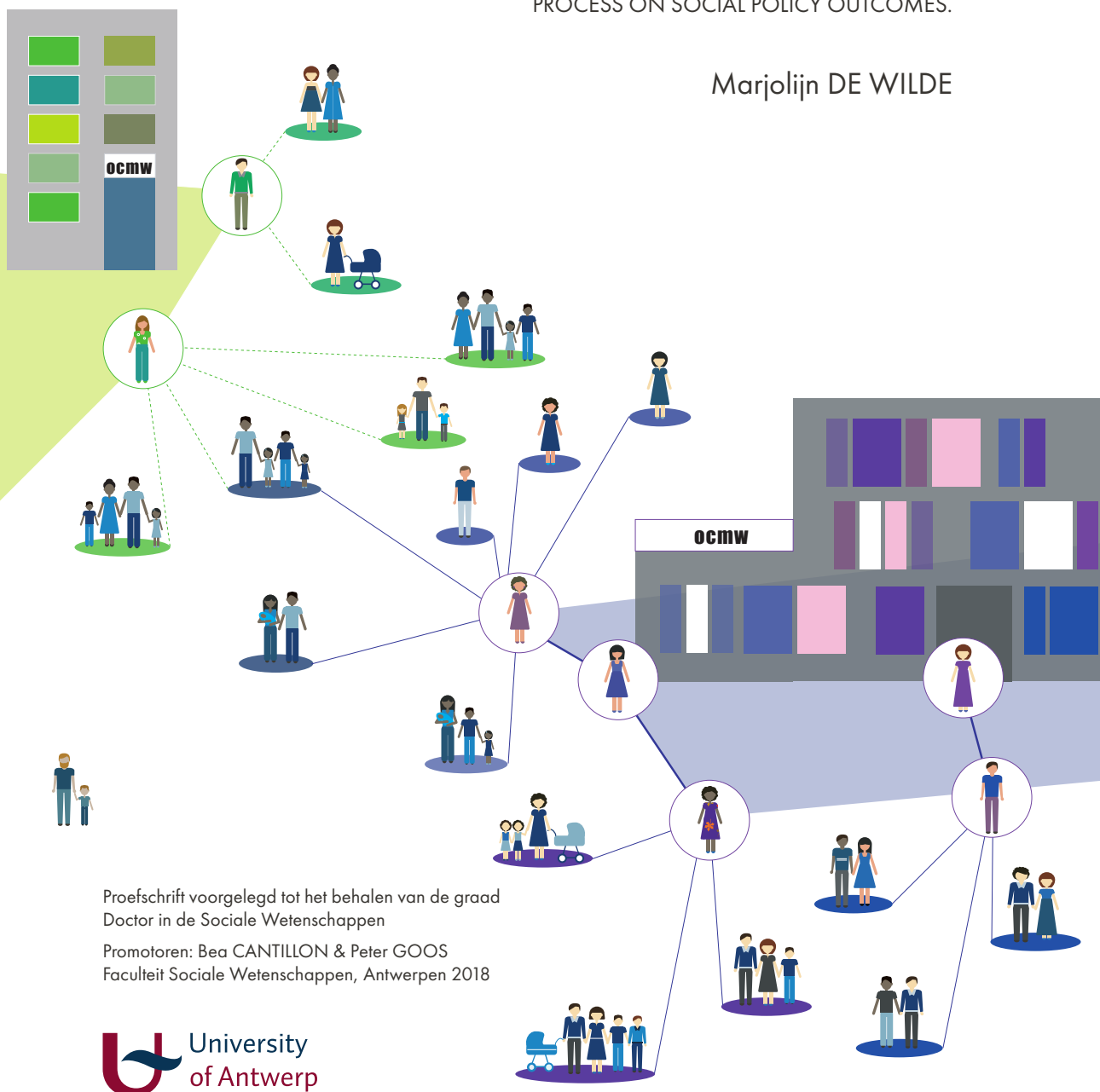


# Between legislation and realisation comes implementation

THE EFFECT OF THE MULTI-LAYERED IMPLEMENTATION PROCESS ON SOCIAL POLICY OUTCOMES.

Marjolijn DE WILDE



Proefschrift voorgelegd tot het behalen van de graad  
Doctor in de Sociale Wetenschappen

Promotoren: Bea CANTILLON & Peter GOOS  
Faculteit Sociale Wetenschappen, Antwerpen 2018



Voor mama, die het begin meemaakte, maar het einde niet.  
Voor Briek en Claire, die het begin niet meemaakten, maar het einde wel.

"T-shirt, de resultaten van de dessertjes, sushi en de levendigheid van de leeuw."  
(Briek's (9 years) translation of 'Third, the results in this dissertation suggest that leaving  
leeway...')



## Table of contents

---

<b>TABLE OF CONTENTS</b> .....	<b>5</b>
<b>LIST OF TABLES</b> .....	<b>7</b>
<b>FIGURES</b> .....	<b>8</b>
<b>ACKNOWLEDGEMENT / DANKWOORD</b> .....	<b>9</b>
<b>TUSSENCOVER</b> .....	<b>11</b>
<b>INTRODUCTION: BETWEEN LEGISLATION AND REALIZATION COMES IMPLEMENTATION.</b> .....	<b>12</b>
1. RESEARCH INTEREST .....	12
2. METHOD AND STUDY DESIGN .....	14
3. OVERVIEW OF THE CHAPTERS .....	15
<b>CHAPTER 1. THE IMPLEMENTATION OF SOCIAL POLICY: A CUSTOMISED FACTORIAL SURVEY APPROACH</b> <b>18</b>	
1. INTRODUCTION .....	19
2. EXISTING METHODS IN SOCIAL POLICY IMPLEMENTATION RESEARCH .....	21
3. THE FACTORIAL SURVEY APPROACH .....	24
3.1. <i>Strengths of the factorial survey with regard to social policy implementation</i> .....	26
3.2. <i>Additional guidelines for traditional factorial survey research</i> .....	26
3.3. <i>Limitations</i> .....	29
4. PROOF-OF-CONCEPT EXAMPLE .....	30
4.1. <i>Accounting for effects on three levels</i> .....	31
4.2. <i>The advantages of multi-level models</i> .....	34
4.3. <i>The advantages of stratified sampling</i> .....	37
5. CONCLUSION .....	40
<b>CHAPTER 2. IN A CATEGORY OF THEIR OWN? A MULTIGROUP SEM COMPARISON OF THE WELFARE STATE ATTITUDES OF SOCIAL WORKERS AND THE GENERAL PUBLIC</b> .....	<b>44</b>
1. INTRODUCTION .....	45
2. LITERATURE REVIEW .....	46
2.1. <i>Welfare state attitudes as a multidimensional concept</i> .....	46
2.2. <i>Comparing social workers and the general public: professional identification or selection?</i> .....	47
3. DATA AND METHODS.....	49
3.1. <i>Data</i> .....	49
3.2. <i>Variables</i> .....	49
3.3. <i>Modelling strategy</i> .....	51
4. RESULTS .....	51
4.1. <i>Descriptive results</i> .....	51
4.2. <i>Measurement equivalence and mean comparison: multigroup SEM with alignment</i> 52	
4.3. <i>Controlling for relevant covariates</i> .....	54
4.4. <i>Restricted sample: only Bachelors and Masters</i> .....	57
5. DISCUSSION AND CONCLUSION .....	58

**CHAPTER 3. WEIGHING UP WORK WILLINGNESS IN SOCIAL ASSISTANCE: A BALANCING ACT ON MULTIPLE LEVELS..... 64**

1. INTRODUCTION ..... 65

2. LITERATURE REVIEW ..... 67

    2.1 *The assessment of client characteristics* ..... 67

    2.2 *Devolution to the local level*..... 69

    2.3 *The discretion of the professional* ..... 70

3. THE BELGIAN SITUATION ..... 71

4. DATA AND METHOD ..... 73

    4.1 *Data* ..... 73

    4.2 *Method*..... 75

5. RESULTS ..... 79

6. DISCUSSION AND CONCLUSION ..... 83

**CHAPTER 4. DESERVINGNESS IN SOCIAL ASSISTANCE ADMINISTRATIVE PRACTICE ..... 90**

1. INTRODUCTION ..... 91

2. THEORY AND HYPOTHESES ..... 91

    2.1. *Deservingness opinions in social assistance legislation* ..... 92

    2.2. *Discretion in deservingness assessment* ..... 93

3. DATA AND METHODS ..... 96

    3.1. *Factorial survey and research setup* ..... 96

    3.2. *Analyses and hypotheses* ..... 97

4. RESULTS ..... 100

5. DISCUSSION AND CONCLUSION ..... 103

**CHAPTER 5. THE USE OF DISCRETION BY SOCIAL ASSISTANCE PROFESSIONALS: A QUANTITATIVE VERIFICATION ..... 106**

1. INTRODUCTION ..... 107

2. LITERATURE REVIEW ..... 108

    2.1. *Manifestation of de facto discretion*..... 108

    2.2. *Factors influencing inter-professional variation* ..... 110

3. THE BELGIAN SITUATION ..... 113

4. METHODS ..... 115

    4.1 *Data* ..... 115

    4.2 *Analyses* ..... 117

        4.2.1 *The models*..... 117

        4.2.2 *The independent variables*..... 121

5. RESULTS ..... 125

    5.1 *Studying the variation*..... 125

    5.2 *Modelling heteroscedasticity*..... 127

6. DISCUSSION..... 132

**CHAPTER 6. LEEFLOON KRIJGEN OF VERLIEZEN: HOE DECENTRAAL WORDT HET BIJSTANDSBELEID GEÏMPLEMENTEERD IN VLAANDEREN? ..... 138**

1. EXTENSIVE ENGLISH SUMMARY ..... 138

2. INLEIDING ..... 140

3. THEORETISCHE ACHTERGROND ..... 141

    3.1 *Leidt decentralisatie tot variatie in de implementatie van sociaal beleid?*..... 141

