholding certain ethical standards in *evaluation research* is to some extent naïve. In practice, it is often impossible to be fully transparent about the research outcomes to respondents or to fully protect their interests and contexts. Evaluating a certain program or policy requires a critical stance and it goes without saying that some research results might affect respondents negatively. For instance, because it turns out that practitioners of a social program do not perform well or that their particular approach is actually counterproductive. Should the qualitative researcher be transparent about such possible outcomes in advance then it would be unlikely he or she could count on much cooperation from respondents in the field.

Finally, apart from the methodological considerations, it would be fruitful to take a step back and analyse the social, cultural and institutional 'embeddedness' of some of these issues (Strassheim & Kettunen, 2014). It would be interesting, for instance, to discern why certain preferences for particular paradigms and purposes for

evaluation seem to correspond with different time periods and sectors. In health research, the personal experience and realities clients and target groups provide to a particular program (constructivist paradigm) certainly has become more important over the last two decades or so, and this development has served as a supplement (or perhaps counterweight) to the dominance of postpositivist investigations focused on the instrumental effectiveness of programs. In community and social work it seems the reverse is at work. Historically highly influenced by postmodern and constructivist schools of thought, programs in community and social work are now increasingly fitted into experimental or 'quantized' research models reminiscent of the old *modus operandi* in the health sector. Moreover, in several European countries institutions and government agencies that are active in community and social work have set up databases of 'effective interventions' analogous to the already established *evidence based* databases in the health sectors (treating interventions as 'independent instruments': postpositivist paradigm). The emancipatory function of evaluation (critical paradigm), prevalent in the 1970s and 1980s, is today again visible in research projects commissioned by the European Union (EU). Most EU research calls demand involvement of 'end-users', 'practitioners' and negotiations with 'stakeholders' and require that proposals elaborate on how such end-users will benefit from the undertaken research.

Sociological and socio-historical research not only can shed light on how and why such sectorial paradigm shifts occur; it could also investigate how these shifts influence ideas on what counts as 'evidence' of particular social programs or policies, and if or how this in turn influences ideas on the role of qualitative information in evaluation.

6.   Studying policy from multiple order perspectives

The main research questions of the study were formulated as follows:

What is the likelihood to which policy measures in the social domain achieve their intended effects? To what degree can claims about those intended effects be substantiated by scientific insights? How can qualitative information be validated in this respect? How can we incorporate reflexive perspectives in studying social policy?

I must first conclude that many social policy measures fail to achieve their intended effects. The Social Support Act did not result in a greater contribution from civil society organizations to social goals and policies, and instigating the social participation of many vulnerable groups led to paradoxical outcomes (increasing their isolation). Few government measures against polarization, radicalization and quality of life issues in deprived neighbourhoods can be substantiated by scientific insights. In fact, only one of the scrutinized measures (neighbourhood watch schemes) can boast a strong scientific basis. The other studied policy measures in these fields have either a moderate or weak scientific basis. Of course, the case studies presented in this book comprise only a particular segment of the social policy domain at a particular time. General conclusions about the effectiveness of social policy measures are therefore unjustified. But if we consider the outcomes of the case studies a random sample, providing at least some indication of the likelihood to which policy measures in the social domain achieve their intended effects, then there is no immediate cause for optimism. Qualitative evaluations can provide meaningful explanations to the (non)workings of social policy measures, provided they follow certain validity procedures. Case study 4 presented a model that combines rigour and flexibility in this respect, taking into account the different worldviews, preferences and perspectives that generally shape evaluation research.

This brings us to the final, and perhaps most fundamental, question of the study; how can we incorporate more reflexive perspectives in studying social policy issues? Niklas Luhmann (1997) theorized that there are multiple ways of looking at reality. Each way of looking comprises a particular viewpoint that has its own blind spot. Luhmann distinguished a 'first', 'second' and 'third' order perspective. A first order perspective entails a direct view from the observer to the observed phenomenon, like an audience in a movie theatre watching a movie. A second order perspective entails an observation

of the observer(s), for example a market researcher observing how audiences respond to certain movie scenes. A third order perspective operates at a more reflexive level; it entails an observation of the observation of the observation. In our example, a media scientist might research how market researchers observe the behaviour of audiences during movies, and what particular analytical framework they use for this.

In my view, policy, too, can be studied from these multiple order perspectives that shift from the micro to the macro level (see Table 10). A first order perspective assesses the policy directly and predominantly deals with instrumental questions; does it work? A second order perspective takes a step back and focuses on methodological issues; how can we best assess if the policy works? A third order perspective is (critical) reflexive; it problematizes the foundations of certain concepts of policy and reflects on what wider social, cultural, political, institutional and paradigmatic factors influence its design, implementation and assessment.

Table 10. Studying policy from multiple order perspectives

Policy Perspective	Level of analysis	Central Question
First order	Micro	Instrumental: does the policy work?
Second order	Meso	Methodological: how to assess if the policy works?
Third order	Macro	Reflexive: what wider factors influence the design, implementation and assessment of policy?

In the remainder of this chapter, I will discuss outcomes of the presented case studies within the framework of these three perspectives.

6.1.   First order perspective

Like an audience in a movie theatre rating the quality of the movie, the goal of case studies 1, 2 and 3 is to directly assess whether certain policies and the assumptions behind those policies are plausible (first order perspective).

Case study 1 showed that social policy can sometimes not only be ineffective, but in fact can produce the opposite of what is intended. Contrary to the objectives of the Social Support Act, a 'revitalisation' of the Dutch civil society – in terms of a greater contribution to social goals and policies – remains problematic, whilst professional entities thrive under the new governmental élan. Explanations for this include a wait-and-see approach by municipalities to actively involve civil society, 'crowding-out' citizenship by government bodies and professional institutions and a restraint and inability of civic organisations to include and share responsibility for vulnerable groups. Other paradoxical outcomes stem from a too-dogmatic approach to the social participation of people with severe mental disabilities. Instigating the socialisation of these groups through mandatory measures can in practice increase their isolation. The problem lies in the assumption that increased socialisation into mainstream contexts, irrespective of the locality, improves the overall well-being and life satisfaction of people with mental impairments and disorders. The case study showed that this causal relationship is not so clear-cut and more often is counterproductive, sending clients further into a downward spiral.

Case studies 2 and 3 (a theory driven approach) showed that dominant policy approaches within the social domain cannot automatically boast strong scientific support, but are more often based on weak to moderate scientific support. In anti-polarization and radicalization measures (cases study 2), social ecological approaches seem to be effective in reducing extremist behaviour in radical youth, but do not necessarily lead to an ideological change. Mediation by peers seems to exert a positive influence in conflict resolution in most cases, but the causal link between the employment of young people and positive outcomes remains unclear. Intergroup contact on average reduces prejudices about other groups, but effect sizes are generally small and there is no evidence for a long-term impact. Positive outcomes of intergroup contact also cannot automatically be generalised to interethnic contact between adolescents. Finally, the theory that enhancing the self-esteem of at-risk youth will prevent radicalisation has a weak scientific basis. It is doubtful that self-esteem enhancement makes (Muslim) youth more resilient against radical tendencies. Most studies report negative behavioural effects as a result of one-sided mental

empowerment methods. Boosting levels of self-confidence can even contribute to greater in-group bias.

As for social policies against neighbourhood disorder (case study 3), only one can currently boast a relatively strong scientific basis: the assumption that neighbourhood watch schemes contribute positively to neighbourhood quality of life. This type of intervention is therefore promising, albeit that the understanding of the processes and conditions surrounding the functioning of the neighbourhood watch is still underdeveloped. For policy assumptions about the contribution of resident representation to resolve disorder issues, the scientific basis can be labelled as 'moderate'. Three of the examined theories have a weak scientific basis. Given the current body of knowledge, policy assumptions in the field of contact stimulation, the collective shaping of behavioural codes of conduct and sports programs will probably contribute little to the quality of life in disadvantaged neighbourhoods, though sport can be used as a vehicle or 'gateway' for alternative behavioural change programs for at-risk youth.

6.2.   Second order perspective

The second order perspective shifts from the micro to the meso level. Like the market researcher observing how the audience rates the movie, it asks the question; how are we (better) to assess the plausibility of policy? Drawing from the case studies, I would like to argue the following.

Utilizing existing literature
Firstly, more focus should be put on evaluating basic assumptions of social policy and specific programs through readily available and existing literature. Rather than relapsing straight into the 'measurement reflex' for each and every individual social intervention, it seems more appropriate to first visit the library. This study has shown that there is a vast reservoir of scientific knowledge from which to build in the evaluation and design of policy. Any gaps in knowledge about a particular approach when consulting the literature can always be closed through empirical research at a later stage. Studying how plausible effects are beforehand - instead of measuring afterwards - not only saves scarce resources. It also opens up a space of useful scientific knowledge which access would otherwise be limited to academics. Also, government institutions and funding agencies can make more stringent requirements on the plausibility of certain assumptions of interventions that they fund in advance, rather than desire 'evidence' afterwards. This prevents practitioners having to provide

empirical evidence for their specific approach (an activity for which they often lack the skills or resources), and at the same time puts the policy in question in a broader epistemological context: what premises does the intervention start from? How plausible are assumptions given earlier scientific knowledge? How credible are assumed causal links?

Applying the 'razor'
However, it is a misunderstanding that the policy theories under scrutiny should always be specified at length, describing all the treatment-as-delivered characteristics, other related exogenous factors, intervening processes and outcome variables. Instead, it is just as valuable to extract the most defining underlying principle of a particular form of intervention along with the postulated relationships between the assumed 'working' components, in the form of an ideal type. Ideal types assist the researcher in analysing and understanding complex social phenomena (Weber, 1971 [1951]; see also Ritzer, 1992). Most social policy interventions possess some ideal typical condition that has to be met for it to work, regardless of the specific context (by which I do not purport that context is not important)[12]. Characteristic of ideal types is that they do not have to represent the most 'pure' form of the object of study, but rather break it down to its essentials, like Ockham's famous reductionist principle of the 'razor'. For example, if a sports program in a deprived neighbourhood aims to mitigate criminal tendencies by offering an alternative social environment where young people learn positive values and norms, then it naturally follows that sport *must* have a significant impact on human social behaviour for it to work in the first place. Or if a meeting is organized between youth groups of different religions to promote mutual understanding, then it naturally follows that intergroup contact *must* have a significant impact on tolerance levels and reduction of stereotypes for it to work. And so on. Consequently, generalizations (ideal types) of the various forms of policy and their underlying assumptions can be constructed which then can be confronted with fundamental scientific insights.

---

[12] Astbury and Leeuw (2010: 374) make a similar point when they argue that mechanisms are often 'portable' in the sense that they are building blocks for middle-range program theories, which may be transferable to different contexts and policy domains. They cite Pawson (2006) who showed that sex offender naming and shaming interventions are not unique. The basic idea is also used by policy makers in relation to school league tables and safety indices for car manufacturers, among many others. According to Astbury and Leeuw, this suggests that although the precise details and context are likely to vary across different policy situations, this does not necessarily mean that we need to start anew every time a 'naming and shaming-type' intervention is launched. Instead of treating all interventions as completely novel, Astbury and Leeuw argue that building a knowledge and theoretical base about 'families of interventions' (Pawson, 2006), including the different types of mechanisms that underlie them may over time reduce 'policy amnesia' and constant reinvention of the wheel.

Reappraising observational research

Second, whilst looking at basic assumptions and applying qualitative methods, observational research deserves some reappraisal (see also Pope & Mays, 2009; Konnerup & Kongsted, 2012). Observations as a qualitative method can be used to 'test' mechanisms and assumptions of policy quite well. It is particularly suitable to refute certain assumptions. After all, if the conditions under which the intervention or measure should work or the required conditions for specific causal links are not present, then the effectiveness of the intervention can already be ruled out in advance. The design of future interventions can be adjusted on the basis of these insights. For instance, many programs and interventions in the social domain depend on collaboration of different institutions and practitioners in the field. Not infrequently, however, it appears in retrospect that the intervention had no effect because the required collaboration did not effectuate. Through observation, such obstacles can already be identified at an early stage, by which the intervention can be adjusted or revised. Observation also provides the opportunity to assess how practitioners (professionals, target groups, stakeholders, etc.) 'respond' to an intervention, and what consequences this has for the implementation of the policy. Pawson (2006) has pointed out that social programs usually are immediately transformed relative to their original design at their implementation in a real-life setting (see also Room, 2013). An intervention is not solely a systematic course of action to achieve a certain goal, it is also - depending on the interest of those involved - seized as a political power, as a resource, as a rhetorical argument, and so on. Interventions are, in the words of Pawson, 'active' and this can produce a gap between the design of the intervention on paper and its actual, contextual implementation. Observational studies can shed light on these processes and expose the mechanisms and conditions that are important herein.

Avoiding a 'one size, fits few' system of valuation

Thirdly, within the social domain it should be more clearly outlined what interventions are suitable for measuring effectiveness through the monocausal experimental model and what types of interventions are not. In the world of health promotion, delinquency, individual care and mental healthcare, methods usually have a sharp delineation in time and intensity, are limited in time and scope and are offered for specific, well defined target groups. This allows for a more consistent connection between the intervention and the intended result. Such interventions are thus more suitable for measurement through quantitative causal modelling (such as Randomized Controlled Trials, RCT). Yet in social work, community work and the realm of neighbourhood development, programs often involve complex methods in various stages of development that also harbour multiple objectives. For these interventions a single quantifiable assessment is

usually not possible (because in fact there is no one 'single' intervention and the policy focuses on different target groups). Here, alternative research methods come into view such as mixed methods research, qualitative inquiry or literature reviews. It is crucial that the social domain takes the differences in design and scope of different policies into account and allows for a more flexible application of research methods and definition of 'evidence'. Rigidly clinging on to the 'golden standard' of the RCT might otherwise produce a 'one size, fits few' system of valuation whereby some policies or programs – given their properties - will never reach the predicate of 'effective' even though they might work.

6.3. Third order perspective

A first and second order perspective evaluate the assumptions of policy and reflect on methodological issues on how to do so. Their view outside of the framework of the policies themselves, however, is limited. Yet like the media scientist reflecting on how market researchers analyse audience responses to movies, the third order perspective enables us to reflect on what wider social, cultural, political, institutional and paradigmatic factors might influence the design, implementation and assessment of policy, along the way problematizing the foundations of certain concepts.

Reflections from the case studies
Case study 1 showed that many of the paradoxical effects of the Social Support Act in effect stem from a problematic notion of 'normalisation'; the ideal that people with mental impairments should enjoy patterns and conditions of everyday living as close as possible to the norms and patterns of the mainstream society (Nirje, 1969; Yates et al., 2008). Yet normalisation *de facto* starts from the assumption that being 'different' is less desirable than being 'normal' and that disabled people should therefore strive to be something other than what they are (Morris, 1991). The study showed that this too-dogmatic idea of normalisation ultimately resulted in increased social isolation instead of participation.

Case study 2 problematized the conceptualisation of 'polarisation' or 'radicalisation'. If these issues are conceived as treatable pathologies (as the four policy approaches discussed in the case study tend to do), then it makes sense to focus on treatment plans (i.e., 'interventions') to tackle radicalisation or polarisation issues. However, such an approach could arguably be dismissed as naive, when no account is offered of sociopolitical factors or perceived injustices fuelling radicalisation or ethnic tensions. It is essential to investigate further to what extent contextual variables are essential

components of interventions seeking a lasting reduction of these problems. Should, for example, in the case of social ecological approaches for Muslim youth, Imams and mosques be seen as desirable 'social attachments' or as potential detractors from mainstream society? In what way do migration factors influence the efficacy of individual counselling programmes for minority ethnic youth? And how strongly do geopolitical developments that influence domestic relations between ethnic groups hinder or accommodate endeavours to promote mutual understanding through intergroup contact?

Case study 3 prompts a debate on *instrumental policy* (policy as a means to an end) versus *symbolic policy* (policy as communicative strategy). The case study was mainly concerned with instrumental policy, and from that perspective, the literature questions the plausibility of several social measures against neighbourhood disorder and crime. But perhaps this is not the 'real' goal of the softer side of urban regeneration policy. Perhaps the real motivation behind such policy measures is their symbolic function, which might serve the interests of administrators, or indirectly lead to positive results or 'by-catches' (see Edelman, 1964). By offering sports programs and building 'Cruyff courts' (named after the most famous Dutch soccer player Johan Cruyff), for example, administrators can show that they find the development of disadvantaged youth important, perhaps thereby empowering these youth indirectly. And through the instigation of resident committees, governments emphasize that the community's voice counts in tackling disorder. That in turn may lead to greater public confidence in local government. Sieber (1981) labels this process *placation* which refers to a policy approach that is aimed at the reassurance of citizens, and is predominantly rhetorical in nature.

Finally, case study 4 posed the question why certain preferences for particular paradigms and purposes for evaluation seem to correspond with different time periods and sectors. In health research, the personal experience and 'multiple realities' clients and target groups provide to a particular program (constructivist paradigm) certainly has become more important over the last two decades or so, and this development has served as a supplement (or perhaps counterweight) to the dominance of postpositivist investigations focused on the instrumental effectiveness of programs. In community and social work it seems the reverse is at work. Historically highly influenced by postmodern and constructivist schools of thought, programs in community and social work are now increasingly fitted into experimental or 'quantized' research models reminiscent of the old *modus operandi* in the health sector. Moreover, in several European countries institutions and government agencies that are active in community and social work have set up databases of 'effective interventions' analogous to the

already established *evidence based* databases in the health sectors (treating interventions as 'independent instruments': postpositivist paradigm). The emancipatory function of evaluation (critical paradigm), prevalent in the 1970s and 1980s, is today again visible in research projects commissioned by the European Union (EU). Many EU research calls demand involvement of 'end-users', 'practitioners' and negotiations with 'stakeholders' and require that proposals elaborate on how such end-users will benefit from the undertaken research. Sociological and socio-historical research not only can shed light on how and why such sectorial paradigm shifts occur; it could also investigate how these shifts influence ideas on what counts as 'evidence' of particular social programs or policies, and if or how this in turn influences ideas on the role of qualitative information in evaluation.

General third order reflections
This section concludes with some general third order reflections. To start with, no matter how detailed a notion one can have of what factors influence the plausibility of policy, this does not mean that those insights can be applied automatically in newly designed policy interventions. The studies on which the conclusions of this thesis are based were often very detailed and differentiated many subtle variations in behaviour and psychological dispositions. This complicates the objective of applying the particular intervention that is consistent with the 'diagnosis' of the problem, that is, if that diagnosis can even be made accurately. For example, within the framework of case study 2 it remains unclear which Muslim youth can or should be labelled as 'radical'. Some Dutch cases painfully illustrate the difficult assessment of the often subtle differences between religious orthodoxy and supposed radicalism of some Muslim youth (see Demant et al., 2008: 58–61). Also, it is one thing to conclude that, for example, the moral education of parents influences xenophobic ideas of at-risk youth; to effectively incorporate such notions in government policy is another matter. And most of the studies on intergroup contact have been conducted in the psychological laboratory. But how might we bring people together under the right conditions in everyday life (see also Wilson, 2011)? These are difficult questions and insight into the possibilities of social policy can thus paradoxically also lead to an understanding of its limitations.

Also, one can question whether it is possible to make statements about the effectiveness of existing approaches solely by examining the literature, as case studies 2 and 3 do. Theory driven evaluation through extant literature provides a reference point of what works, but not a recipe. From the literature, insights emerge about effective elements based on empirical findings. Whether such potential effectiveness takes shape in practice depends on a variety of social, political and administrative

constraints and contextual factors (see, for example, Spicker, 2011; Pawson & Tilley, 1997). However helpful, literature reviews by no means yield unequivocal policy verdicts (Boaz & Pawson, 2005; see also Chambers et al., 2012). The viewpoint of confronting causal assumptions with scientific evidence from the literature can never paint a complete picture of the efforts required to address issues such as neighbourhood disorder, extremism and ethnic tensions. Moreover, no intervention occurs in a social vacuum. Knowledge obtained from scientific research rarely flows in its most pure form from the source to the receiver. Politicians and administrators want to prove that their policy is effective, and this has implications for the interpretation of data. Every policy researcher is familiar with the phenomenon of incorrectly interpreting research findings by the media or selectively 'shopping' from research reports by politicians. In particular junctures, certain policy and political paradigms are simply dominant, and often certain scientific insights are ignored because they do not suit politicians and administrators (Engbersen, 2012). In this sense, political realities impose limits on policy interventions, regardless of their plausibility. Moreover, practitioners in the field can have their own interpretation of what mechanisms ought to contribute to specified goals. The interpretations of these 'street-level bureaucrats' (Lipsky, 1980) can deviate from problem definitions and goal specifications put forth in official documents. Practitioners may narrow the goals of policy makers or elaborate upon them, or substitute entirely different goals. Such a transformation can produce a gap between the (plausible) design of a particular programme and its actual implementation.

Finally, to be clear, the plausibility of policy deals with only one aspect of the spectrum of policy analysis. Hemerijck (2003) argues that a legitimate analysis of policy not only comprises its instrumental effectiveness (does it work?), but also its social acceptability (do citizens accept the measures at hand?), its institutional feasibility (is the policy implementable?) and its constitutional legitimacy (does the law allow it?). Naturally, these spectra relate to one another in a big way. For instance, putting a police officer on every street corner might be effective in reducing crime, but it will not score very high on institutional feasibility or social acceptability. To add to Hemerijck's conceptual framework, I would like to highlight the symbolic function of policy; does it serve a rhetorical purpose for those in power? In fact, this might well be one of the most important motives behind government policy design and implementation, and serves as a possible explanation for Jan Kees de Jager's reluctance to act on the critical report of the Dutch Court of Audit on the lack of scrutiny of government subsidies (see chapter 1; introduction). Perhaps De Jager foresaw that an acknowledgement of the Court's instrumental critique on government subsidies would render much of government policy ineffective, thereby distorting its rhetorical function and indirectly undermining the legitimacy of those in power. The fashionable political

discourse of 'participation' implemented at the cost of the psychological safety of mentally disabled clients (discussed in chapter 2), is another illustrative example of the symbolic power of policy. In the real world, then, research outcomes are not the only source of policy design. The plausibility of the policy itself must always compete with social, institutional, constitutional and symbolic aspects as well.

6.4.   Themes for future research

In conclusion, I would like to suggest the following themes for future research:

- The interaction between an 'intervention' oriented approach against social problems and wider sociopolitical factors. For example, in radicalisation/polarisation issues, in what way do migration factors influence the efficacy of individual counselling programmes for minority ethnic youth? How strongly do geopolitical developments that influence domestic relations between ethnic groups hinder or accommodate endeavours to promote mutual understanding through intergroup contact? (see e.g. current developments in Syria and Iraq attracting some Muslim youth in Western Europe); and how much of the variance of positive developments in quality of life issues in deprived neighbourhoods can/should be explained by policy, and how much by extra-local factors (demographical, economic, educational etc.)?

- The interrelation between policy theories and scientific theories; what mechanisms guide this interrelation? Much policy seems to be influenced by scientific theories, but the reverse also applies; science, especially social science, is just as influenced by policy questions[13]. Moreover, often scientific theories obtain a different meaning when they are translated into policy interventions and *vice versa*. In other cases, the boundaries between scientific theories and policy theories are not clear to begin with and a policy theory does not necessarily have to fall back on a scientific underpinning at all. Can we discern patterns in this seemingly chaotic picture?