3.2	<i>Wat verklaart de variatie in begeleiding tussen verschillende lokale overheden?</i>	143
4.	METHODOLOGIE.....	144
4.1	<i>Opzet onderzoek</i> .....	144
4.2	<i>Analyses</i> .....	146
5.	RESULTATEN.....	148
5.1	<i>Verscheidenheid tussen gemeenten</i> .....	148
5.2	<i>Verklaringen voor de variatie</i> .....	150
6.	NAAR EEN MEER ADMINISTRATIEVE TOEKENNING VAN HET LEEFLOON? .....	153
<b>CHAPTER 7. CONCLUSION: BETWEEN LEGISLATION AND REALISATION COMES IMPLEMENTATION.....</b>		<b>158</b>
1.	THE FACTORIAL SURVEY: AN INNOVATIVE RESEARCH STRATEGY.....	161
2.	IMPLEMENTING THE WORK WILLINGNESS CONDITION .....	163
2.1	<i>The work willingness condition</i> .....	163
2.2	<i>Exemptions from work willingness</i> .....	164
3.	THE IMPORTANCE OF THE PROFESSIONAL LEVEL.....	165
3.1	<i>Legislation vagueness increases discretion</i> .....	166
3.2	<i>Variation as an inevitable result of the human process?</i> .....	167
3.3	<i>Can discretion be managed?</i> .....	168
3.4	<i>Awareness of the use of discretion</i> .....	170
4.	THE RELATIVE IMPORTANCE OF LOCAL POLICYMAKING .....	171
4.1	<i>Low implementation variation across municipalities</i> .....	171
4.2	<i>Does decentralization live up to its promises?</i> .....	173
5.	POLICY QUESTIONS.....	174
5.1	<i>Guaranteeing of human and civil rights</i> .....	174
5.2	<i>Guaranteeing the intended implementation of legislation</i> .....	176
6.	POLICY PROPOSALS .....	177
6.1	<i>Administratization of the eligibility test</i> .....	177
6.2	<i>Re-evaluation of discretion with regard to other forms of treatment</i> .....	179
7.	METHODOLOGICAL IMPLICATION: COMBINING THE STUDY OF OUTCOMES AND IMPLEMENTATION PROCESSES	180
<b>REFERENCES .....</b>		<b>181</b>
<b>APPENDICES .....</b>		<b>197</b>

### List of tables

TABLE 1-1. MULTI-LEVEL REGRESSION RESULTS WITH REGARD TO SANCTION LIKELIHOOD .....	34
TABLE 1-2. MULTI-LEVEL VARIANCE COMPONENTS WITH REGARD TO SANCTION LIKELIHOOD .....	36
TABLE 1-3. MULTI-LEVEL RESULTS FOR SELECTION OF MUNICIPALITIES AND PROFESSIONALS WITH REGARD TO SANCTION LIKELIHOOD .....	39
TABLE 2-1. OVERVIEW OF LATENT VARIABLES AND DESCRIPTIVE STATISTICS OF MULTIPLE INDICATORS .....	50
TABLE 2-2. DESCRIPTIVE RESULTS FOR THE INDEPENDENT VARIABLES .....	52
TABLE 2-3. ESTIMATES OF THE STRUCTURAL PART OF THE MODEL (SOCIO-DEMOGRAPHIC AND IDEOLOGICAL CHARACTERISTICS), LATENT MEANS AND INTERCEPT DIFFERENCES.....	56
TABLE 3-1. INDEPENDENT VARIABLES AND EXPECTED EFFECTS WITH REGARD TO SANCTIONING LIKELIHOOD .....	77

TABLE 3-2. MULTI-LEVEL RESULTS FOR PROFESSIONAL AND AGENCY/MUNICIPALITY CHARACTERISTICS WITH REGARD TO SANCTION LIKELIHOOD .....	81
TABLE 3-3. OVERVIEW OF AFFIRMED AND NOT AFFIRMED HYPOTHESES .....	86
TABLE 4-1: DESERVINGNESS PREDICTIONS BASED ON THE CLIENT DESCRIPTIONS .....	100
TABLE 4-2. MULTI-LEVEL REGRESSION RESULTS OF SOCIAL ASSISTANCE BENEFIT DESERVINGNESS .....	102
TABLE 5-1. OVERVIEW OF MODELS USED IN THE ANALYSES .....	121
TABLE 5-2. OVERVIEW OF RANDOMIZED CLIENT VARIABLES IN MODEL 2 WITH EXPECTED EFFECTS .....	122
TABLE 5-3.OVERVIEW OF RANDOMIZED PROFESSIONAL VARIABLES IN MODEL 3 AND EXPECTED EFFECTS.....	123
TABLE 5-4. OVERVIEW OF RANDOMIZED MUNICIPALITY VARIABLES IN MODEL 4 AND EXPECTED EFFECTS.....	125
TABLE 5-5. FREQUENCY TABLE OF LIKELIHOOD OF RECEIVING OR KEEPING SOCIAL ASSISTANCE BENEFITS .....	126
TABLE 5-6. RESULTS OF MULTI-LEVEL 0-MODELS .....	127
TABLE 5-7. OVERVIEW OF THE DIFFERENCE AT LEVEL 2 – VARIANCE BASED ON RANDOM SLOPES MODELS (RANDOMIZATION OF LEVEL 1 VARIABLES AT LEVEL 2) .....	129
TABLE 5-8. OVERVIEW OF THE SIGNIFICANCE OF RANDOMIZED PROFESSIONAL VARIABLES .....	130
TABLE 5-9. OVERVIEW OF THE SIGNIFICANCE OF RANDOMIZED MUNICIPALITY VARIABLES .....	131
TABLE 5-10. OVERVIEW OF CONFIRMED AND UNCONFIRMED HYPOTHESES .....	135
TABLE 6-1. OVERZICHT KENMERKEN OCMW EN GEMEENTE.....	148
TABLE 6-2. VARIATIE IN BEGELEIDING NAAR GELANG DE ACTOR (CLIËNT, MAATSCHAPPELIJK WERKER, OCMW) ...	149
TABLE 6-3. EFFECTEN VAN VERSCHILLENDE GEMEENTE EN OCMW-KENMERKEN .....	153

## Figures

FIGURE 1-1. OVERVIEW OF THE LEVELS DETERMINING SOCIAL POLICY IMPLEMENTATION .....	20
FIGURE 1-2. STUDY DESIGN AND SAMPLE .....	31
FIGURE 2-1: ALIGNMENT MODEL WITHOUT EXPLANATORY VARIABLES (MODEL 1).....	55
FIGURE 3-1. OVERVIEW OF THE LEVELS DETERMING SOCIAL POLICY IMPLEMENTATION.....	66
FIGURE 3-2. STUDY DESIGN AND SAMPLE .....	74
FIGURE 5-1. HYPOTHETICAL ILLUSTRATIONS OF WITHIN-GROUP AND BETWEEN-GROUP VARIATION. ....	109
FIGURE 5-2. EXAMPLE OF INTER-PROFESSIONAL VARIATION RELATED TO CLIENT CHARACTERISTICS .....	111
FIGURE 5-3. HYPOTHETICAL EXAMPLE OF INTER-PROFESSIONAL VARIATION RELATED TO AGENCY CHARACTERISTICS.....	113
FIGURE 5-4. STUDY DESIGN AND SAMPLE .....	116
FIGURE 5-5. BOX PLOTS OF PROFESSIONAL-SPECIFIC PREDICTIONS BASED ON MULTI-LEVEL 0 MODELS .....	127
FIGURE 5-6. RESIDUAL VARIANCE AT THE SECOND LEVEL BY THE RANDOMIZED VARIABLES .....	130
FIGURE 5-7. RESIDUAL VARIANCE AT THE SECOND LEVEL BY THE RANDOMIZED VARIABLES .....	132
FIGURE 6-1. DE GEMIDDELDE KANS OP HET VERLIEZEN VAN HET LEEFLOON IN 89 VLAAMSE OCMW'S.....	150
FIGURE 7-1. OVERVIEW OF THE VARIOUS LEVELS THAT DEFINE CONDITIONALITY OR DESERVINGNESS .....	159
FIGURE 7-2. OVERVIEW OF THE STUDY DESIGN .....	161



## Acknowledgement / dankwoord

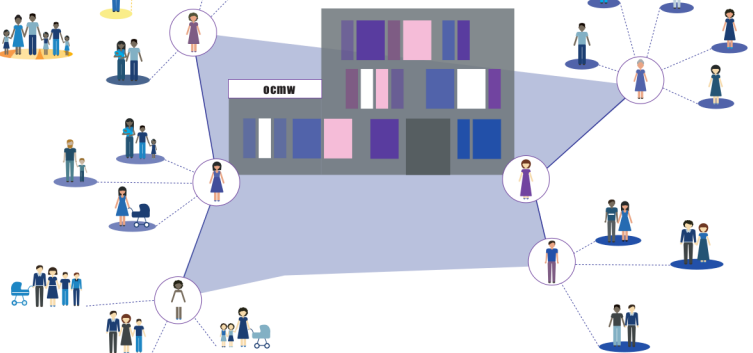
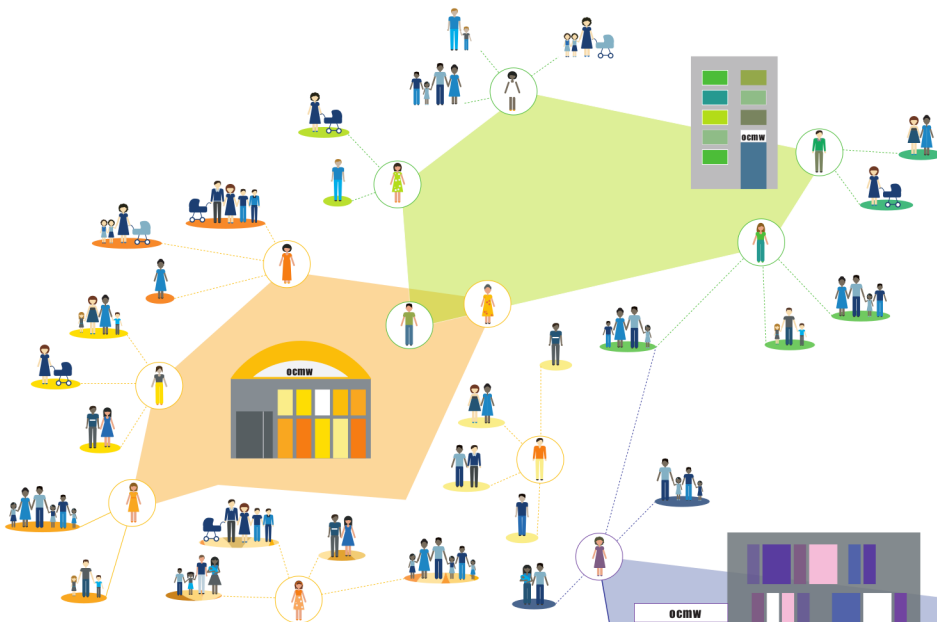
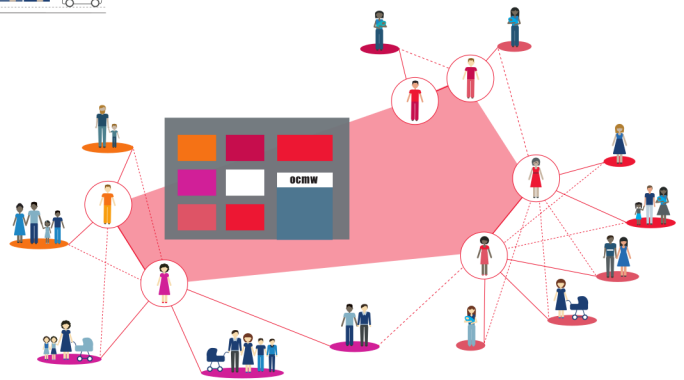
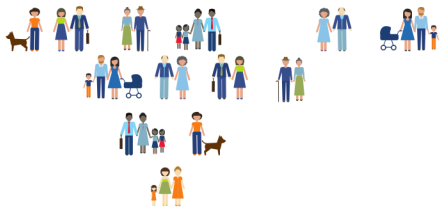
---

Een mens heeft niet veel nodig voor het schrijven van een doctoraat: een thema dat blijft boeien; een begeleider die met vijf woorden, twee gebaren en dé blik een proces op gang brengt (in haar kielzog wil je varen); een methodologische gids die (als hij niet aan het lopen is) met al zijn verstand ook gewoon naast je komt staan; een inspirator aan de zijlijn waarbij je met een klein hartje binnenloopt (te slim en te beslist) en bijwie je altijd net iets slimmer en beslist buitenwandelt; een blik op die andere wereld van wetten en rechtspraak (en lange anekdote-overladen gesprekken); een instantie die iets ziet in je vage ideeën (en die wil financieren of onderschrijven ook); proefpersonen die zich via een labeloos experiment aan je overleveren (en die je nadien het ‘grote gebruik van discretionaire ruimte’ kan aanwrijven); hun leidinggevendenden die zich makkelijk laten verleiden tot nieuwsgierigheid; mensen die je vertrouwen en je meenemen naar de plek waar ze hoopten nooit heen te moeten gaan (als je iets leert, dan is het daar); people who accept (and correct) your disabled English; iemand die de theoriën in beelden kan vatten (multi-level? dat teken je zó); mensen die schijnbaar blijven aanvaarden dat je langer op hun werkplek rondloopt dan gepland (en die maken dat je het debat over academische concurrentie niet begrijpt); liefsten die (hoewel ze niet állemaal graag gezelschapsspellen spelen) adem en ruimte zijn; natuurlijke compagnons de route waarvan je (omwille van de vanzelfsprekendheid) bijna zou vergeten dat als er een thema is dat boeit, als er begeleiders zijn die vertrouwen, als er een rijke vriendenkring is en als er fijngevoeligheid is voor de realiteit, dat dat dankzij hen is; een naaste die tegen wil en dank meer over dit doctoraat weet dan wie ook (en wiens kus de allerliefste is); en twee kinderen die denken dat dit het beste doctoraat ooit is (ssjt!).

Ik dank het fenomeen discretionaire ruimte, de promotoren Bea Cantillon en Peter Goos, de commissieleden Natascha Van Mechelen en Daniël Cuypers, de universiteit Antwerpen, het FWO en de VVSG, de 600 maatschappelijk werkers die deelnamen aan dit onderzoek, de meest betrokken OCMW-leidinggevendenden Mathias Vaes, Thomas Maesele, Marianne Decanne, Leen De Meyer, Koen Dresselaers, Dominique Vandevort en Guy Ongenaert, de OCMW-ervaringsdeskundigen Aurélie, Joke, Kenny, Kim, Mustafa, Tim, Tijn, Valérie en Zaki, het vertaalbureau Linguapolis en in het bijzonder Samantha Brunt, het designbureau Studio Ravie, de collega’s Leen, Sarah, Sarah, Josephine, Tess, Bérénice, Tim, Gerre, Wim, Jeroen, Koen en Sintia, de vrienden Anne, Moira, Veerle, Eef, Roos, Ivo en Pieter, de papa Dirk, de mama Marie-Claire en de broer Jeroen, de echtgenoot Krist en de kinderen Briek en Claire.



# Introduction



It is every social policy maker's fear: to see her or his well-considered policy is being implemented differently from how it was planned. And still, this is also every policy maker's experience. A one to one policy implementation is virtually non-existent. Several factors contribute to this phenomenon. First of all, regulations or policy goals cannot be detailed in legislation and implementation procedures cannot control for all possible deviations. So, regulations and procedures provide room for implementation variation. Second, the extent to which sub-national levels (region, municipality, agency, team within the agency) are explicitly provided with policy-making and implementation discretion (Kazepov & Barberis, 2013) and the extent to which they use it (Minas e.a., 2014), might result in treatment variation across implementing teams, agencies and municipalities, and thus variation in the actual treatment of clients. Finally, as social policies are often implemented by local professionals, the degree of discretion offered to or used by these individual professionals is a third influential factor (Evans, 2012; Kazepov & Barberis, 2013; Lipsky, 1980). As long as implementation is (partly) done by humans, deviations from the actual purpose of legislation and variation across professionals and agencies will occur.

### **1. Research interest**

All aspects of this implementation process have been studied before. However, no study has managed to investigate the different levels within one study in a way that allows comparisons between levels. This is exactly what I do in this dissertation. In light of this, it makes a unique contribution to social policy, social work as well as public administration research. Several studies in these research traditions are focused on legislation and, based on this legislation, simulate what policies mean for individuals or proto-typical families. Other studies start from the actual outcome for citizens and clients and try to explain such outcomes by means of readily available administrative data (Bargain, Immervoll, & Viitamäki, 2012; Carpentier, 2016; Keiser, Mueser, & Choi, 2004; Soss, Fording, & Schram, 2011). Finally, a fair amount of research is concerned with implementation. In such studies the implementation processes are described, usually without considering the possible outcomes of these processes. The overall research conclusions concerning the various policymaking and implementation levels (without going into detail), are:

- (1) Client characteristics determine the treatment that clients receive (=similar across professionals and agencies);

- (2) The professional is of decisive importance for treatment (= the use of discretion);
- (3) the level of decentralization of the policy determines treatment variation

My project is unique in that it includes the main relevant levels (client, professional and agency / municipality) as explicators for treatment variation and treats the quantitative outcomes for clients as the dependent variables. The main hypothesis in this dissertation is that doing so might raise concerns with regard to the above premises, premises that previously have been taken for granted. As the different levels interact with one another, I expect the effect at each different level to be smaller or different than in case not all levels are taken into account. In the remainder of this section, I give some alternative hypotheses that guided my research.

First, I expect that client characteristics do indeed influence treatment, but not for every professional in the same way. In this professional specific assessment, conditionality and professionalism or discretion meet, with as a result that treatment outcomes for clients can not be predicted based on conditionality in legislation. Deservingness assessment and professional specific deservingness assessment is an important, complementary factor (De Wilde, 2017; van Oorschot, Meuleman, Roosma, & Reeskens, 2017).

Another sub-theme concerns the role that professionals or local circumstances and authorities play in implementation processes. Effects of decentralisation tendencies on treatment variety are, generally, expected to be high. An oft-neglected question is: what is driving this variety? Is it local socio-demographics and political choices, or rather the professionals with specific preferences in a certain region. The hypothesis in this manuscript is that both factors are equally important and that studying local treatment variety solely as an effect of the decentralisation of policy making to local governors is too limited.

A final example concerns the attention paid to the effect that organisational features (e.g., management arrangements) have on the use of discretion. It is commonly thought that increasing the control by managers will decrease the use of discretion by social policy professionals. Evans and Harris (2012; 2004) opposed this premise, and claimed that discretion remained present in strongly managerially controlled settings. However, these opposing views have been made relying on statements of professionals and managers. In the social policy field, this was never checked using, first, client level data, and second, quantitative data. It is only by including the three levels in one analytic model that such questions can be answered in a more objective way.

## 2. Method and study design

In order to study the effect of the three levels, I used an experimentally designed survey, namely a factorial survey. Factorial survey experiments require each respondent to read experimentally varied stories about individuals or situations (vignettes – level 1). After viewing the vignettes, respondents are typically asked to answer a set of rating questions. I asked the employees to rate nine vignettes concerning hypothetical clients with regard to the likelihood that these clients would receive social assistance benefits, and the likelihood that they would lose their benefits if they refused a job offer. I selected the vignette characteristics on the basis of existing literature and interviews with case managers, team leaders and academics, keeping the deservingness criteria in mind (Nybom, 2013; Taylor, 2012; van Oorschot e.a., 2017; Wallander, 2009) (see Table 4-1 for the client characteristics). The surveys were distributed to a stratified sample of 90 Flemish municipalities (level 3) among 640 stratified samples of employees involved in eligibility decision-making (level 2). In the first chapter, Peter Goos and I empirically show the usefulness of this experimental case study method to study social policy implementation in all its dimensions. We claim that, studying issues on conditionality, discretion, decentralization and, ultimately, international policymaking within the same research design, offers a unique and more complete view on the reality of policy implementation reality.

I used Belgium as a case. Social assistance in Belgium is regulated by federal law (RMI law) and implemented by local welfare agencies. One of the conditions for eligibility is the client's willingness to work but the RMI law does not detail how social workers should interpret willingness to work. It also refers to reasons for exemption from activation as 'health reasons' or 'fairness reasons', but no clearly described categories are included (Vande Lanotte, 2002a). Clearly, the RMI law affords important leeway to local regulations and preferences and the professional opinions and assessments of social workers. At the local level, each Belgian municipality has a social welfare agency headed by a politically composed non-professional board, which decides – within the framework of the law – on eligibility for financial assistance. Funding provision for the benefits is divided between the federal and local levels. The professionals who come into contact with the clients are always qualified social workers. In sum, the country's decentralization, rather vague descriptions of work willingness in legislation and the explicit recognition of the professional as a qualified professional with important discretion make Belgium an interesting case with which to look at the various determinants of eligibility (and their interaction).

### **3. Overview of the chapters**

As in the first chapter the focus is on the methodological requirements, chapter three makes a similar, but more content rich claim. Sarah Marchal and I give a descriptive overview of the importance of the three levels, with regard to one of the dependent variables, namely the risk of losing social assistance benefits in cases where clients refuse a job offer. First, we study how the actual implementation of the work willingness condition takes place in light of specific client characteristics, circumstances and behaviour. Second, we test whether the interpretation of such behaviour is similar across social assistance professionals and municipalities. In the introductory second chapter, co-authored by Bart Meuleman, Koen Abts and me, we zoom in on one of the the social workers' characteristic that is salient in explaining treatment variation found in chapter three, namely the welfare state attitudes of social workers. We use multi-group structural equation models to relate this important explaining characteristic of social workers relates to the same characteristic in the general public.