---

[13] In the United Kingdom a debate has come up in recent years under the term 'policy-based evidence making' that questions some of the so-called evidence-based policy making in the spheres of government, warning for perverse effects. Boden and Epstein (2006) argue that many evidence-based approaches in effect are fundamentally flawed by the fact that the government seeks to control the production knowledge to the point where this type of research might best be described as 'policy-based evidence'; the commissioning of research with the sole aim of supporting policy that has already been decided upon, instead of genuinely collecting evidence on its effectiveness.

- Paradigm shifts through the dimensions of time and space; why do certain preferences for particular scientific paradigms and purposes for evaluation seem to correspond with different time periods and sectors? How and why do such sectorial paradigm shifts occur? How do these shifts influence ideas on what particular type of 'science' is consulted or what counts as 'evidence' of particular social programs or policies? And how does this in turn influence ideas on the preferred methodology of particular evaluation designs?

- Instrumental policy (policy as means to an end) vs. symbolic policy (policy as a communicative/rhetorical strategy); social policy will probably always balance between the forces of symbolism and instrumentalism. But when do the scales tip in favour of one or the other? And how does this (im)balance influence the design of social policy in particular situations?

References

Aberson, C.L., Healy, M. & Romero, V. (2000). Ingroup bias and self-esteem: A meta analysis, *Personality and Social Psychology Review*, 4, 157-173.

Allport, G.W. (1954). *The nature of prejudice*, Reading, MA: Addison Wesley.

Astbury, B. & Leeuw, F. (2010). Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. *American Journal of Evaluation*, 31, 3, 363-381.

Baker, T. & Evans, H. (2002). *Resident Services Organisations: Lessons from the Pilot Projects.* Manchester: PEP.

Bauman, Z. (2004). *Wasted Lives. Modernity and its Outcasts*, Cambridge: Polity.

Baumeister, R.F., Smart, L. & Boden, J.M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem, *Psychological Review,* 103, 5-33.

Bekhuis, H., Ruiter, S. & Coenders, M. (2009). Xenofobie onder jongeren: de invloed van interetnisch contact, *Mens en Maatschappij*, 84, 72-85.

Bennett, T.H. (1990). *Evaluating neighbourhood watch. Cambridge Studies in Criminology LXI.* Aldershot, UK: Gower.

Bennett, T., Holloway, K. & Farrington, D.P. (2006). Does neighborhood watch reduce crime? A systematic review and meta-analysis. *Journal of Experimental Criminology,* 2, 437-458.

Bickmore, K. (2002). Peer Mediation Training and Program Implementation in Elementary Schools: Research Results, *Conflict Resolution Quarterly,* 20, 137-160.

Bjørgo, T., Carlsson, Y. & Haaland, T. (2005). Conflict processes between youth groups in a Norwegian City: Polarization and Revenge, *European Journal of Crime, Criminal Law and Criminal Justice*, 13, 1, 44-78.

Blascovich, J., Mendes, W.B., Hunter, S.B. & Lickel, B. (2000). 'Stigma, threat, and social interactions', in: T. F. Heatherton, R.E. Kleck, M.R. Hebl and J.G. Hull (eds.), *The social psychology of stigma,* New York: Guilford Press.

Boaz, A. & Pawson, R. (2005). The Perilous Road from Evidence to Policy: Five Journeys Compared, *Journal of Social Policy*, 34, 2, 175-194.

Boden, R. & Epstein, D. (2006). Managing the research imagination? Globalisation and research in higher education. *Globalisation, societies and education*, 4, 2, 223-236.

Bongers, K. & Langeveld, N. (2011). *Evaluatie projecten gedragscodes. Ervaringen van professionals en bewoners.* Utrecht: Centrum voor Veiligheid en Criminaliteitspreventie.

Borduin, C.M., Mann, B.J., Cone, L.T., Henggeler, S.W., Fucci, B.R., Blaske, D.M. & Williams, R.A. (1995). Multisystemic treatment of serious juvenile offenders:long-term prevention of criminality and violence, *Journal of Consulting and Clinical Psychology*, 63, 569–578.

Borduin, C.M., Schaeffer, C.M. & Heiblum, N. (2009). A randomized clinical trial of multisystemic therapy with juvenile sexual offenders: Effects on youth social ecology and criminal activity, *Journal of Consulting and Clinical Psychology*. 77, 1, 26-37.

Borduin, C.M., Schaeffer, C.M. & Heiblum, N. (2009). A randomized clinical trial of multisystemic therapy with juvenile sexual offenders: effects on youth social ecology and criminal activity, *Journal of Consulting and Clinical Psychology*, 77, 1, 26–37.

Boston, J. M., Pallot, J. & Walsh, P. (1996). *Public Management: The New Zealand Model*, Auckland: Oxford University Press.

Boutellier, H. & Marissing, E. van (2011). Veiligheidsarrangementen in IJburg. Over de praktijk van de besturing van veiligheid. *Tijdschrift voor Veiligheid,* 10, 1, 59-68.

Bouzar, D. (2006). *Quelle éducation face au radicalisme religieux?,* Paris: Editions Dunod.

Bradnum M., Niewoudt, J. & Tredoux, C. (1993). Contact and the Alteration of Racial Attitudes in South Africa, *South African Journal of Psychology,* 23, 4, 204-211.

Brighton, S. (2007). British Muslims, multiculturalism and UK foreign policy: 'integration' and 'cohesion' in and beyond the state, *International affairs,* 83, 1, 1-17.

Brink, G. van den. (2006). *Culturele contrasten. Het verhaal van de migranten in Rotterdam*, Amsterdam: Bert Bakker.

Browning, C.R., Feinberg, S.L. & Dietz, R. (2004). The Paradox of Social Organisation: Networks, Collective Efficacy, and Violent Crime in Urban Neighborhoods. *Social Forces*, 83, 2, 503-534.

Bullock C.S. (1976). Interracial Contact and Student Prejudice. The Impact of Southern School Desegregation, *Youth and Society* 7, 3, 271-310.

Bursik, R.J.& Grasmick, H. (1993). *Neighborhoods and Crime: The Dimensions of Effective Community Control.* New York: Lexington Books.

Burton, P., Croft, J., Hastings, A., Slater, T., Goodlad, R., Abbott, J. & Macdonald, G. (2004). *What works in community involvement in area-based initiatives? A systematic review of the literature.* Home Office Online Report 53/04.

Cadat, M. & Engbersen, R. (2006). *Leefbaarheid op het (verstedelijkt) platteland: anticiperen op multi-etniciteit. Een verkenning naar culturele spanningen op het platteland*, Utrecht: NIZW.

Cameron, L., Rutland, A., Brown R. & Douch, R. (2006). Changing Children's Intergroup Attitudes Toward Refugees: Testing Different Models of Extended Contact, *Child Development,* 77, 1208-1219.

Campbell, D.T. (1965). Ethnocentric and other altruistic motives, in D. Levine (ed.), *Nebraska Symposium on Motivation* (p. 94-172), New York: McGraw-Hill.

Carmichael, D. (2008). *Youth Sport vs. Youth Crime. Evidence that youth engaged in organized sport are not likely to participate in criminal activities.* Ontario: Active Healthy Links Inc.

CCV. (2010). *Stappenplan voor het opzetten van een project jongerenbuurtbemiddeling*, Utrecht: Centrum Criminaliteitspreventie en Veiligheid.

Chambers, D., Grant, R., Warren, E., Pearson, S. & Wilson, P. (2012). Use of evidence from systematic reviews to inform commissioning decisions: a case study, *Evidence & Policy*, 8, 2, 141-148.

Chekroun, P. (2008). Social Control Behavior: The Effects of Social Situations and Personal Implication on Informal Social Sanctions. *Social and Personality Psychology Compass*, 2, 6, 2141-2158.

Chen, H.T. (1990). *Theory-Driven Evaluation*, Beverly Hills, CA: Sage.

Chen, H.T. & Rossi, P.H. (1983). Evaluating with sense. The Theory-Driven Approach, *Evaluation Review*, 7, 3, 283-302.

Clark, D. & Southern, R. (2006). Comparing institutional designs for neighbourhood renewal: neighbourhood management in Britain and the Re´gies de Quartier in France, *Policy and Politics*, 34, 1, 173–191.

Cohen, G.L., Aronson, J. & Steele, C.M. (2000). When beliefs yield to evidence: Reducing biased evaluation by affirming the self, *Personality and Social Psychology Bulletin,* 26, 1151-1164.

Cohen, R. (1999). *The School Mediator's Field Guide. Prejudice, Sexual Harassment, Large Groups And Other Daily Challenges*, Watertown: School Mediation Associates.

Cook, S.W. (1985). Experimenting on Social Issues. The Case of School Desegregation, *American Psychologist.* 40, 4, 452-460.

Cooper, R., & Slavin, R.E. (2004). 'Cooperative learning: an instructional strategy to improve intergroup relations', in: Stephen, W.G. & Vogt, W.P. (eds.), *Education programs for improving intergroup relations*, New York: Teachers College Press.

Creswell, J.W, & Miller, D.L. (2000). Determining validity in qualitative inquiry, *Theory into Practice*, 39, 124-130.

Crocker, J., Thompson, L.L., McGraw, K.M. & Ingerman, C. (1987). Downward comparison, prejudice, and evaluations of others: Effects of self-esteem and threat', *Journal of Personality and Social Psychology,* 52, 907-916.

Culham, A., & Nind, M. (2003). Deconstructing normalisation: Clearing the way for inclusion, *Journal of Intellectual and Developmental Disability,* 28, 1, 65–78.

Curtis, N.M., Ronan, K.R. & Borduin, C.M. (2004). Multisystemic Treatment: A Meta-Analysis of Outcome Studies, *Journal of Family Psychology*, 18, 3, 411-419.

Damen, M. (2010). De invloed van de omgevingskwaliteit op de beoordeling van de leefbaarheid door bewoners. *Tijdschrift voor Veiligheid*, 91, 22-43.

Daniels, E. & Leaper, C. (2006). A longitudinal investigation of sport participation, peer acceptance, and self-esteem among adolescent girls and boys. *Sex Roles,* 55, 875-880.

Davies, D, & Dodd J. (2002). Qualitative Research and the Question of Rigor, *Qual Health Res*, 12, 279-289.

Deakin, N. (1996). The Devils in the Detail: Some Reflections on Contracting for Social Care by Voluntary Organizations, *Social Policy and Administration,* 30, 1, 20-38.

Dean, H. (2006). *Social Policy. Short introductions*. Cambridge: Polity Press.

Demant, F., Slootman, M., Buijs, F. & Tillie, J. (2008). *Teruggang en uittreding. Processen van radicalisering ontleed*, Amsterdam: IMES.

Demant, F., Wagenaar, W. & Donselaar, J., van. (2009). *Deradicaliseren in de praktijk. Monitor Racisme en Extremisme*, Amsterdam: Pallast Publications.

Denzin, N.K, & Lincoln, Y.S. (eds.), 2005, *The Sage Handbook of Qualitative Research* (3rd ed.), Thousand Oaks, CA: Sage

Diekstra, R.F.W. (2006). *Van straataanpak tot straatburgerschap*. Rapportage van de straataanpak. Verschenen als rapport B4 in de serie Sociale integratie. Rotterdam: Gemeente Rotterdam.

DiMaggio, P.J. & Powell, W.W. (1983). The iron cage revisited: institutional isomorphism and collective rationality in organizational fields, *American Sociological Review,* 48, 2, 147-160.

Donnellan, M.B., Trzesniewski, K.H., Robins, R.W., Moffitt, T.E. & Caspi, A. (2005). Low self-esteem is related to aggression, antisocial behavior and delinquency, *Psychological Science*, 16, 328-335.

Duyvendak, J.W. (2005). 'De buurt: guur oord of kuuroord? Bedenkingen bij de extramuralisering'. In: *Inleiding op de conferentie 'Beter thuis?', Stuurgroep Wonen-Zorg-Dienstverlening*, Amsterdam.