In the last three chapters, I study in more detail each of the three levels included in the project. First, in chapter four, I test whether eligibility for welfare benefits could be explained by deservingness assessment. The interaction between public policy and public opinion is a recurrent issue in the study of welfare deservingness (Swaan, Manor, Øyen, & Reis, 2000). However, very few studies link opinions about deservingness to the actual treatment of welfare claimants. Several studies provide insights into the treatment of claimants and their characteristics, which determine their eligibility for benefits. These analyses are nevertheless mostly based on official registration data (Bargain e.a., 2012; Carpentier e.a., 2014) that include static information, such as gender, age, nationality background, parenthood, housing situation, level of education and employment history. Important elements including attitude, motivation, mental health and addiction are seldom recorded officially, in spite of the fact that they align with the traditional CARIN deservingness criteria: control, attitude, reciprocity, identity and need (van Oorschot e.a., 2017). In the fourth chapter, I link this deservingness assessment to the predicted eligibility treatment of deserving and non-deserving clients.

In chapter five, I focus on the discretion used by professionals. Discretion is – as mentioned in the first section of this introduction – mostly studied qualitatively, based on statements by professionals or clients (Sandfort, 2000; Watkins-Hayes, 2011). Quantitative investigations predominantly rely on the degree of freedom that professionals feel they have when making treatment decisions (Tummers & Bekkers, 2014). As the main conclusion of Chapter 3 was that inter-professional variation is high, I explore in this chapter what influences the amount of variation (= the use of discretion):

professional characteristics (e.g. older professionals make more use of discretion than younger colleagues), organisation characteristics (e.g. in agencies with strict internal rules professionals use less discretion than in agencies with almost no extra regulations) or client characteristics (professionals agree more on clients without children than on clients with children). This novel approach of an old academic theme calls to supplement qualitative studies of organizations with a more quantitative approach that links organizational settings with client outcomes.

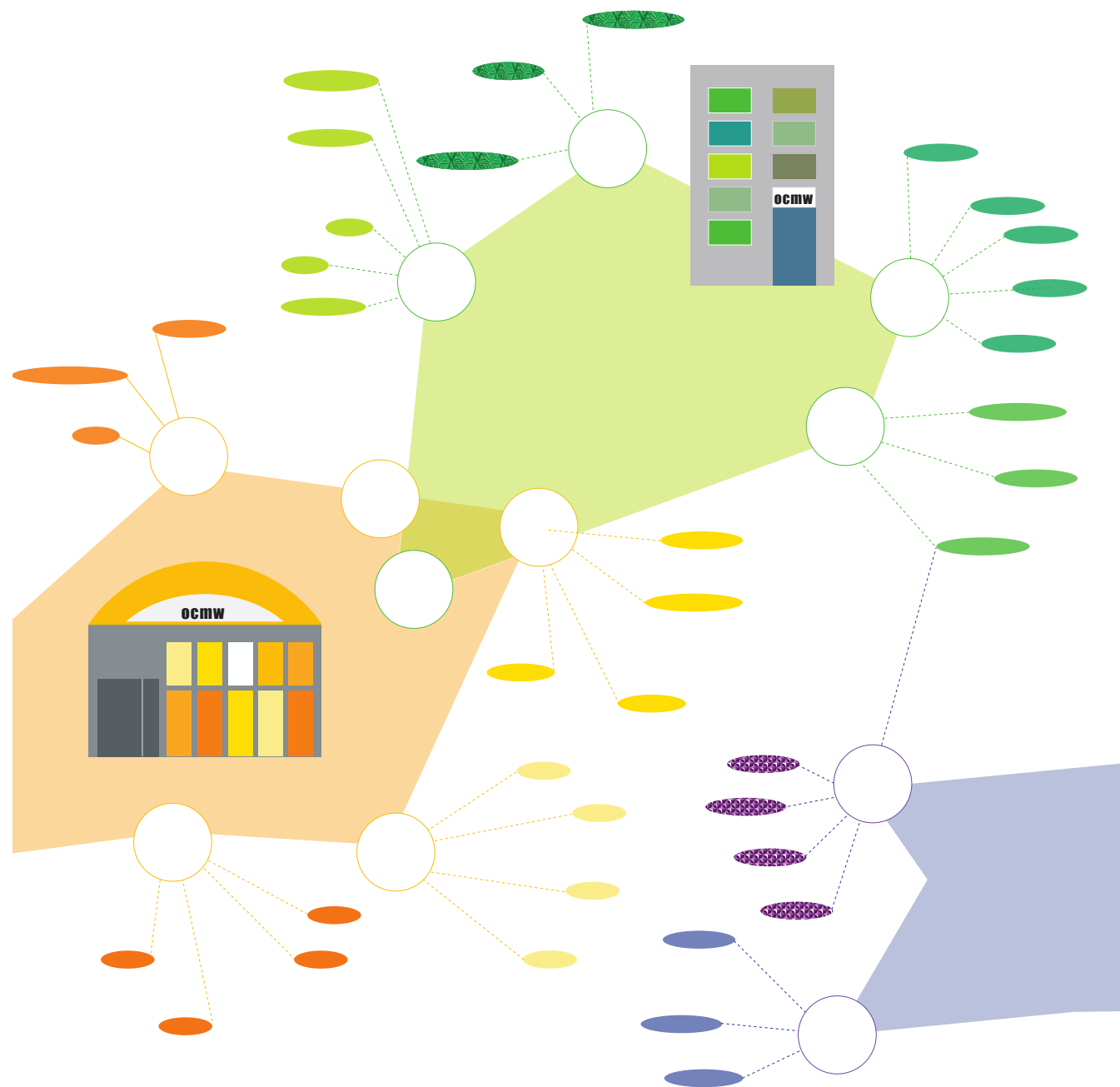
And finally, in chapter six, the level of decentralization is brought into context with passive subsidiarity. Decentralization is typically described as the intended shifting of policy-making power or budgets to local government levels. However, the term 'passive subsidiarity' refers to the spread of responsibilities over national and sub-national entities, in the absence of targeted resources or ideological preference for decentralised policy making (Bonoli & Trein, 2016). This chapter describes the feedback loops between traditional social security and social assistance in Belgium and examines the consequences of an intended or non-intended trend towards decentralization.

In the conclusion, I argue – based on the results of the previous chapters – that the integration of different viewpoints in social policy, social work and public administration calls into question premises that are taken for granted in these different fields of study. Hereby the focus is on the wide use of discretion, which – again – stresses the importance of social policy implementation. I invite academics and politicians to, at the one hand, – as they design or advice policies – be aware of the professional's use of discretion, and at the other hand to problematize and to cherish it. Policy proposals aim, first, at an eligibility check for social assistance benefits that is as administrative and as simple as possible; and second, at a reevaluation of discretion in the trajectory following the eligibility assessment.



# The implementation of social policy:

A CUSTOMISED FACTORIAL SURVEY APPROACH



## Chapter 1. The implementation of social policy: A customised factorial survey approach

---

*Marjolijn De Wilde and Peter Goos*

*Submitted to Journal for Social Policy*

### **Abstract**

The implementation of social policies has a multidimensional character. To properly understand implementation conditionality (client level), discretion (social worker and organisation level), decentralisation (municipality/region level) and international policymaking (country level) should be addressed. None of the existing methods is able to account for all these dimensions. Standard approaches rely on qualitative methods, with content rich descriptions of parts of the implementation process. These approaches fail to capture all relevant processes between legislation and actual treatment. Other researchers rely on administrative outcome data, that do not contain information about the actors carrying out the implementation.

We, therefore, present a new way of using the experimental factorial survey method. Factorial survey studies in social policy research are increasing. However, the focus is restricted to testing client specific eligibility conditionality or attitudes and preferences of professionals. We propose some adjustments to tailor it to the study of social policy implementation. The adjustments involve asking for expected rather than preferred treatment, adding a questionnaire about respondents and their organisation, and an adjusted way of using multi-level techniques for analysis. We are the first to empirically show that the factorial survey is appropriate to collect large scale datasets on several implementation levels. Our proof-of-concept example involves a survey among Belgian social assistance professionals, who were asked to predict the likelihood of experimentally varied hypothetical clients being sanctioned (level3=organisation [n=79]; level2=respondent [n=593]; and level1=client descriptions [n=4846]).

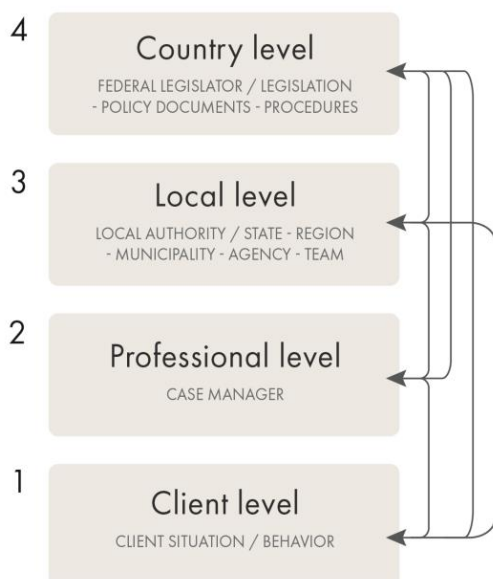
## 1. Introduction

The implementation of social policy legislation has a multidimensional character. It is contingent on several levels, for example, the country, the organisation, the social worker and the client (Priem, Walters, & Li, 2011; Rice, 2013) (see Figure 1-1). First of all, implementation depends on existing legislation (country) and the characteristics of clients who are deemed eligible for the specific policy (client) (De Wilde, 2017; van Oorschot e.a., 2017). Second, the extent to which decision-making power is decentralised to subnational levels (region, municipality, organisation, team in the organisation) and the degree to which these subnational levels use their policy-making and implementing discretion<sup>1</sup> result in treatment variation across implementing teams, organisations and municipalities, and thus on the actual treatment of eligible clients (Kazepov, 2010; Kazepov & Barberis, 2013). Furthermore, as social policies are often implemented by local professionals, the degree of discretion offered to or used by these individual professionals is a third overarching and influential factor (Evans and Harris 2004; Lipsky 2010). A final level that, while interesting, is not discussed further in this paper, is the European coordination of national policies. It is in this context that the term ‘multi-level-governance’ is often used to refer to the distribution of decision-making power over several actors both vertically (national, supranational and subnational) and horizontally (several actors at one level) (Benz 2000; Hooghe 1996; Hooghe and Marks 2003; Piattoni 2009; Stephenson 2013). It is noteworthy that, to the best of our knowledge, the decision-making power of individual professionals is seldom mentioned in the multi-level-governance literature.

---

<sup>1</sup> In this manuscript we understand the term ‘discretion’ as ‘individual’ or ‘locally based’ decision making. Such discretion can be studied in two different ways. First, it can be studied as the room that is allowed to lower-level actors (e.g. professionals) by higher management levels (e.g. managers). This sort of discretion is called *de jura* discretion by Evans (2012). However, the factorial survey only enables us to test the extent to which professionals differ in their treatment proposals and thus use their discretion, which is *de facto* discretion (Evans, 2012).

Figure 1-1. Overview of the levels determining social policy implementation



The aim of this paper is to show that it is particularly difficult to account for all of these levels using one research method. This is due partly to data-gathering limitations, but also to paradigm differences between research fields. In this paper, we demonstrate how factorial surveys have the potential to overcome methodological limitations, and, in doing so, provide opportunities to study the multi-level nature of social policy implementation.

In factorial surveys, respondents are presented with experimentally varied hypothetical stories and asked to judge the situations. We illustrate our claim that factorial surveys have much potential with an example in which almost 600 professionals from 79 social assistance agencies in Flanders (Belgium) were presented with unique sets of nine client descriptions. In this example, we focused on the question whether labour market activation is as intense among clients who have children, including sick children, as it is among clients who do not have children. We were able to demonstrate that the answer to this question depends both on the client characteristics in terms of ‘parenthood’ and on the characteristics of the professional treating the client, as well as on several characteristics inherent to the organisation and the municipality in which the professional works. Furthermore, we observed extensive variation between professionals from the same organisation and only little variation between organisations.

The remainder of the paper consists of three parts. The first part reviews the literature on the conditionality of social assistance, the decentralisation of decision-making power to subnational levels and individual professionals' discretion. We illustrate that, because one or more of the levels described above were ignored in published work, the conclusions in the literature are in fact biased. In the second part, we outline the factorial survey approach, arguing it offers the potential to overcome the shortcomings of other methods provided some adjustments are made to the traditional methodological guidelines. In the third part, we describe the first successful implementation of the customised factorial survey approach in social policy research.

## **2. Existing methods in social policy implementation research**

For a long time, policy makers and researchers believed that policies were decided upon at national level and implemented at lower levels as intended by the legislators. Any local influences have been neglected (Kazepov & Barberis, 2013; Tabin & Perriard, 2016). Two different evolutions have altered this perception. The first evolution is the growing importance of other policy-making levels (Europe, regions, municipalities, etc.). The second relates to research on street-level bureaucracy, which acknowledges the impact of street-level workers in policy implementation and in policy-making.

At present, two methods are commonly used to study and compare the national legislation and implementation of social assistance policies. The first method involves the study of national cash welfare expenditure and activation-related spending (e.g. active labour market programmes or ALMPs) (Champion & Bonoli, 2011). A comparison of the two types of spending provides information about governmental choices and, indirectly, about policy implementation. Although still widely used and valuable to a certain extent, this approach has serious limitations. We discern two main problems. The first concerns the classification of spending packages. Several studies have shown that, by classifying the same programs differently, the results change considerably (De Deken & Kittel, 2007; Kittel & Obinger, 2003). Second, these studies provide no information about how programs are actually implemented. The same program may be implemented entirely differently by different organisations or professionals.

The second method to compare national legislations involves interviewing country experts about standard families and about the interpretation of the country's legislation when it comes to families with a given composition and a particular set of problems (Eardley, Bradshaw, Ditch, Gough, & Whiteford, 1996; Kazepov, 2010; Kazepov & Barberis, 2013; Marchal, Marx, & Van Mechelen, 2014; Marchal & van Mechelen, 2017; Van Mechelen, Marchal, Goedemé, Marx, & Cantillon, 2011). This powerful method

provides insight into similarities and differences in legislation across countries and over time within countries. However, the method ignores the way in which policies are embedded locally. Marchal and Van Mechelen (2017) circumvent this problem in countries with high levels of local discretion by interviewing extra experts from local municipalities or states (e.g., a stakeholder in the city of Antwerp in the case of Belgium). Yet, the authors do conclude that a limitation of their study is that they need to rely on one particular case to draw conclusions about an entire country or region (Marchal & van Mechelen, 2017). In this sense, both methods have the drawback that it is difficult to study local embedment, which is particularly relevant. Some authors suggest, indeed, that, compared to other forms of policy, social policy is one of the domains in which the importance of the local level is high (Ripley & Franklin, 1982 as cited in Hasenfeld & Brock, 1991). This is mainly due to the fact that national legislation merely provides a framework and does not have immediate applicability. To a large extent, discretion is left to local organisations and its social workers (Wallander, 2012a).

This is where decentralisation comes in. The interplay between the national and local policy levels is very country-specific. Kazepov (2013) describes four different ways of organising social policy-making: “(1) countries with strong local autonomy which is centrally framed; (2) countries with a strong national/central frame; (3) countries with strong regional (or federal) frame; (4) countries with mixed frames in transition from one frame to another” (Kazepov & Barberis, 2013). Differences between these systems are notable both in the aspect of legislative power and in the degree of financial autonomy at each level. Besides Italy, there are very few countries in Europe without national framework (Kazepov & Barberis, 2013). Yet, the details of each set of regulations differ from country to country and may range from establishing the right to social assistance to detailed regulations concerning eligibility criteria, eligibility duration and sanction measures. Furthermore, the regulations may be implemented by locally embedded organisations that depend on the state level, or by individual municipalities themselves, at local level. Decentralisation dynamics are often studied with the aim of establishing how implementation is organised (Hölsch & Kraus, 2006; Minas, Whrighth, & Van Berkel, 2012; van Berkel, 2006). To date, however, very few studies have examined the effect of such decentralisation on actual implementation or actual client treatment (Carpentier, 2016; Fording, Soss, & Schram, 2007; Soss, Schram, Vartanian, & O’Brien, 2001).

Furthermore, recent history teaches us that, in many branches of social policy, the degree of discretion organisations and professionals are permitted to use has increased due to the new focus on integration and activation rather than on the provision of

financial support (Kazepov & Barberis, 2013; Vando Borghi & van Berkel, 2007). In order to decide whether a client is in financial need, a social worker is required to perform a means test. The norms for this test can be established at national level. However, decisions about clients' integration trajectories are likely to be tailored more closely to the needs and abilities of the client in question. National guidelines concerning activation are more difficult to formulate than norms for means tests. Consequently, the allocation of resources and activation measures depends to a large extent on the individual social worker (Lipsky, 2010).

Several methods are currently being used to study the implementation of social policies by professionals and local organisations. The most straightforward of these is the observation method, as it is less biased by the professional's or client's view than survey research. A disadvantage of observation, however, is that it is very time-consuming, meaning that researchers need substantial human and financial resources to overcome the main weakness of existing studies, which is that they typically involve only a few municipalities (Meyers, Glaser, & Donald, 1998; Thoren, 2008a).

More generally, most studies that address the territorial dimension, by interviewing either field workers or clients themselves, are qualitative in nature. When the interviews are held face-to-face and the questions are open-ended, a substantial amount of information can be collected (Bergmark, Bäckman, & Minas, 2017; Hermans, 2005; Moller & Stone, 2013; Nybom, 2013). This information can result in meaningful hypotheses about the reasons behind specific ways of implementation, but – because it is not quantitative – it does not allow for these hypotheses to be tested. Qualitative research on social policy implementation is limited to comparisons of agencies or municipalities within single countries (Bergmark e.a., 2017; Carpentier, 2016; Hermans, 2005) or among small numbers of municipalities in multiple countries (Evans, 2007; Saraceno, 2002). Therefore, no general view exists on the treatments received by clients.

Some studies have collected quantitative local data, however. The most promising of these relied on registered data. Bargain and colleagues (2012) used a Finnish survey to measure the use of social assistance. The survey in question collects administrative data on 0.5% of the population of Finland. It allows the researchers to carry out a systematic check of how many citizens with low incomes are receiving benefits. Carpentier, Neels and Van den Bosch (2014) used a Belgian database that collects registration data from various institutions, including information on demographics, unemployment benefits, social assistance benefits and health insurance. They had access to five years of data on one third of the individuals who started receiving social assistance benefits in 2004.

Fording et al (2007) and Soss et al. (2001) investigated administrative datasets on welfare sanctioning in one American state. This type of data makes it possible to follow a large number of clients over long assistance pathways, without having to interview them multiple times. It is, however, unclear whether the data registered is suitable for investigating all aspects of policy implementation. Information about clients' motivation and life experiences or about professionals' decision making is often missing. Nybom (2013) has attempted to overcome this shortcoming by combining administrative data (400 client records from four municipalities) with interviews. Again, this is a very time-consuming method and does not permit extensive general investigations.

Another problem with large-scale registration or survey data is that the selection bias caused by differences in client profiles across municipalities impedes the comparison of local treatment policies. A strong activation practice in a certain municipality may be caused either by local policy or by specific claimant characteristics. Moreover, certain types of clients might never occur in certain municipalities, thus hindering comparison across municipalities (Blommesteijn, van Geuns, Groenewoud, & Slotboom, 2012). Finally, most registration data are country- or state-specific, which prevents comparisons between countries and states.

It should be clear that each of the existing methods has major weaknesses. We believe that using the factorial survey approach would mean a step forward, because it overcomes some of the shortcomings of the approaches described above.

### **3. The factorial survey approach**

Factorial surveys incorporate the positive features of experimental research. In a traditional comparative experiment, the effect of one factor on a dependent variable is tested. In a factorial experiment, however, at least two factors are varied at once, enabling the researchers to investigate the effect of multiple factors and their 'interaction effects' on the dependent variable<sup>2</sup>. Each factor has two or more levels or categories (e.g. female and male for gender). When all possible combinations of all levels across all factors are considered, a full factorial design is obtained. For example,

---

<sup>2</sup> The term 'interaction effect' refers to a situation in which the impact of one factor depends on the level of one or more other factors. For example, overall, offensive behaviour may have a minor impact on the frequency of support. However, if we study the interaction between offensive behaviour and gender, it might be that women who exhibit offensive behaviour are supported as frequently as other women, but that men who exhibit offensive behaviour are supported significantly less frequently than other men. Thus, offensive behaviour has an impact on support treatment, but only among men. In that case, we say there is an interaction effect between the variables offensive behaviour and gender.



when an experiment consists of three factors and all these factors have two possible levels, then the full factorial design consists of eight possible factor-level combinations ( $2 \times 2 \times 2$ ).