Edelman, M.J. (1964). *The Symbolic Uses of Politics*. University of Illinois Press.

Elliott, J. (1988). Response to Patricia Broadfoot's presidential address, *British Educational Research Journal,* 14, 2, 191–194.

Endresen, I.M. & Olweus, D. (2005). Participation in power sports and antisocial involvement in preadolescent and adolescent boys. *Journal of Child Psychology and Psychiatry*, 46, 5, 468-478.

Engbersen, G. (2012). Wat werkelijk telt. Hoe politieke grenzen gekozen remedies bepalen, in: Uitermark, J., Gielen, A.J., & Ham, M. (eds.), *Wat werkt nu werkelijk? Politiek en praktijk van sociale interventies* (p. 251-262). Amsterdam: Van Gennip.

Engbersen, G. (2009). *Fatale remedies. Over onbedoelde gevolgen van beleid en kennis.* Amsterdam: Pallast Publications.

Englund, C. (2002). *Exit Motala - Case Study. An example of how to use integration into the society as a process of healing for different groups of people living in the margin*, Stockholm: National Focal Point of the European Monitoring Centre on Racism and Xenophobia (EUMC).

Eshel, Y. & Dicker, R. (1995). Congruence and Incongruence in Perceived Ethnic Acceptance Among Israeli Students, *The Journal of Social Psychology*, 135, 2, 251-262.

Ewijk, H. van (2010). *European Social Policy and Social Work. Citizenship-based Social Work*, London: Routledge.

Finlay, K.A. & Stephan, W.G. (2000). Improving Intergroup Relations: The Effects of Empathy on Racial Attitudes, *Journal of Applied Social Psychology,* 8, 1720-1737.

Fischer, C.S., Jackson, R.M., Stueve, C.A., Gerson, K., Jones L.M. & Baldassare M. (1977). *Networks and places. Social relations in the urban setting*. New York: Free Press.

Fletcher-Morgan, P.C, & Leyland, K. (2010). Making People More Responsible: The Blair Governments Programme for Changing Citizens' Behaviour, *Political Studies,* 58, 427-449.

Gibbons, S., Ebbeck, V., & Weiss, M. (1995). Fair play for kids: effects on the moral development of children in physical education, *Res Q Exerc Sport*, 66, 247-255.

Gielen, A.J. (2009). *Een kwestie van identiteit. Evaluatie training Identiteit en Weerbaarheid voor moslima's,* Amsterdam: A.G. Advies.

Green, B.L. & McAllister, C. (2002). Theory-Based Participatory Evaluation: A Powerful Tool for Evaluating Family Support Programs, *The Bulletin of the National Center for Zero to Three,* 30-36.

Greenberg, S.W., Rohe, W.M. & Williams, J.R. (1985). *Informal citizen action and crime prevention at the neighborhood level*. Washington, D.C.: Government Printing Office.

Grol, R. (2001). Improving the quality of medical care: building bridges among professional pride, payer profit, and patient satisfaction, *Journal of the American Medical Association*, 286, 2578-85.

Guba, E.G, & Lincoln, Y.S. (1981). *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches*, San Francisco, CA: Jossey-Bass.

Guba, E.G, & Lincoln, Y.S. (1994). Competing paradigms in qualitative research, in N.K. Denzin & Y.S. Lincoln (eds.), *Handbook of qualitative research*, Thousand Oaks: Sage, 105-117.

Gurin, P., Peng, T., Lopez, G., & Nagda, B. (1999). 'Context, identity and intergroup relations', in: Prentice, D.A. & Miller, D.T. (eds.), *Cultural Divides. Understanding and Overcoming Group Conflict*, New York: Russel Sage Foundation.

Hall, E. (2004). Social geographies of learning disability: Narratives of exclusion and inclusion, *Area,* 36, 3, 298–306.

Hall, S. & Hickman P. (2011). Resident Participation in Housing Regeneration in France. *Housing Studies*, 26, 6, 827-843.

Hammersley, M. (2007). The issue of quality in qualitative research. *International Journal of Research & Method in Education*, 30, 3, 287-305.

Hardyns, W., Pleysier, S. & Pauwels, L. (2010). Oorzaken van het mijden van onveilige situaties bij mannen en vrouwen. Een contextuele analyse op basis van de 'collectieve efficacy'-theorie. *Tijdschrift voor Criminologie*, 52, 4, 374-394.

Harter, S. (2003). 'The development of self-representations during childhood and adolescence', in: Leary, M.R. & Tangney, J.P. (eds.), *Handbook of self and identity*, New York: Guilford Press.

Hartmann, D. & Massoglia, M. (2007). Reassessing the relationship between high school sports participation and deviance: Evidence of Enduring Bifurcated Effects. *The Sociological Quarterly*, 48, 485-505.

Heatherton, T.F. & Vohs, K.D. (2000). Interpersonal evaluations following threats to self: Role of self-esteem, *Journal of Personality and Social Psychology,* 78, 725-736.

Heitmeyer, W. & Sitzer, P. (2008). *Right-wing extremist violence among adolescents in Germany* in New Directions for Youth Development No. 119, 169-185. Jossey-Bass.

Heitmeyer, W., J. Müller & Schröder, H. (1997). *Verlockender Fundamentalismus. Türkische Jugendliche in Deutschland*, Frankfurt am Main: Suhrkamp.

Hemerijck, A. (2003). Vier kernvragen van beleid. *Beleid en Maatschappij*, 30, 3-19.

Hortulanus, R. (2004). *Het belang van de Wet maatschappelijke ondersteuning*, Utrecht: MOgroep.

Hunter, A. (1985). 'Private, Parochial and Public Social Orders: The Problem of Crime and Incivility in Urban Communities', in: Suttles, G. & Zalt, M. (eds.), *The Challenge of Social Control*. Norwoord, NJ: Alblex.

Jacobs, J. (1961). *The Death and Life of Great American Cities*. New York: Random House.

Johnstone, A.A. (1991). *Rationalized Epistemology. Taking Solipsism Seriously*, Albany: State University of New York Press.

Jones, T.S. (2004). Conflict Resolution Education: The Field, The Findings and the Future, *Conflict Resolutions Quarterly,* 22, 233-267.

Kleinhans, R.J. & Bolt, G. (2011). Like Chicken and Eggs. The Intricate Interplay between neighbourhood disorder and collective efficiency, in: Duyvendak, J., Nicholls, W., de Olde, C., Roussel, S. and Hirschel I. (eds.), *The struggle to belong. Dealing with diversity in 21st century urban settings* (p. 1-29). Amsterdam: Research Committee 21 (RC21), Sociology of urban and regional development, International sociological association.

Konnerup, M. & Kongsted, H.C. (2012). Do Cochrane reviews provide a good model for social science? The role of observational studies in systematic reviews, *Evidence & Policy*, 8, 1, 79-96.

Kreager, D.A. (2007). Unnecessary Roughness? School Sports, Peer Networks, and Male Adolescent Violence. *American Sociological Review*, 72, 705-724.

Kunneman, H.P. (2000). Levenskunst en burgerschap in een technopolis, *Humanistiek*, 1, 66-83.

Kwekkeboom, M.H. & Koops H. (2005). *Socialization of care*, The Hague: Netherlands Institute for Social Research.

Lane-Garon, P. (2000). Practicing Peace: The Impact of a School-Based Conflict Resolution Program on Elementary Students, *Peace and Change,* 25, 467-483.

'Learning from grant evaluations', Dutch Court of Audit, 2011.

Leeuw, F.L. (2003). Reconstructing program theories: methods available and problems to be solved, *American Journal of Evaluation*, 24, 5-20.

Leeuw, T. de. & Swaaningen, R. van (2011). Veiligheid in veelvoud: beeld, beleid en realiteit in Rotterdams Oude Westen. *Tijdschrift voor Veiligheid* (10)1, 26-42.

Lelieveldt, H. (2004). Helping Citizens Help Themselves, *Urban Affairs Review*, 39(5), 531- 551.

Lelieveldt, H. & Kolk, H. van der (2005). 'Structureel sociaal kapitaal en het oplossen van buurtproblemen', in: Volker, B. (ed.), *Burgers in de buurt. Samenleven in school, wijk en vereniging.* Amsterdam: Amsterdam University Press, 25-40.

Lincoln, Y.S, & Guba, E.G. (1985). *Naturalistic inquiry*, Thousand Oaks: Sage.

Linders, L. (2010). *What has neighbourhood cohesion to do with informal care?*, ESPAnet Social Policy Research Day 2010, University of Amsterdam: UvA.

Lindo, F. (2008). *Interetnische contacten tussen scholieren in het voortgezet onderwijs. Een analyse en waardering van het onderzoek binnen het paradigma van de contacthypothese*, Amsterdam: IMES.

Lipsey, M.W. (1993). 'Theory as Method: Small Theories of Treatments', in: L.B. Sechrest & A.G. Scott (eds.), *Understanding Causes and Generalizing About Them, New Directions for Program Evaluation*, no. 57, 5-38.

Lipskey, M. (1980). *Street-Level Bureaucracy. Dilemmas of the Individual in Public Services*. New York: Russel Sage Foundation.

Lub, V. (2005). *Mensen maken de stad. Een belevingsonderzoek op buurtniveau naar sociale cohesie en actief burgerschap in twee Rotterdamse achterstandswijken.* Rotterdam: Sociaal-wetenschappelijke afdeling SoZaWe. Doctoraalscriptie Erasmus Universiteit Rotterdam.

Lub, V. (2009). *Stimulering van maatschappelijke binding van jongeren. Een verkennning naar sociale interventies op het terrein van radicalisering en culturele spanningen.* Utrecht: MOVISIE.

Lub, V. (2013a). Polarisation, Radicalisation and Social Policy. Evaluating the Theories of Change, *Evidence and Policy,* 9, 2, 165-183.

Lub, V. (2013b). *Schoon, heel en werkzaam? Een wetenschappelijke beoordeling van sociale interventies op het terrein van buurtleefbaarheid.* Den Haag: Boom Lemma.

Lub, V. (2014). *Kwalitatief evalueren in het sociale domein. Mogelijkheden en beperkingen.* Den Haag: Boom Lemma.

Lub, V., Sprinkhuizen A. & Cromwijk, R. (2010). *Trendrapport WMO 2010. Het spel op het maatschappelijk middenveld.* Utrecht: MOVISIE.

Lub, V. & Uyterlinde M. (2012). Evaluating State-Promoted Civic Engagement and Participation of Vulnerable Groups. The Paradoxical Policies of the Social Support Act in the Netherlands, *Journal of Social Policy*, 41, 2, 373-390.

Lub, V., Arum, S. van & Sprinkhuizen, A. (2009). Voor het verval. Belemmeringen en voorwaarden voor vroegsignalering en bewonersparticipatie in wijkveiligheid. *Journal of Social Intervention. Theory and Practice*, 18, 4, 62-82.

Lub, V., N. de Groot & J. Schaafsma (2011). *Polarisatie en radicalisering: de onderbouwing van sociale interventies getoetst.* Utrecht: MOVISIE / Universiteit van Tilburg.

Lub, V., Sprinkhuizen, A., & Alblas M. (2008). *Trendrapport: de uitvoering van de Wmo in beeld,* Utrecht: MOVISIE.

Luhmann, N. (1997). *Die Gesellschaft der Gesellschaft*. Frankfurt am Main: Suhrkamp.

Marissing, E. van (2008). *Buurten bij beleidsmakers. Stedelijke beleidsprocessen, bewonersparticipatie en sociale cohesie in vroeg-naoorlogse stadswijken in Nederland*, KNAG / Faculteit Geowetenschappen Universiteit Utrecht.

Marissing, E. van & Meere, F. de (2009). Afwachtende gemeenten. Wmo is nog geen stimulans om sociale contacten te bevorderen, *Tijdschrift voor sociale vraagstukken.* 10/11, 18-21.

Marrinetto, M. (2003). Who Wants to be an Active Citizen? The Politics and Practice of Community Involvement, *Sociology*, 37, 103-120.

Maxwell, J.A. (1996). *Qualitative research design: An interactive approach,* Thousand Oaks: Sage.

Mays, N. & Pope, C. (2000). Assessing quality in qualitative research, *BMJ*, 320, 50-52.

McClenahan C., Cairns, E., Dunn, S. & Morgan, V. (1996). Intergroups Friendships. Integrated and Desegregated Schools in Northern Ireland, *The Journal of Social Psychology* 136, 5, 549-558.

McConville, M. & Shepherd, D. (1992). *Watching Police, Watching Communities*. London: Routledge.

McPhedran, S. (2011). Disability and Community Life: Does Regional Living Enhance Social Participation? *Journal of Disability Policy Studies,* 20, 5, 1-15.

Meere, F. de e.a., Verwey-Jonker Instituut (2004). *Sociale integratie… zo doen we dat. Rotterdamse straten op de agenda*. Verschenen als B2 in de serie Sociale integratie. Rotterdam: gemeente Rotterdam.

Mendes, W.B., Blascovich, J., Lickel, B. & Hunter, S., (2002). Challenge and threat during social interaction with White and Black men, *Personality and Social Psychology Bulletin,* 28, 939-952.

Metz, J. (2010). Ervaring in Dordrecht: professionals verdringen burgers, *Tijdschrift voor sociale vraagstukken,* 1/2, 12-15.

Milner, P. & Kelly, B. (2009). Community participation and inclusion: people with disabilities defining their place, *Disability and Society,* 24, 1, 47-62.

Ministerie van BZK. (2007). *Actieplan Polarisatie en Radicalisering 2007-2011*, Den Haag: BZK.

Ministerie van BZK. (2008). *Trendanalyse Polarisatie en Radicalisering december 2008*, Den Haag: BZK.

Mogroep. (2007). 'Wet maatschappelijke ondersteuning een brede participatiewet'. Brief aan de informateur, 24-01-2007, Mogroep.

Molina, E.L. & Wittig, M.A. (2006). Relative importance of contact conditions in explaining prejudice reduction in a classroom context: separate and equal?, *Journal of Social Issues*, 62, 3, 489–509.

Morenoff, J.D., Sampson, R.J. & Raudenbusch, S. (2001). Neighborhood Inequality, Collective Efficacy and the Spacial Dynamics of Urban Violence, *Criminology*, 39, 517-560.

Morris, J. (1991). *Pride against prejudice.* London: The Women's Press.

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.

Müller, T. (2011). Zwerfafval als teken van wanorde en onveiligheid? Een kwalitatief onderzoek naar het weggooien van zwerfafval door jongeren. *Tijdschrift voor Criminologie*, 53, 1, 40-54.

Muir, K., Fisher, K.R., Abello, D., & Dadich, N. (2010). 'I didn't like just sittin' around all day': Facilitating Social and Community Participation Among People with Mental Illness and High Levels of Psychiatric Disability', *Journal of Social Policy,* 39, 3, 375-39.

Nirje, B. (1969). 'The normalization principle and its human management implications', in: Kugel R. & Wolfensberger, W. (eds.), *Changing patterns in residential services for the mentally retarded,* Washington, DC: Presidential Committee on Mental Retardation.

Nixon, H.L. (1997). Gender, sport, and aggressive behaviour outside sport. *Journal of Sport and Social Issues*, 21, 379-391.

Noije, L. van & Wittebrood, K. (2008). *Sociale veiligheid ontsleuteld. Veronderstelde en werkelijke effecten van veiligheidsbeleid*. Den Haag: Sociaal en Cultureel Planbureau.

Notten, T. & Witte, T. (2011). Radicalisering moslimjongeren in Rotterdam?, *Journal of Social Intervention: Theory and Practice*, 20, 54-69.

Oliver, K, Aicken, C, & Arai, L. (2013). Making the most of obesity research: developing research and policy objectives through evidence triangulation, *Evidence & Policy*, 9, 2, 207-23.

Onwuegbuzie, A.J, & Leech, N. (2007). Validity and Qualitative Research: An Oxymoron? *Quality & Quantity*, 41, 233–249.

Ostrom, E. (2000). Crowding out Citizenship, *Scandinavian Political Studies,* 23, 3-16.

Pawson, R. (2006). *Evidence-based Policy. A Realist Perspective*. London: Sage Publications.

Pawson, R. & Tilley N. (1997). *Realistic Evaluation*, Londen: Sage Publications.

Pels, T. & Vollebergh, W. (2006). *Diversiteit in opvoeding en ontwikkeling. Een overzicht van recent onderzoek in Nederland.* Amsterdam: Aksant.

Pescosolido, B.A. & Rubin, B.A. (2000). The web of group affiliations revisited: social life, postmodernism and sociology, *American Sociological Review*, 65, 52-76.

Petrosino, A., Turpin-Petrosino., C & Buehler J. (2002). *'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency*. Cochrane Database of Systematic Reviews 2002, Issue 2. Art. No: CD002796.

Pettigrew, T.F. & Tropp, L.R. (2006). A Meta-Analytic Test of Intergroup Contact Theory, *Journal of Personality and Social Psychology,* 5, 751-783.

Pettigrew, T.F. & Tropp, L.R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators, *European Journal of Social Psychology,* 38, 922-934.

Phalet, K. (2003). 'De constructive van etnisch-religieuze identiteit en alteriteit: Beeldvorming en believing van de islam in Nederland', in: Foblets, M.C. & Cornelis, E. (eds.), *Identiteitsbeleving bij allochtone jongeren*, Leuven: ACCO.

Pope, C. & Mays, N. (2006). *Qualitative research in health care*. Oxford: Blackwell Publishing.

Pope, C. & Mays, N. (2009). Critical reflections on the rise of qualitative research, *British Medical Journal*, 339, 737-739.

Pratt, T. & Cullen, F. (2005). 'Assessing Macro-Level Predictors and Theories of Crime: A Meta-Analysis', in: Tonry, M. (ed.), *Crime and Justice: A Review of Research. Chicago*: University of Chicago Press: p. 373-45.

Prior, D. & Barnes, M. (2009). 'Subversion' and the Analysis of Public Policy', in: Barnes, M. & Prior, D. (eds.), *Subversive Citizens. Power, Agency and Resistance in Public Services*, Bristol: The Policy Press.

Rauch, A. & Dornette, J. (2009). Equal Rights and Equal Duties? Activating Labour Market Policy and the Participation of Long-term Unemployed People with Disabilities after the Reform of the German Welfare State, *Journal of Social Policy,* 39, 1, 53-70.

Ritzer, G. (1992). *Metatheorizing*. New York: Sage.

RMO. (2008). *Verkenning participatie. Arbeid, vrijwillige inzet en mantelzorg in perspectief.* Den Haag: SWP.

RMO. (2009). *Polarisatie. Bedreigend en verrijkend*. Den Haag: Raad voor Maatschappelijke Ontwikkeling.

Roberts, A.R. & Yaeger K.R. (2006). *Foundations of Evidence Based Social Work Practice*. New York: Oxford University Press.

Roche, M. (2002). 'Social Citizenship: Grounds of Social Change', in: Isin, E. & Turner, B. (eds.), *Handbook of Citizenship Studies,* Londen: Sage Publications.

Roe, D. & Lysaker, P.H. (2012). Concerns and issues that have emerged with the evolution of evidence-based practice, *Journal of Mental Health*, 21, 5, 427-429.

Rolfe, G. (2006). Validity, trustworthiness and rigour: quality and the idea of qualitative research, *Journal of Advanced Nursing,* 53, 3, 304–310.

Romance, T., Weiss, M. & Bockoven J. (1986). A program to promote moral development through elementary school physical education, *J Teach Phys Educ*, 5, 126-136.

Rommelspacher, B. (2006). *Der Hass hat uns geeint. Junge Rechtsextreme und ihr Ausstieg aus der Szene.* Frankfurt/Main: Campus Verlag.

Rooijen, M. van (2004). *Gemeenten hebben geen benul van psychiatrie*, in: www.psy.nl , nr. 7.

Room, G. (2013). Evidence for agile policy makers: the contribution of transformative realism, *Evidence & Policy*, 9, 2, 225-244.

Rosenbaum, D.P. (1987). The theory and research behind neighborhood watch: Is it a sound fear and crime reduction strategy? *Crime and Delinquency*, 33, 103-134.

Roseth, C., Johnson, D., & Johnson, R. (2008). Promoting early adolescents' achievements and peer relationships: The effects of cooperative, competitive, and individualistic goal structures, *Psychological Bulletin*, 134, 223-246.

Ross, C.E. & Mirowsky, J. (2001). Neighborhood disadvantage, disorder, and health, *Journal of Health and Social Behavior*, 42, 3, 258–276.

Sampson, R.J. (2009). Disparity and diversity in the contemporary city: social (dis)order revisited, *British Journal of Sociology*, 60, 1, 1–31.

Sampson, R., Raudenbusch, S.W. & Earls, F. (1997). Neighborhoods and violent crime: a multilevel study of collective efficacy, *Science,* 277 (15 August), 918-924.

Sampson, R.J. (2011). 'The community', in: Wilson J.Q. & Petersilia, J. (eds.), *Crime and Public Policy*: p. 210-236. New York: Oxford University Press.

San, M. van. (2006). 'Hoe drugssmokkel acceptabel werd in de Curacaose gemeenschap', in: Tonkens, E. Uitermark, J. & Ham, M. (eds.), *Handboek moraliseren. Burgerschap en ongedeelde moraal*. Amsterdam: Van Gennip.

Sandelowski, M. (1993). Rigor or rigor mortis: the problem of rigor in qualitative research revisited, *Advances in Nursing Science,* 16, 2, 1–8.

Sandelowski, M. & Barroso J. (2002). Reading qualitative studies, *International Journal of Qualitative Methods,* 1, 1.

Schafer, W.E. (1969). Participation in Interscholastic Athletics and Delinquency: A preliminary Study, *Social Problems*, 17, 1, 40-47.

Schout, G.H. (2007). *Zorgvermijding en zorgverlamming. Een onderzoek naar competentieontwikkeling in de Openbare Geestelijke Gezondheidszorg (proefschrift),* Groningen: Rijksuniversiteit Groningen.

SCP. (2009). *Jaarrapport Integratie*. Den Haag: Sociaal en Cultureel Planbureau.

Shadish, W.R. jr. (1987). 'Program Micro- and Macrotheories: A Guide for Social Change', in: Bickman L. (ed.), *Using Program Theory in Evaluation. New Directions for Program Evaluation*, no. 33, 93-108.

Shaw, C.R. & McKay, H.D. (1942). *Juvenile delinquency and urban areas.* Chicago: University of Chicago Press.

Shelton, J. & Richeson, J. (2005). Intergroup contact and pluralistic ignorance, *Journal of Personality and Social Psychology*, 88, 91-107.

Shelton, J. & Richeson, J. (2006). Interracial interactions: A relational approach, in: Zanna, M.P. (ed.), *Advances in Experimental Social Psychology*, San Diego, CA: Elsevier Academic Press.

Sherif, M. (1966). *In common predicament: social psychology of intergroup conflict and cooperation,* Boston: Houghton Mifflin.

Sherman, L.W. (2011). 'Democratic Policing on the Evidence', in: Wilson, J.Q. & Petersilia, J. (eds.), *Crime and Public Policy*: p. 589-618. New York: Oxford University Press.

Sieber, S. (1981). *Fatal Remedies. The Ironies of Social Interventions*. New York: Plenum.

Silver, E. & Miller L.L. (2004). Sources of Informal Social Control in Chicago Neighborhoods, *Criminology*, 42, 3, 551-583.

Skogan W.G. (2006). *Police and Community in Chicago. A Tale of Three Cities.* Oxford New York: Oxford University Press.

Skogan, W. & Hartnett, S.M. (1997). *Community Policing, Chicago Style*. New York and Londen: Oxford University Press.