In a factorial survey, the experiment requires each respondent to read a story about a hypothetical person or situation and rate the person or the situation according to well-defined dependent variables. This story or situation is usually referred to as a vignette (see Appendix 1 for an example of a vignette and related dependent variables). Like every experimental test in a factorial experiment, every story involves only one level of each factor. In some cases, all respondents rate all possible stories (= the full factorial universe). Most commonly, however, a sample is taken from the vignette population. Samples can be taken at random, but a more appropriate approach – especially if few respondents are involved in the survey – is to select a D-efficient sample. A D-efficient sample ensures that maximum information is obtained about the effects of the experimental factors. It is possible to construct a D-efficient sample for a model with main effects only, but it is also possible to create one for a model with interactions. If there are no practical constraints in the selection of a sample, then the D-efficient sample consists of an orthogonal design in which, for each factor, the levels occur equally often (level balance). An orthogonal design ensures that a linear regression or ANOVA model relating the factors to the ratings does not suffer from multicollinearity (Atzmüller & Steiner, 2010; Auspurg & Hinz, 2014; Dülmer, 2007, 2016) and allows the impact of the experimental factors on the dependent variable to be quantified with maximum precision. Practical constraints, such as the fact that certain combinations of factor levels are unrealistic and can therefore not be used in the survey, or the fact that only limited numbers of vignettes are practically feasible, may make it impossible to construct a perfectly orthogonal design. The D-efficient survey design approach is flexible, however, in the sense that it allows for the generation of a survey with maximum information content under the given constraints. The resulting D-efficient sample is then as close as possible to being orthogonal, for main effects as well as for interactions.

Typically, a factorial survey is sent out to various respondents, and each respondent evaluates only a subset of all of the vignettes in the D-efficient sample. An attractive feature of the D-efficient survey design approach is that it allows researchers to determine the best possible allocation of vignettes to respondents, in the sense that the influence of both the effects of the experimental factors and the effects of the respondent characteristics can be quantified with maximum precision. This allocation of

vignettes to respondents is known as 'blocking' in the experimental design literature (see, e.g., Goos & Jones, 2011).

### 3.1 Strengths of the factorial survey with regard to social policy implementation

Using the factorial survey approach has the potential to solve some of the problems in social policy implementation research listed above. It makes it possible to target a large number of respondents in multiple municipalities (and countries), thus allowing for generalisation without ignoring local processes (Aguinis & Bradley, 2014). Further, the possibility to manipulate the independent variables (i.e. the vignette attributes) and measure their independent effects on the dependent variable allows for conclusions about causal relationships. Lastly, randomisation of the assignment of vignette blocks to respondents prevents selection bias. This might be compared to studies on registered real data where selection bias constitutes a problem, as in reality some client problems might be over- or underrepresented in certain areas. The possibility to randomise cases over professionals and over organisations thus makes local policies truly comparable.

Studying social policy implementation is related to what is called 'measuring professional judgment' (Jasso, 2006; B. J. Taylor, 2006; Wallander, 2012a), but it goes beyond the main research topic in the professional judgment literature, which is the way in which social workers make decisions in their everyday practice. While the primary aim of professional judgment research is to unravel professional agreements concerning specific cases or problems, social policy implementation research also aims to account for professional disagreement within organisations or countries and across organisations and countries. One of the most important authors in this field, Wallander (2012a), has made a case for studying professional disagreement too, but this has not yet been picked up in existing research. At present, in the literature, we observe an intense use of the factorial survey approach to examine client characteristics as the cause of specific treatments (Morley, 2010; Ortega, Baz, & Sánchez, 2012; Skarlicki & Turner, 2014; Stokes & Schmidt, 2012; Webster, O'Toole, O'Toole, & Lucal, 2005), while characteristics of client managers and agencies are largely overlooked. An adjusted use of the factorial survey would, however, help to obtain a fuller picture of social policy implementation. Below, we outline four valuable additions that we suggest to the existing guidelines of Wallander (2012a) and Taylor (2006).

### 3.2. Additional guidelines for traditional factorial survey research

First, the dependent variables (or at least one of them) should relate to plausible acts or treatments (e.g. starting an activation programme with a client) and not merely to an opinion about the situation of the client (e.g. the client is ready for the regular labour market). Further, these treatments should be framed in the context of the organisation,

meaning that the respondent is asked to disclose what should happen to a certain client in the organisation where she is working. The respondent should not be asked what she would personally prefer or what she would personally do, but should be encouraged to consider all elements that could influence the decision concerning treatment. We are not aware of a single study that has framed the survey in this broader perspective on treatment. In our view, this means that existing studies neglect the decision-making process in organisations, a process which is rarely straightforward. A teacher's decision, for example, to report the abuse of a pupil by her parents might be influenced by the situation surrounding the abuse (vignette attributes) and by the intuitions of the teacher (= respondent), but also by the fact that the teacher must discuss her decisions with a supervisor. If the respondent is not asked to take her organisation's decision-making process into account, the researcher runs the risk of missing the information she is looking for.

Secondly, in order to be able to account for local decision-making processes or policies, researchers should include a questionnaire about the respondent and the organisation where she works (Wallander, 2012a). This questionnaire should collect information about the partners involved in making decisions about the treatment (Degenholtz, Kane, Kane, & Finch, 1999; Wallander & Blomqvist, 2008). This may simply be the respondent herself, but could potentially involve the head of the team, the entire team or an external committee. Such decision structures reveal a great deal about the discretion a social worker has in deciding what treatment is offered. The structure can vary substantially depending on the organisation in question. It should be noted that, whereas the factorial survey enables collecting comparable data on the vignette / client level, the collection of data on the respondent or organisation level suffers from the same comparison problems as traditional surveys. Several factorial survey studies in areas other than professional judgment indeed included independent variables other than the vignette characteristics (Auspurg, Hinz, & Sauer, 2017; Byers & Zeller, 1998; Jasso & Milgrom, 2008; Shlay, 2010), but such studies in the field of professional judgement or directed to social policy implementation research are scarce (Baughman e.a., 2012; Degenholtz e.a., 1999; Mion e.a., 2010; Wallander & Blomqvist, 2008).

The third valuable addition to existing guidelines would be to perform appropriate sampling of respondents so that the discretion used by an organisation and/or its social workers can be measured properly (Aguinis & Bradley, 2014). The theoretical literature on professional judgment research draws a range of conclusions about sampling. Taylor (2006) suggests taking random samples from the respondent population. Auspurg and Hinz (2014) distinguish between studies tailored at describing universal mechanisms and

studies accounting for diversity. They propose convenience samples for the former type of study and population wide sampling for the latter. Wallander (2012a) pays more attention to the organisational level and suggests sampling multiple respondents from several workplaces to be able to account for disagreement in treatment practice across and within organisations. As external validity is more important in social policy implementation research than in professional judgement research, respondents should be distributed evenly over the region being investigated and it should also be possible to compare agencies and respondents (with different characteristics) (Auspurg & Hinz, 2014; Byers & Zeller, 1998). This means that a stratified sample is required (Bryman, 2012; Groves e.a., 2009; Heeringa & Berglund, 2010). Cluster samples can be chosen to facilitate the research, as it might be impossible to sample respondents in all existing agencies. Researchers considering the use of a factorial survey to investigate social policy implementation will mostly opt for a multi-stage sampling technique that combines clustered and stratified sampling on several levels (Webster e.a., 2005). While population based random sampling is frequently used in existing research, it remains uncommon to consider the local level by clustering or stratifying.

Fourth, multi-level regression or ANOVA models are appropriate for quantifying and testing how respondents' ratings relate to the experimental factors and to respondent and agency characteristics. Multi-level modelling takes into account statistical dependencies between ratings given by the same respondent and ratings coming from the same agency (Aguinis & Bradley, 2014; Auspurg & Hinz, 2014; Hox, Kreft, & Hermkens, 1991). Traditional guidelines stress the importance of distinguishing within and between respondent variation (Auspurg & Hinz, 2014; Wallander, 2009). As a matter of fact, when implementing the above recommendations on the acquisition of information about the agency and the use of clustered/stratified sampling, vignettes are not only clustered in respondents, but respondents are also clustered into agencies. This leads to a second kind of dependency in the individual vignette ratings: ratings given by two different respondents from the same agency tend to be correlated and are therefore not independent. This is because every agency is attempting to implement certain local policies and because all of the respondents from a given agency are subject to similar unobserved environmental/cultural/political influences. A multi-level regression model allows for the inclusion of known organisational or municipal features, such as population parameters or the number of clients with the problem being studied (e.g. people on social assistance). Further, it is possible to study the unexplained variance at all levels included in the model. For instance, when the results indicate large differences between agencies or municipalities, this means that much of the discretionary freedom is being used at agency level, with local regulations or historically

grown treatment traditions guiding the decisions of social workers. The main advantages of multi-level models have been discussed in several factorial survey papers (Hox e.a., 1991), but they have never been used in their full capacity in former studies (Baughman e.a., 2012; Corazzini, 2003; Degenholtz e.a., 1999; Killick & Taylor, 2012; Mion e.a., 2010; Müller-Engelmann, Krones, Keller, & Donner-Banzhoff, 2008; Wallander & Blomqvist, 2008).

### 3.3. Limitations

The vignette method does have a number of drawbacks, however. Most of them concern validity: questions arise as to whether or not the method measures what it sets out to. Internal validity, meaning the ability to draw causal conclusions, is high since the method operates on an experimental basis (B. J. Taylor, 2006; Wallander, 2012a) and ensures that no systematic errors occur. External validity, in all its forms, is more problematic (Auspurg & Hinz, 2014). Most of the problems and corresponding solutions have already been discussed in the existing literature, and, due to space limitations, we refer the reader to other articles. Examples of external validity problems are: the risk that vignettes might lack realism (Aguinis & Bradley, 2014; Aiman-Smith, Scullen, & Barr, 2002; Karren & Barringer, 2002; Wallander, 2009); the risk that the respondent's answers lack realism (Armacost, Hosseini, Morris, & Rehbein, 1991; Eifler, 2010; Kirwan, Chaput de Saintonge, Joyce, & Currey, 1983; Langley, Tritchler, Llewellyn-Thomas, & Till, 1991; Mohan e.a., 2014; Peabody, Luck, Glassman, Dresselhaus, & Lee, 2000a); and the potential over- and under-complexity of the vignettes (Aiman-Smith e.a., 2002; Auspurg, Hinz, Liebig, & Sauer, 2015; Auspurg & Jäckle, 2015; Caussade, Ortúzar, Rizzi, & Hensher, 2005; DeShazo & Fermo, 2002; Johnson, 2006; B. J. Taylor, 2006; Wallander, 2009). The type of factorial survey used to study social policy implementation has the same weaknesses as traditional factorial survey research. For instance, because, as stated in Section 3.2, the respondent should be asked about what would happen in her organisation and not about her own individual decisions, there is an increased risk of response error. This risk can be countered by replacing the question by two questions, one of which asks the professional what she would personally advise, while the other asks what she thinks that would happen in reality. As this would make the questionnaire longer, it may contribute to respondent fatigue and reduce the quality of the respondents' answers. However, in defence of the factorial survey approach, several scholars believe that responses to experimentally created stories, as in a factorial survey, are closer to reality than those given in interviews directly by respondents (Armacost e.a., 1991; Kirwan e.a., 1983; Peabody e.a., 2000a).