Skogan, W.G. (1990). *Disorder and decline. Crime and the spiral of decay in American cities.* New York: Free Press.

Small, M. (2009). *Unanticipated Gains: Origins of Network Inequality in Everyday Life*. New York: Oxford University Press.

SMN. (2008). *Peer preventie bij radicalisering – projectbeschrijving*. Utrecht: Samenwerkingsverband van Marokkaanse Nederlanders.

Spicker, P. (2011). Generalisation and Phronesis: Rethinking the Methodology of Social Policy, *Journal of Social Policy*, 40, 1, 1-19.

Squires, P. (1990). *Anti-social Policy: Welfare, Ideology and the Disciplinary State*, Hemel Hempstead: Harvester Wheatsheaf.

Stephan, W.G. & Finlay, K.R. (1999). The Role of Empathy in Improving Intergroup Relations, *Journal of Social Issues*, 55, 4, 729-743.

Stephens, M. (2011). Comparative Housing Research: A 'System- Embedded' Approach, *International Journal of Housing Policy*, 11, 4, 337-355.

Steptoe, A. & Butler, N. (1996). Sports participation and emotional wellbeing in adolescents. *The Lancet*, 347, 1789-92.

Stokkom, B. van & Toenders, N. (2010). *De sociale cohesie voorbij. Actieve burgers in achterstandswijken*. Amsterdam: Pallast Publications.

Strassheim, H, & Kettunen, P. (2014). When does evidence-based policy turn into policy-based evidence? Configurations, contexts and mechanisms, *Evidence & Policy*, 10, 2, 259-77.

Stucke, T.S. & Sporer, S.L. (2002). When a grandiose self-image is threatened: Narcissism and self-concept clarity as predictors of negative emotions and aggression following ego-threat, *Journal of Personality, 70,* 509-532.

Teddlie, C. & Tashakkori, A. (2009). *Foundations of Mixed Methods Research.* Thousand Oaks: Sage.

Thomaes, S., Bushman, B., Orobio de Castro B., Cohen, G. & Denissen, J.J.A. (2009). Reducing narcissistic aggression by buttressing self-esteem: An experimental field study, *Psychological Science,* 20, 1536-1542.

Tonkens, E. (2010). *Tussen onderschatten en overvragen. Actief burgerschap en activerende organisaties in de wijk.* Amsterdam: SUN.

Uitermark, J. & Duyvendak, J.W. (2006). *Ruimte maken voor straatburgerschap.* Verschenen als rapport B3 in de serie Sociale integratie. Rotterdam: Gemeente Rotterdam.

Verhoeven, I. & Ham, M. (2010). 'De overheid op zoek naar brave burgers', in: Verhoeven, I. & Ham, M. (eds.), *Brave burgers gezocht. De grenzen van de activerende overheid.* Amsterdam: Van Gennip.

Verkuyten, M. (1996). Personal Self-Esteem and Prejudice among Ethnic Majority and Minority Youth, *Journal of Research in Personality,* 30, 248-263.

Verplanke, L. & J.W. Duyvendak (2010). *Onder de mensen? Over het zelfstandig wonen van psychiatrische patiënten en mensen met een verstandelijke beperking.* Amsterdam: University Press.

Vertovec, S. & Rogers, A. (1999). *Muslim European Youth. Reproducing Ethnicity, Religion, Culture*, Aldershot: Ashgate.

VNG. (2008). *De toekomst van de Wmo. VNG-visie op de toekomst van de Wmo in relatie tot de AWBZ,* Vereniging van Nederlandse Gemeenten.

Vornberg J.A. & Grant R.T. (1976). Adolescent Cultural Acquaintance Experiences and Ethnic Group Attitudes, *Adolescence* 11, 4, 601.

Wagner U., Hewstone M. & Machleit U. (1989). Contact and Prejudice Between Germans and Turks: A Correlational Study, *Human Relations*, 42, 7, 561-574.

Webber, M. (1963). 'Order in diversity: communities without proprinquity', in: Wingo, L. (ed.), *Cities and space*. Baltimore: The John Hopkins Press.

Weber, M. (1971) [1951], *Die Objektivität sozialwissenschaftlicher und sozial-politischer Erkentniss*, Tübingen: J.C.B. Mohr.

Weber, M. (1951) *Die Objektivität sozialwissenschaftlicher und sozial-politischer Erkentniss*, in: Weber 1951 (ed. 1972).

Weiss, C.H. (1995). 'Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families', in: J. Connell, A. Kubbisch, L. Schorr & C. Weiss (eds.), *New Approaches to Evaluating Community Initiatives*. Washington, DC: The Aspen Institute.

Werdmölder, H. (2005). *Marokkaanse lieverdjes. Crimineel en hinderlijk gedrag onder Marokkaanse jongeren*. Amsterdam: Balans.

Wilson, J.Q. & Kelling, G.L. (2003). 'Broken windows: the police and neighbourhood safety', in: McLaughlin, E. Muncie, J. & Hughes, G. (eds.), *Criminological Perspectives: Essential Readings*, p. 400–412, London: Sage.

Wilson, T.D. (2011). *Redirect. The Surprising New Science of Psychological Change*. London: Penguin Books.

Wycoff, M.A. & Skogan, W. (1993). *Community Policing in Madison: Quality from the Inside Out. An Evaluation of Implementation and Impact. Research Report*. Washington, D.C.: National Institute of Justice.

Yates, S., Dyson. S & Hiles, D. (2008). Beyond normalization and impairment: theorizing subjectivity in learning difficulties – theory and practice, *Disability and Society,* 23, 3, 246-258.

Yin, R.K. (1994). *Casestudy Research, Design and Methods*. Thousand Oaks: Sage.

Summary

This thesis provides an analysis of the plausibility of policy based on four separate case studies from the social domain. Case study 1 takes a mixed method approach to assess whether state-promoted civic engagement and promoting the participation of vulnerable groups (e.g. people with mental disabilities, psychiatric patients) is an appropriate government endeavour, and, in the process, exposes some of the 'dark sides' of social policy (real and potential) that are often overlooked by politicians and administrators. Case studies 2 and 3 take a theory driven approach to evaluate the plausibility of dominant policy assumptions in subsequently anti-radicalization/polarization and quality of life issues in deprived neighbourhoods. These case studies in effect take an opposite approach to the multi-layered mixed methods research of case study 1 by focussing on basic policy assumptions (i.e. ideal types of theories of change) and confronting them with the scientific literature. With these two studies I aim to show that strong explanations about the plausibility of policy can be provided if they focus on theories that do not posit too much in them, like Ockham's famous reductionist principle of the 'Razor'. Finally, case study 4 reflects on the question of validity in qualitative evaluation. Qualitative research is sometimes seen as too subjective and biased (allowing too many degrees of freedom in causal inferences and construction of conclusions) which of course has its problems when applied to evaluation purposes in social policy. Yet one can equally argue that the increased importance given to qualitative information in policy evaluation simply requires a more precise conceptualization of validity criteria that can be linked to evaluations purposes, and in the case study I explore the possibilities to this end.

Case study 1: a mixed methods approach
Case study 1 reports on the effects of the Social Support Act (Wmo) in the Netherlands, implemented in 2007. The Dutch Act resembles the current European political interpretation of citizenship: stressing self-responsibility in one's personal life, fighting against elements of the welfare state that are believed to be invasive and implying a shared responsibility of government and civil society for the care of socially vulnerable groups. Data were collected on the basis of two surveys evaluating the effects of the Wmo for voluntary organisations and professional non-profit institutions in social care and welfare (2007: $N = 383$ and 2009–10: $N = 389$). In addition, in-depth interviews with stakeholders and qualitative case studies were carried out. The study yields several paradoxical policy outcomes. Contrary to the objectives of the Social Support Act, a 'revitalisation' of the Dutch civil society – in terms of a greater contribution to social goals and policies – remains problematic, whilst professional

entities thrive under the new governmental élan. Other paradoxical outcomes stem from a too-dogmatic approach to the social participation of people with severe mental disabilities. Instigating the socialisation of these groups through mandatory measures can in practice increase their isolation. To reduce unintended effects, I argue that the Social Support Act should take into account the divergent capacities of vulnerable groups and prioritise the psychological safety of clients over political and administrative policy objectives.

Case study 2: a theory driven approach
Case study 2 evaluates the validity of 'theories of change' of anti-polarisation and anti-radicalisation interventions. Assumptions of four dominant social policies in the Netherlands are confronted with the literature: social ecological interventions (1), peer mediation (2), intergroup contact interventions (3), self-esteem enhancement (4). The most defining underlying principle of a particular form of intervention was extracted along with the postulated relationships between the assumed 'working' components. For this purpose, generalisations (ideal types) of the various forms of interventions and their underlying assumptions were constructed. The programme theories were reconstructed by analysis of written documents, such as programme descriptions of professional and civic organisations, government position papers and strategy outlines of funding agencies. To review the theories of change, the international literature was systematically searched. The scientific resources and impact assessments had to be related to one of the four themes and the assumed underlying causal assumptions. The literature yielded over 150 relevant international studies and impact assessments. None of the investigated policy approaches is strongly supported by the literature, though some theories of change display a moderate scientific validity. In addition, epistemological issues are discussed. Notions of 'what works and why', do not equate to straightforward application in anti-polarisation and anti-radicalisation methods. Also, the potential of the four policy approaches depends in great measure on their conceptualisation of 'polarisation' or 'radicalisation'. A focus on interventions could be dismissed as naive, when no account is offered of wider sociopolitical factors fuelling radicalisation or ethnic tensions. I present the interaction between these two notions as a relevant topic for future research.

Case study 3: a second theory driven approach
Case study 3 evaluates the validity of 'theories of change' of social policies against neighbourhood disorder. In addition to physical measures and stricter law enforcement, deprived neighbourhoods in the Netherlands display many social and community-based interventions to tackle disorder and improve quality of life. The study evaluates the plausibility of such interventions by confronting them with insights from the scientific

literature. Four policy categories were assessed: promoting neighbourhood contacts (1), resident representation and participation (2), community-based behavioural codes programs (3), neighbourhood watch schemes (4), and neighbourhood sports (5). The program theories were reconstructed by coding-analysis of written documents, such as government policy papers, descriptions of quality of life programs from professional and civic organizations and strategy outlines of funding agencies. First, an inventory was made of all active Dutch social programs designed to tackle neighbourhood disorder (ranging into the dozens). To this end, both national and local government databases were searched as well as various other databases containing descriptions of neighbourhood programs. The respective programs were then lumped into more general themes and categories of which subsequently the most defining underlying principle of a particular policy category along with the postulated relationships between the assumed 'working' components was extracted. In other words, as in case study 2, generalizations (ideal types) of the various policy themes and their underlying theories were constructed, which then served as a directive for the literature assessment. The literature yielded circa 100 relevant international studies and impact assessments. According to the literature, only neighbourhood watch schemes can boast a strong scientific basis. I discuss implications for urban policy and theory, such as the need for a shift from residential networks (*horizontal* relations) to building organizational capacity (*vertical* relations) and the interrelation between instrumental and symbolic disorder measures.

Case study 4: a discussion on qualitative evaluation
Case study 4 provides a theoretical and methodological discussion of the question of validity in qualitative evaluation. Since roughly the 1970s increasing criticism of the reliability and objectivity of qualitative research has resulted in a growing interest in establishing more rigorous criteria and methodological standards. Within the academic community the idea seems to be dominant that qualitative researchers must demonstrate in one way or another that their research results are valid. Elaborating on epistemological and theoretical conceptualizations by Guba and Lincoln (1994) and Creswell and Miller (2000) the case study explores aspects of validity of qualitative research with the explicit objective of connecting them with aspects of evaluation. I will argue that different purposes of qualitative evaluation in social policy can be linked with different scientific paradigms and perspectives, thus transcending unproductive paradigmatic divisions as well as providing a theoretical framework for researchers and reviewers of qualitative evaluations.

Studying policy from multiple order perspectives
Policy can be studied from multiple order perspectives that shift from the micro to the macro level. A first order perspective assesses the policy directly and predominantly deals with instrumental questions; does it work? A second order perspective takes a step back and focuses on methodological issues; how can we best assess if the policy works? A third order perspective entails a (critical) reflexive focus; it problematizes the foundations of certain concepts of policy and reflects on what wider social, cultural, political, institutional and paradigmatic factors might influence its design, implementation and assessment.

The goal of case studies 1, 2 and 3 is to directly assess whether certain policies and the assumptions behind those policies are plausible (first order perspective). Yet like the market researcher observing how the audience rates the movie, one can also ask the question; how are we (better) to assess the plausibility of policy (second order perspective)? Drawing from the case studies, I would like to argue that more focus should be put on evaluating basic assumptions of social policy and specific programs through readily available and extant literature. Rather than relapsing straight into the 'measurement reflex' for each and every individual social intervention, it seems more appropriate to first visit the library. However, it is a misunderstanding that the policy theories under scrutiny should always be specified at length, describing all the treatment-as-delivered characteristics, other related exogenous factors, intervening processes and outcome variables. Instead, it is just as valuable to extract the most defining underlying principle of a particular form of intervention along with the postulated relationships between the assumed 'working' components. whilst looking at basic assumptions and applying qualitative methods, observational research deserves some reappraisal (see also Pope and Mays, 2009; Konnerup & Kongsted, 2012). Observations as a qualitative method can be used to 'test' mechanisms and assumptions of policy quite well. It is particularly suitable to refute certain assumptions. Finally, within the social domain it should be more clearly outlined what interventions are suitable for measuring effectiveness through the monocausal experimental model and what types of interventions are not. It is crucial that the social domain takes the differences in design and scope of different policies into account and allows for a more flexible application of research methods and definition of 'evidence'. Rigidly clinging on to the 'golden standard' of the RCT might otherwise produce a 'one size, fits few' system of valuation whereby some policies or programs – given their properties - will never reach the predicate of 'effective' even though they might work.

As for the third order perspective, Case study 1 showed that many of the paradoxical effects of the Social Support Act in effect stem from a problematic notion of 'normalisation'; the ideal that people with mental impairments should enjoy patterns and conditions of everyday living as close as possible to the norms and patterns of the mainstream society (Nirje, 1969; Yates et al., 2008). Normalisation *de facto* starts from the assumption that being 'different' is less desirable than being 'normal' and that disabled people should therefore strive to be something other than what they are (Morris, 1991). The study showed that this too-dogmatic idea of normalisation ultimately resulted in increased social isolation instead of participation. Case study 2 problematized in great measure the conceptualisation of 'polarisation' or 'radicalisation'. If these issues are conceived as treatable pathologies (as the four policy approaches discussed in the case study tend to do), then it makes sense to focus on treatment plans (i.e., 'interventions') to tackle radicalization or polarization issues. However, such an approach could arguably be dismissed as naive, when no account is offered of sociopolitical factors or perceived injustices fuelling radicalisation or ethnic tensions. Case study 3 elicits a debate on *instrumental policy* (policy as a means to an end) versus *symbolic policy* (policy as communicative strategy). Finally, case study 4 posed the question why certain preferences for particular paradigms and purposes for evaluation seem to correspond with different time periods and sectors. Sociological and socio-historical research not only can shed light on how and why such sectorial paradigm shifts occur; it could also investigate how these shifts influence ideas on what counts as 'evidence' of particular social programs or policies, and if or how this in turn influences ideas on the role of qualitative information in evaluation.

The thesis concludes with some general third order reflections. No matter how detailed a notion one can have of what factors influence the plausibility of policy, this does not mean that those insights can be applied automatically in newly designed policy interventions. Insight into the possibilities of social policy can thus paradoxically also lead to an understanding of its limitations. Also, one can question whether it is possible to make statements about the effectiveness of existing approaches solely by examining the literature, as case studies 2 and 3 do. Theory driven evaluation through extant literature provides a reference point of what works, but not a recipe. From the literature, insights emerge about effective elements based on empirical findings. Whether such potential effectiveness takes shape in practice depends on a variety of social, political and administrative constraints and contextual factors. However helpful, literature reviews by no means yield unequivocal policy verdicts. Finally, the plausibility of policy deals with only one aspect of the spectrum of policy analysis. A legitimate analysis of policy not only comprises its instrumental effectiveness (does it work?), but also its social acceptability (do people accept the measures at hand?), its

institutional feasibility (is the policy implementable?) and its constitutional legitimacy (does the law allow it?). I also highlight the symbolic function of policy; does it serve a rhetorical purpose for those in power? In sum, research outcomes are not the only legitimate source of policy, and the plausibility of the policy itself must always compete with social, institutional, constitutional and political aspects as well.

Presented themes for future research include the interaction between an 'intervention' approach against social problems and wider sociopolitical factors; the interrelation between policy theories and scientific theories (what mechanisms guide this interrelation?); paradigm shifts through the dimensions of time and space (why do certain preferences for particular paradigms and purposes for evaluation seem to correspond with different time periods and sectors?) and instrumental policy vs. symbolic policy (when do the scales tip more often in favour of one or the other and how does this (im)balance influence the design of social policy in particular situations?).

Samenvatting

Het proefschrift geeft een analyse van de plausibiliteit van beleid op basis van vier afzonderlijke case studies uit het sociale domein. Case studie 1 hanteert een meervoudige aanpak om te beoordelen of het bevorderen van maatschappelijk engagement en de participatie van kwetsbare groepen (zoals mensen met een verstandelijke handicap, psychiatrische patiënten) door de staat een passend streven is, en onthult daarbij een aantal van de 'donkere kanten' van sociaal beleid (reële en potentiële) die vaak over het hoofd worden gezien door politici en bestuurders. Case studies 2 en 3 hanteren een theoriegestuurde aanpak om de plausibiliteit van dominante aannames in anti-radicalisering / polarisatie en leefbaarheid in achterstandswijken te evalueren. Deze case studies hanteren in feite een tegenovergestelde aanpak van het meervoudige *mixed methods* onderzoek van case studie 1 door te focussen op fundamentele beleidsuitgangspunten (dat wil zeggen, ideaal typische veranderingstheorieën) en hen te confronteren met de wetenschappelijke literatuur. Met deze twee studies tracht ik aan te tonen dat een sterke verklaring van de aannemelijkheid van beleid kan worden verstrekt indien ze zich richt op theorieën die niet te veel poneren, zoals het beroemde reductionistische principe van het 'scheermes' van Ockham. Ten slotte reflecteert casus 4 op de geldigheid van kwalitatieve evaluatie. Kwalitatief onderzoek wordt soms gezien als te subjectief en vooringenomen (waardoor te veel 'vrijheidsgraden' in causale gevolgtrekkingen en constructie van conclusies zouden ontstaan) hetgeen natuurlijk problemen geeft wanneer kwalitatief onderzoek wordt toegepast voor evaluatiedoeleinden van sociaal beleid. Men kan echter eveneens stellen dat het toegenomen belang dat aan kwalitatieve informatie in beleidsevaluatie wordt gehecht, vraagt om een meer precieze conceptualisering van validiteitscriteria die kunnen worden gekoppeld aan evaluatiedoeleinden, en case studie 4 onderzoekt de mogelijkheden hiertoe.

Case studie 1: een meervoudige benadering
Case studie 1 rapporteert over de effecten van de Wet maatschappelijke ondersteuning (Wmo) in Nederland, geïmplementeerd in 2007. De Nederlandse wet haakt aan op de huidige Europese politieke interpretatie van burgerschap; nadruk op de eigen verantwoordelijkheid in het persoonlijke leven, het aanvechten van elementen van de verzorgingsstaat staat die worden verondersteld te invasief te zijn en een gedeelde verantwoordelijkheid van de overheid en het maatschappelijk middenveld voor de zorg van sociaal kwetsbare groepen. Er werden gegevens verzameld op basis van twee survey-onderzoeken bij vrijwilligersorganisaties en professionele non-profit instellingen in de sectoren zorg en welzijn (2007: N = 383 en 2009-10: N = 389).

Daarnaast werden diepte-interviews met stakeholders en kwalitatieve case studies uitgevoerd. De studie presenteert verschillende paradoxale beleidsuitkomsten. Tegen de doelstellingen van de Wet maatschappelijke ondersteuning in, blijft een 'revitalisering' van het Nederlandse maatschappelijk middenveld - in termen van een grotere bijdrage aan de sociale doelstellingen en het beleid - problematisch, terwijl professionele instituties gedijen onder het nieuwe gouvernementele elan. Andere paradoxale uitkomsten komen voort uit een te dogmatische benadering van de maatschappelijke participatie van mensen met een ernstige verstandelijke handicap. Stimulering van de socialisatie van deze groepen door middel van verplichte maatregelen kan in de praktijk leiden tot meer sociaal isolement. Om onbedoelde effecten te verminderen, beargumenteer ik dat de Wmo rekening moet houden met de uiteenlopende mogelijkheden van kwetsbare groepen en de prioriteit moet liggen bij de psychologische veiligheid van cliënten in plaats van bij politieke en bestuurlijke beleidsdoelstellingen.

Case studie 2: een theoriegestuurde benadering
Case studie 2 evalueert de geldigheid van 'veranderingstheorieën' van anti-polarisatie en anti-radicaliseringsinterventies. Veronderstellingen van vier dominante beleidsbenaderingen in Nederland worden geconfronteerd met de literatuur: sociale ecologische interventies (1), *peer-mediation* (2), intergroepscontact (3), mentale weerbaarheidsversterking (4). Voor de toetsing werd het meest bepalende uitgangspunt van een bepaalde vorm van interventie geëxtraheerd samen met de gepostuleerde relaties tussen veronderstelde 'werkende' componenten. Hiertoe werden generalisaties (ideaaltypen) van de verschillende vormen van interventies en hun onderliggende aannames geconstrueerd. De programmatheorieën werden gereconstrueerd door analyse van schriftelijke documenten, zoals programmabeschrijvingen van maatschappelijke organisaties, overheid *position papers* en strategische beleidsdocumenten van financieringsinstellingen. Ter toetsing van de veranderingstheorieën werd de internationale literatuur systematisch doorzocht. De wetenschappelijke bronnen moesten kunnen worden gerelateerd aan een van de vier thema's en/of de veronderstelde onderliggende causale veronderstellingen. De literatuur leverde meer dan 150 relevante internationale studies en effectstudies op. Geen van de onderzochte veranderingstheorieën wordt sterk ondersteund door de literatuur, hoewel sommige theorieën kunnen bogen op een matige wetenschappelijke validiteit. Noties van 'wat werkt en waarom' zijn bovendien niet als vanzelf toe te passen in anti-polarisatie en anti-radicalisering methoden. Ook het potentieel van de vier beleidsbenaderingen hangt in grote mate af van hun conceptualisering van 'polarisatie' of 'radicalisering'. Een focus op interventies zou kunnen worden gekenschetst als naïef, wanneer geen rekening wordt gehouden met bredere sociaal-

politieke factoren die kunnen aanzetten tot radicalisering of etnische spanningen. Ik presenteer de interactie tussen deze twee zaken (interventiegerichte aanpak vs. sociaal-politieke factoren) als een relevant onderwerp voor toekomstig onderzoek.