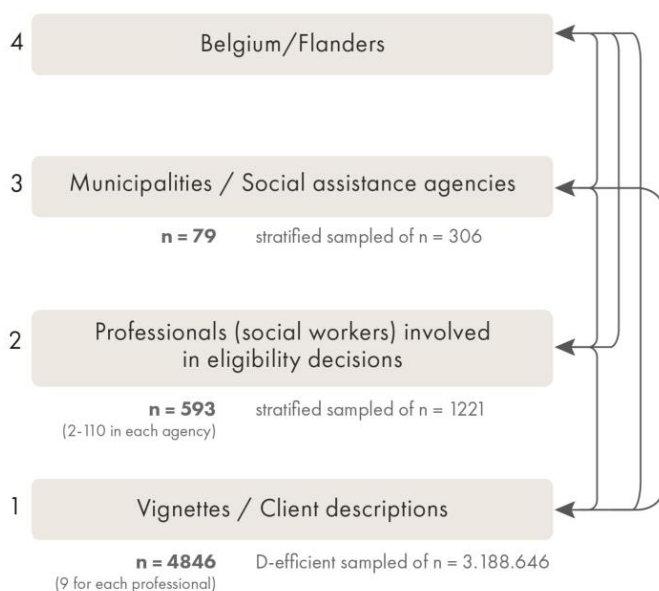
#### 4. Proof-of-concept example

In this section, we demonstrate the potential of the factorial survey method with an example. The policy being investigated in the example is social assistance in Flanders. We asked almost 600 professionals working in 79 social assistance organisations to respond to a unique<sup>3</sup> set of nine descriptions of social assistance claimants (see Figure 1-2). The main question was whether clients with and without children are treated similarly when it comes to activation requirements (= direction towards the labour market). The dependent variable under investigation concerned the likelihood (measured on a 7-point Likert scale) that a client would lose her social assistance benefits if she refused a job or activation offer that started at 5 o'clock in the morning. In each of the following three paragraphs, we address one of the claims made above. First, we demonstrate the effects of the client characteristic 'parenthood', as well as the professional's and organisational characteristics. Second, we demonstrate the usefulness of the multi-level technique. In the third paragraph, we elaborate on the sampling procedure.

---

<sup>3</sup> In total, there were 400 unique questionnaires (decks), spread over almost 600 professionals. Each deck was rated at least once.

Figure 1-2. Study design and sample



4

#### 4.1. Accounting for effects on three levels

The vignettes used in this study portrayed single clients who had no work, no money, no savings and no contact with close relatives. The clients in the vignettes differed with regard to 14 client characteristics, including gender, nationality, level of education, language abilities, mental health, addiction, motivation to work, work and activation experiences and attitude (see Appendix 2 for a detailed outline of the vignette characteristics). The characteristic that interests us most in this paper is parenthood. We included three levels: a client with no children, a client with a healthy two-year-old child and a client with a sick two-year-old child (immunity disorder, meaning that the child falls ill often and unpredictably). Based on the literature (Eardley e.a., 1996; Kazepov, 1999; Lødemel & Trickey, 2001; Lorentzen, Dahl, & Harsløf, 2012; Rice, 2013; van Oorschot, Uunk, & Jeene, 2008) and on interviews with social assistance stakeholders and academics about the factors influencing eligibility and sanction decisions, we expect

---

<sup>4</sup> Note with regard to the population sizes: 3: Flanders has 306 municipalities and thus 306 social assistance agencies; 2: from each of the 90 municipalities that contributed to the survey, I received lists of the professionals involved in social assistance eligibility decisions, a total of 1221 professionals. 3: the vignette population consisted of more than 3 million experimental combinations of the 14 client characteristics.

that clients with children are less likely to lose their benefit. Likewise, we expect this likelihood to drop even further for clients with sick children. In Belgian legislation (Vande Lanotte, 2002b), the condition 'parenthood' as such does not exempt from job or activation acceptance. Claimants younger than 25 years are required to cooperate in the activation process, unless fairness or health reasons apply. Such fairness reasons are not described in legislation. Whether being a parent (of sick children) is considered as a fairness reason might depend on general tendencies, across professionals and municipalities, to do so. If such general tendencies are present, we should see a strong effect of including the client characteristic 'parenthood' in the regression model. The fairness assessment could, however, also be influenced by individual considerations of professionals (i.e., discretion) or by local legislation or practice (i.e., decentralisation). To test all these possibilities, we used multi-level models (see the section below) and we include explanatory variables of the client, the professional and the organisation level.

In Model 1, we included all 14 client characteristics, but, in Table 1-1 we only report results relevant for the parenthood variable. The effect of having a healthy or a sick child on the likelihood of being sanctioned after refusing a job or activation offer that starts early in the morning is clearly seen in Table 1-1 in the 'Model 1' column. Having a child reduces the likelihood of being sanctioned by around 14 (healthy child) to 19 (sick child) percentage-points. Although this is not shown in the table, whether or not having a child is the client characteristic that had the largest estimated effect on the dependent variable.

Up to this point, our example does not contain new methodology. The use of models such as Model 1 is the traditional way in which factorial surveys are used: to detect the effects of vignette attributes on the dependent variable. In Model 2, however, we introduce explanatory variables related to the respondents rating the vignettes. These were 593 professionals from 79 Flemish social assistance offices. We wanted to enter two sorts of respondent characteristics that might influence the respondents' approach to sanctioning clients with children. The first characteristic was parenthood: we asked whether the respondent had children herself (binary variable). To check whether respondents with children reacted differently when treating clients who also had children, we included an interaction term between the two variables. We also included the age of the respondent, because, otherwise, any observed parenthood effects could be due to an age effect, as older respondents are more likely to have children. The second variable added to the model was an opinion question. We asked the respondents to rate the following statement from 1 (disagree) to 5 (agree): "welfare clients should be sanctioned more often if they do not comply with agreements". This



variable was drawn from a broader standardised opinion questionnaire completed by all respondents and should not normally be used on its own. As a matter of fact, the two variables entered in the model are only given for the purpose of illustration. In a thorough analysis, we would include a larger number of variables. It is true that the effects measured in this paper might have been influenced by variables we did not take into account, but, for clarity of exposition, we have decided to keep the analyses simple in our proof-of-concept example.

The results of the modified model are shown in the 'Model 2' column in Table 1-1. There was no direct effect of respondent parenthood, but there did seem to be an age effect. Older respondents predicted less sanctioning of clients than did their younger colleagues. Furthermore, professionals with children predicted less sanctioning of clients with sick children than their childless colleagues did. The professional's attitude towards sanctioning in general also influenced the results. The more the professional believed that additional sanctioning is positive in general, the more likely she was to predict that the hypothetical clients would be sanctioned.

In Model 3, we added variables related to the organisation and the municipality in which it is located. First, we entered a variable indicating the availability of childcare, based on figures from the organisation governing childcare provision in Flanders (Kind en Gezin). The figures reflect the percentage of children aged between 0 and 3 for whom childcare is available in the municipality. Across the 79 municipalities surveyed in this research, the figures ranged from 19% to 69% with a mean of 42%. We combined this variable with the client parenthood variable in an interaction term. We also added each organisation's mean opinion of sanctioning (see respondent variable described above). In this way, we were able to test whether the common culture in an organisation influences the predictions of sanctioning and potentially overrules the rather strong effect of professionals' personal opinions.

No significant, meaningful results were found in Model 3. Childcare availability is significant, but not in the way that we expected: municipalities with high childcare availability seem to sanction less. This unexpected result may be due to one or more other characteristics not included in our analysis here. Also the interaction with client parenthood – the variable that interested us most – was not found to be significant. The same was true for the mean opinion on sanctioning in the municipality (Table 1-1).

This example, which demonstrates the inclusion of variables from three of the levels that are deemed decisive for social assistance implementation, shows that factorial surveys are a powerful method. In principle, it should also be possible to investigate several countries using one set of vignettes. Country-dummies could be added to the

multi-level regression model, additively and in interaction terms with variables from the three other levels, to detect country-specific effects. If enough countries are included in the study, country characteristics can be added too. This would result in a degree of comparability that is unique in social sciences.

Table 1-1. Multi-level regression results with regard to sanction likelihood

<b>Likelihood to loose social assistance if refusing a work offer in percentages</b>	<b>M0</b>	<b>M1</b>	<b>M2</b>	<b>M3</b>
Intercept	46.6***	53.63***	47.61***	50.17***
C parenthood ( <i>ref. = C no child</i> )		C Healthy child -13.74***	-14.46***	-19.42***
		C Sick child -18.68***	-19.82***	-22.95***
R age			-0.27**	-0.27**
R parenthood ( <i>ref. = R no child</i> )		R child(ren)		
			-2.36	-2.45
<i>Interaction</i> C parenthood & R parenthood		C healthy child * R child(ren)	-1.83	-1.94
		C sick child * R child(ren)	-3.14*	-3.13*
M Availability of childcare				-0.29*
<i>Interaction</i> C parenthood & M availability of childcare		C healthy child * M availability of childcare		0.11
		C sick child * M availability of childcare		0.07
R Opinion concerning sanctioning of welfare clients			4.91***	4.62***
M Opinion concerning sanctioning of welfare clients				3.38
N municipalities	79	79	79	79
N professionals	593	593	593	593
N	4846	4846	4846	4846

Model 0 = multi-level model with no predictor variables;

Model 1 = M0 + client characteristics as fixed effects

Model 2 = M1 + respondent characteristics as fixed effects

Model 3 = M2 + organisation and municipality characteristics as fixed effects

C = client/vignette variables; R = professional/respondent variables; M = municipality / agency variable

\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

#### 4.2. The advantages of multi-level models

As pointed out, all of the analyses presented in Section 4.1 involved multi-level models. A major advantage of multi-level modelling is that it offers the possibility to investigate which level (organisation – professional – client) influences treatment most. In multi-level models, the unexplained variance of the regression model is divided into as many parts as there are levels in the model. The resulting variance components are shown in Table 1-2. The null model (Model 0) provides insight into the division of the unexplained variability when no explanatory variables are considered. As we can see, most of the unexplained variance in the null model appears at the first and at the second level, namely at the client and the respondent level. Only 6% of all unexplained variance appears at the municipality level (Row 2). Most of the unexplained variance in the null

model is due to predicted treatment differences for client vignettes rated by a given professional or respondent (49% - Row 11). A similar proportion of the unexplained variation occurs between professionals (45% - Row 5). Before evaluating these percentages, it is important to know whether the variation in the responses is substantial. As each possible value from 1 to 7 was chosen almost equally often by the respondents in our example, that variation is certainly large. The fact that the variation in the responses is so high, and situated at the respondent and client levels, means that predicted treatment mainly depends on client characteristics (as should be expected) and on professional preferences. If equal treatment for similar clients is targeted, this is an important result which cannot be established without using multi-level techniques.