Case studie 3: een tweede theoriegestuurde benadering
Case studie 3 beoordeelt de geldigheid van 'veranderingstheorieën' van sociale beleidsmaatregelen op het terrein van buurtleefbaarheid. Naast fysieke maatregelen en een strengere rechtshandhaving, vertonen Nederlandse achterstandswijken veel sociale interventies om wanorde aan te pakken en de leefbaarheid te verbeteren. De studie evalueert de aannemelijkheid van dergelijke interventies door hen te confronteren met inzichten uit de wetenschappelijke literatuur. Vier beleidscategorieën werden beoordeeld: het bevorderen van bewonerscontacten (1), bewonersvertegenwoordiging en burgerbestuur (2), gedragscodeprojecten (3), buurttoezicht door vrijwillige buurtwachten (4), en wijksport (5). De programmatheorieën werden gereconstrueerd door het analyseren van schriftelijke stukken, zoals overheidsbeleidsstukken, beschrijvingen van leefbaarheidsprogramma's van professionele en maatschappelijke organisaties en strategiedocumenten van financieringsinstellingen. Eerst werd een inventarisatie gemaakt van alle actieve Nederlandse sociale programma's op het terrein van buurtleefbaarheid. Daartoe werden zowel nationale als lokale databases van de overheid doorzocht en verschillende andere databases met beschrijvingen van leefbaarheidsprogramma's. De respectieve programma's werden vervolgens samengetrokken tot meer algemene thema's en categorieën waaruit vervolgens het meest bepalende onderliggende principe van een bepaalde beleidscategorie samen met de vooropgestelde relaties tussen de veronderstelde 'werkende' componenten werd geëxtraheerd. Evenals in casus 2, werden generalisaties (ideaaltypen) van de verschillende beleidsthema's en hun onderliggende theorieën geconstrueerd, die vervolgens dienden als richtlijn voor de literatuurbeoordeling. De literatuur leverde circa 100 relevante internationale studies en effectstudies op. Volgens de literatuur kunnen alleen buurtwachten een sterke wetenschappelijke basis voorleggen. De wetenschappelijke onderbouwing van de overige interventievormen moeten worden gekenschetst als matig of zwak. Ik bespreek implicaties voor stedelijk beleid en theorie, zoals de noodzaak tot een verschuiving van residentiële netwerken (horizontale relaties) naar het opbouwen van organisatorische capaciteit (verticale relaties) in achterstandswijken en de verhouding tussen instrumentele en symbolische maatregelen op het gebied van buurtleefbaarheid.

Case studie 4: kwalitatieve evaluatie en validiteitseisen
Case studie 4 biedt een theoretische en methodologische bespreking van de kwestie van de validiteit in kwalitatieve evaluatie. Sinds ongeveer de jaren 1970 heeft een

toenemende kritiek op de betrouwbaarheid en objectiviteit van kwalitatief onderzoek geresulteerd in een groeiende belangstelling voor de vaststelling strengere criteria en methodologische standaarden. Thans lijkt binnen de academische gemeenschap het idee dominant dat kwalitatieve onderzoekers op een of andere manier moeten aantonen dat hun onderzoeksresultaten geldig te zijn. Voortbouwend op epistemologische en theoretische conceptualiseringen van Guba en Lincoln (1994) en Creswell en Miller (2000) onderzoekt de casestudie aspecten van de validiteit van kwalitatief onderzoek met de expliciete doelstelling deze te verbinden met aspecten van evaluatie. Ik betoog dat verschillende doeleinden van kwalitatieve evaluatie in sociaal beleid kunnen worden gekoppeld aan verschillende wetenschapsparadigma's en perspectieven. Zodoende kunnen onproductieve paradigmatische verschillen worden overbrugd, alsmede een theoretisch kader worden verschaft voor onderzoekers en beoordelaars van kwalitatieve evaluaties.

Beleidsbestudering vanuit meerdere perspectieven
Beleid kan worden bestudeerd vanuit meerdere perspectieven, die verschuiven van het micro- tot macroniveau. Een eersteordeperspectief beoordeelt het beleid rechtstreeks en houdt zich hoofdzakelijk bezig met instrumentale vragen; werkt het? Een tweedeordeperspectief doet een stap terug en richt zich op methodologische kwesties; hoe kunnen we het beste beoordelen of het beleid werkt? Een derdeordeperspectief hanteert een (kritische) reflexieve blik; het problematiseert de fundamenten van bepaalde beleidsbegrippen en reflecteert op welke bredere sociale, culturele, politieke, institutionele en paradigmatische factoren het ontwerp, uitvoering en beoordeling beïnvloeden.

Het doel van de case studies 1, 2 en 3 is om direct te beoordelen of bepaalde beleidsmaatregelen en de veronderstellingen achter dit beleid plausibel zijn (eersteordeperspectief). Maar vanuit het tweedeordeperspectief kan men ook de vraag stellen; hoe kunnen we (beter) de plausibiliteit van het beleid beoordelen? Puttend uit de case studies, pleit ik voor meer aandacht voor de evaluatie van de basisuitgangspunten van sociaal beleid en specifieke programma via direct beschikbare en bestaande literatuur. In plaats van direct in de 'metingsreflex' te schieten voor iedere afzonderlijke sociale interventie, lijkt het meer opportuun om eerst de bibliotheek te bezoeken. Het is echter een misverstand dat de beleidstheorieën die onder de loep worden genomen altijd tot in detail moeten worden gespecificeerd, met alle behandeling-als-geleverde kenmerken alsook andere gerelateerde exogene factoren, tussenliggende processen en uitkomstvariabelen. In plaats daarvan, is het evenzo waardevol het meest bepalende onderliggende principe van een bepaalde vorm van interventie te extraheren, samen met de veronderstelde relaties tussen de

veronderstelde 'werkende' componenten. Kijken we naar de uitgangspunten en toepassingen van kwalitatieve methoden, dan verdient observationeel onderzoek enige herwaardering. Observatie als kwalitatieve methode kan heel goed worden gebruikt om mechanismen en veronderstellingen van beleid te toetsen. Het is in het bijzonder geschikt om bepaalde aannames te weerleggen. Tot slot zal in het sociale domein duidelijker moeten worden aangegeven welke interventies geschikt zijn voor het meten van de effectiviteit middels een monocausaal experimenteel model en welke soorten interventies hier niet voor geschikt zijn. Het is cruciaal dat het sociale domein de verschillen in opzet en reikwijdte van verschillende typen beleidsmaatregelen erkent en zorgt voor een meer flexibele toepassing van onderzoeksmethoden en definitie van 'bewijs'. Star vasthouden aan de 'gouden standaard' van de RCT (*Randomized Controlled Trial*) zou anders een *one size, fits few* systeem van waardebepaling kunnen produceren waarbij bepaalde beleidsmaatregelen en programma's - gelet op hun kenmerken - nooit het predicaat van 'effectief' kunnen bereiken, hoewel ze wel zouden kunnen werken.

Betreffende het derdeordeperspectief laat case studie 1 zien dat veel van de paradoxale effecten van de Wet Maatschappelijke Ondersteuning voortvloeien uit een problematische notie van 'normalisering'; het ideaal dat mensen met verstandelijke beperkingen patronen en omstandigheden van het dagelijks leven moeten genieten die zo dicht mogelijk staan bij de normen en patronen van de gewone maatschappij. Normalisatie gaat in feite uit van de veronderstelling dat 'anders zijn' minder wenselijk is dan 'normaal zijn' en dat mensen met een handicap daarom moeten streven naar iets anders dan wat ze werkelijk zijn. De case studie toont aan dat dit al te dogmatische idee van normalisering uiteindelijk kan resulteren in een versterkte sociale isolatie in plaats van maatschappelijke participatie. Case studie 2 problematiseert de begripsvorming van 'polarisatie' of 'radicalisering'. Als deze problemen worden opgevat als behandelbare pathologieën (zoals de vier beleidsbenaderingen besproken in de case studie neigen te doen), dan is het zinvol om zich te concentreren op de behandelingswijzen (dat wil zeggen 'interventies') om radicalisering of polarisatieproblemen aan te pakken. Echter, een dergelijke aanpak zou kunnen worden afgedaan als naïef, indien geen rekening wordt gehouden met sociaal-politieke factoren of vermeende onrechtvaardigheden die radicalisering of etnische spanningen kunnen vergroten. Case studie 3 provoceert een debat over de instrumentele kant van beleid (beleid als middel om een doel te bereiken) tegenover symbolisch beleid (beleid als communicatieve strategie). Ten slotte werpt casus 4 de vraag op waarom bepaalde voorkeuren voor bepaalde paradigma's en evaluatiedoeleinden lijken te corresponderen met verschillende tijdsgewrichten en sectoren. Sociologisch en sociaalhistorisch onderzoek kan niet alleen licht werpen op hoe en waarom dergelijke sectorale

paradigma verschuivingen optreden; het kan ook onderzoeken hoe deze veranderingen van invloed zijn op ideeën over wat telt als 'bewijs' van bepaalde sociale programma's of beleid, en of en hoe dit op zijn beurt ideeën beïnvloedt over de rol van kwalitatieve informatie in evaluatie.

Het proefschrift sluit af met enkele algemene derdeordereflecties. Men kan nog zo'n gedetailleerd beeld hebben van welke factoren de plausibiliteit van het beleid beïnvloeden, dit betekent niet dat die inzichten automatisch kunnen worden toegepast in nieuw te ontwerpen beleidsinterventies. Inzicht in de mogelijkheden van sociaal beleid kan paradoxaal genoeg dus ook leiden tot een beter begrip van zijn beperkingen. Ook kan men zich afvragen of het mogelijk is om uitspraken te doen over de effectiviteit van bestaand beleid, uitsluitend via raadpleging van de literatuur, zoals case studies 2 en 3 doen. Theoriegestuurde evaluatie via gegeven wetenschappelijke kennis biedt een referentiepunt van wat werkt, maar niet een receptuur. Uit de literatuur vloeien inzichten voort over werkzame elementen op basis van empirische resultaten. Of dergelijke potentiële (plausibele) effectiviteit vorm krijgt in de praktijk is afhankelijk van een verscheidenheid van sociale, politieke, bestuurlijke en contextuele factoren. Hoe behulpzaam ook, literatuurstudies legitimeren niet per se absolute uitspraken over het beleid. Ten slotte gaat de plausibiliteit van het beleid over slechts één aspect van het spectrum van beleidsanalyse. Een legitieme analyse van beleid omvat niet alleen de instrumentele effectiviteit (werkt het?), maar ook de maatschappelijke aanvaardbaarheid (accepteren burgers de maatregelen?), de institutionele haalbaarheid (is het beleid uitvoerbaar?) en haar constitutionele rechtmatigheid (laat de wet het beleid toe?). Ik voeg hieraan toe dat ook de symbolische functie van beleid nadruk behoeft; dient het een retorisch doel voor bewindvoerders? Kortom, wetenschap is niet de enige legitimatiebron van beleid. De plausibiliteit van het beleid zelf zal altijd concurreren met sociale, institutionele, constitutionele en politieke aspecten van beleid.

Als thema's voor toekomstig onderzoek presenteer ik de interactie tussen een 'interventie' gerichte aanpak van sociale problemen en bredere sociaal-politieke factoren; het verband tussen beleidstheorieën en wetenschappelijke theorieën (welke mechanismen sturen hun verwevenheid?); paradigmaverschuivingen door de dimensies van tijd en ruimte (waarom corresponderen bepaalde voorkeuren voor paradigma's en evaluatiedoelen met verschillende perioden en sectoren?) en instrumenteel beleid versus symbolisch beleid (wanneer slaat de balans vaker om in het voordeel van het een of de andere? En hoe beïnvloedt deze (im)balans het ontwerp van sociaal beleid in specifieke situaties?).

CURRICULUM VITAE

Vasco Lub (1978) studied sociology at Erasmus University Rotterdam. After working for several institutes, he founded the Bureau for Social Argumentation in 2011, of which he is director and sole senior researcher. His research focuses on social policy analysis of various kinds, quality of life issues in urban neighbourhoods and methodological innovation of evaluation research in the social domain. Besides (co)authoring over 60 research reports and articles, and two books (*Schoon, heel en werkzaam*? (2013), *Kwalitatief evalueren in het sociale domein* (2014)), he has published academic papers in: Journal of Social Policy, Evidence & Policy and Beleid en Maatschappij.