A second advantage of using multi-level techniques is that they enable us to check the impact of adding explanatory variables (fixed effects or random intercept regressions) to the model on the magnitude of the unexplained variance at each level. As we can see in Table 1-2– Model 1, the unexplained variance at the first level (the vignette) decreased by 76.2 (Row 10), which is a reduction of 18%-point (Row 12). In total, 13%-point of the original total unexplained variance is explained by adding the client characteristics (Row 15). This reduction is completely situated at the vignette level. Of the 13%-point decline in the total unexplained variance due to entering the client variables in the model, 12%-point is due to the parenthood variable (not shown in Table 1-2). This means that one client variable explains almost all variation.

When adding the variables at the respondent level (Table 1-2, Model 2), the unexplained variance at the second level decreased by 10%-point (Row 6) and the total unexplained variance further decreased by 5%-point (when compared to Model 1 - Row 15). This decrease is entirely located at the second level, which means that the respondent characteristics selected do not influence the results at the other levels. Lastly, the addition of the organisation and municipality variables causes a reduction of 30%-point of the unexplained variance at the first level (Table 1-2, Model 3, Row 3). However, this only represents 1% of the total variance (Row 15).

Table 1-2. Multi-level variance components with regard to sanction likelihood

Variance components		Model 0	Model 1	Model 2	Model 3	Model 4
Level 3 –	Real value	51.8*	47.8*	46.4*	32.6*	36.6*
Municipality	% of total variance	5.7%	/	/	/	/
	%-point decrease compared to the previous Model	/	-7.7%	-2.9%	-29.7%	/
Level 2 –	Real value	446.6***	407.7***	369.0***	372.1***	/
Respondent	% of total variance	49.0%	/	/	/	/
	%-point decrease compared to the previous Model	/	-8.7%	-9.5%	0.8%	/
	Real value C no children <sup>5</sup>	/	/	/	/	403.7***
	Real value C healthy child	/	/	/	/	453.4***
	Real value C sick child	/	/	/	/	472.5***
Level 1 –	Real value	413.4***	337.2***	336.8***	336.5***	263.3***
Client vignette	% of total variance	45.3%	/	/	/	/
	%-point decrease compared to the previous Model	/	-18.4%	-0.1%	-0.1%	/
Total variance	Real value	911.8	792.7	2.66	741.2	/
	%-point decrease compared to Model 0	/	-13.1%	-5.1%	-1.4%	/
N		4846	4846	4846	4846	4846

Model 0 = multi-level model with no predictor variables; Model 1 = M0 + client characteristics as fixed effects ; Model 2 = M1 + respondent characteristics as fixed effects ; Model 3 = M2 + organisation and municipality characteristics as fixed effects ; Model 4 = M3 + random slopes for client characteristic at the second level  
 C = client/vignette variables; R = professional/respondent variables; M = municipality / agency variable  
 \* p<0.05, \*\* p<0.01, \*\*\* p<0.001

A final advantage of multi-level techniques is shown in the last column of Table 1-2 (Model 4): the ability to add random slopes. Adding random slopes for the client parenthood variable, for instance, at the second level (the respondent) means we do not expect the effect of client parenthood to be the same for all respondents. After all, given that the preferences of the professionals do matter (see variances in the null model), the likelihood that clients with children will be sanctioned less than clients without children probably varies from respondent to respondent. When calculating the unexplained variance at the second level for respondents rating clients without children (404 – Row 7), we do indeed see less unexplained variance than when respondents are rating clients who do have children (453 – Row 8), especially compared to clients with a sick child (473 – Row 9). This means that clients with children are sanctioned less than clients who do not have children (see above for client effect), but that professionals differ in their degree of sanctioning.

<sup>5</sup> The variance for the specific client parenthood levels is calculated as follows: var at the respondent level for a specific C parenthood level= (var of the intercept at the respondent level) + 2\*covariance(intercept\C healthy child)\*C health child + 2\*covariance(intercept\C sick child)\*C sick child + 2\*covariance(C healthy child\C sick child)\*C healthy child\*C sick child+var(C healthy child)<sup>2</sup>\*C healthy child + var(C sick child)<sup>2</sup>\*C sick child

#### 4.3. The advantages of stratified sampling

We selected the 79 municipalities and almost 600 professionals surveyed in Flanders in 2015 using multi-stage stratified sampling. The first stratification level was composed by grouping the municipalities (each municipality in Belgium has one social assistance organisation). The strata were based on the number of inhabitants (three levels: <50,000; 50,000–100,000; and >100,000), the percentage of social assistance claimants among all inhabitants (four levels: <0.5%; 0.5–1%; 1–2%; and >2%) and an 16-cluster-index based on socio-economic and socio-demographic characteristics (Belfius, 2007). As these indices overlapped to a certain extent, several of the 192 strata remained empty. We established 36 strata in total. Next, guided by power analyses based on a pilot study, we randomly selected a third of the municipalities in each stratum. This resulted in 90 municipalities. If a social assistance organisation's contact person declined the invitation to participate, we replaced the corresponding municipality with another one from the same stratum. In total, we invited 104 municipalities, 90 of which actually participated in our study. In 13 of these municipalities, only one person eventually responded. In two municipalities there was only one professional working in the particular position. These municipalities were kept in the analyses. We excluded the other 11 municipalities from our analysis, because having only one respondent from a municipality with multiple professionals endangers the representativeness.

In each municipality, we sampled one third of the professionals involved in the decision-making process concerning eligibility for benefits. This means that social workers specialised in debt counselling or activation trajectories were not surveyed. We asked the social assistance organisations to provide lists of these professionals and to include information about gender, age group (4 levels) and position (team manager or not). Based on these characteristics, we constructed 16 strata, from which we drew a proportional one-third sample, while making sure that at least one person was selected from each stratum available in each municipality. In the event a certain respondent did not participate in the study after two reminders, we approached a similar professional from the same municipality. In total, 839 professionals received the questionnaire, 681 started the survey and 593 completed all data used in the analyses (see Figure 1-2).

In the remainder of this section, we demonstrate how generalisation of results would not have been possible if we had not used random or stratified samples. Paying insufficient attention to the respondent and municipality levels tends to give rise to two rather common practices, both of which cause bias in the results. The first of these is only studying some of the organisations (generally the largest in size), while the second is surveying only those professionals who volunteer to participate. Here, we wish to

demonstrate that, even if a researcher only wants to take the effect of vignette characteristics into account, he or she should take a random or stratified sample of all possible organisations and a random or stratified sample of respondents from those organisations.

In order to demonstrate the potential bias, we added two new models to our study (see Table 1-3). Model 5 and 6 respectively demonstrate the need for stratified sampling of municipalities and professionals. Model 5 adds several variables to Model 1 (only client characteristics): seven dummy variables indicating the largest municipalities in the study (more than 80,000 inhabitants = Aalst, Antwerp, Bruges, Ghent, Hasselt, Leuven and Ostend) and the interaction of these dummies with the client parenthood variable. As shown in Table 1-3, one of the municipalities entered using a dummy has a direct effect on the likelihood of sanctioning. Professionals in Municipality 6 tend to predict sanctioning less often than those in smaller municipalities and in at least three of the other large municipalities. Some of the municipalities also display a specific effect with regard to client parenthood. For Municipalities 3 and 7, the negative effect of having children on the likelihood of being sanctioned is less strong than for the smaller municipalities. In Municipality 4 and 6, the positive interaction effects (almost) overrule the negative main effect of the municipality. This means that these municipalities sanction less in general (not significant for Municipality 4), but (almost) as often if the client has children. These results show that, should a researcher only survey these large cities, the results would not be generalizable.

Table 1-3. Multi-level results for selection of municipalities and professionals with regard to sanction likelihood

Likelihood to loose social assistance		Model1	Model 5	Model 6
Intercept		4.2***	55.95***	55.34***
C Parenthood (ref. = no child)	C healthy ch	-0.82***	-16.45**	-16.07***
	C sick ch	-1.12***	-21.62***	-21.02***
Municipalities – dummies	M1		3.58	
	M2		-5.97	
	M3		0.02	
	M4		-9.21	
	M5		-0.9	
	M6		-16.24***	
	M7		3.29	
Interactions M & C parenthood	M1 * C healthy ch		2.4	
	M1 * C sick ch		3.24	
	M2 * C healthy ch		2.98	
	M2 * C sick ch		-0.12	
	M3 * C healthy ch		7.45*	
	M3 * C sick ch		11.73***	
	M4 * C healthy ch		8.13*	
	M4 * C sick ch		7.67*	
	M5 * C healthy ch		2.45	
	M5 * C sick ch		-1.16	
	M6 * C healthy ch		7.54***	
	M6 * C sick ch		9.11***	
	M7 * C healthy ch		12.09	
	M7 * C sick ch		13.75*	
Type of respondent (ref = R female, 30-40, not TM)	R = male – 40-50 – not TM			-13.92*
Interaction R type & C parenthood	R = male – -30 – not TM * C healthy ch			11.61**
	R = male – -30 – not TM * C sick ch			10.42**
	R = male – 30-40 – not TM * C healthy ch			6.79*
	R = male – 30-40 – not TM * C sick ch			8.93**
	R = male – 40-50 – TM * C healthy ch			12.5***
	R = male – 40-50 – not TM * C sick ch			18.54**
	R = male – 40-50 – TM * C healthy ch			16.15**
	R = male – +50 – TM * C sick ch			8.73*

Model 1 = multi-level model with client characteristics as fixed effects; Model 5 = M1 + dummy variables for 7 municipalities + interactions of these dummies with client parenthood ; Model 6 = M1 + respondent-strata-variable + interactions of this variable with client parenthood

C = client/vignette variables; R = professional/respondent variables; M=municipality/agency variables; TM=team manager; ch=child

\* p<0.05, \*\* p<0.01, \*\*\* p<0.001

In Model 6, we added a different sort of variable to demonstrate how a non-random or non-stratified sample of professionals might distort the results. As described above, we stratified the respondent population (= all professionals working on social assistance eligibility files in the 79 selected municipalities) by three characteristics, namely age (four groups), gender and position (two groups each), resulting in 16 strata. One of these strata was absent in both the population and the sample, namely male team managers

younger than 30. Most professionals in the respondent population were female, aged between 30 and 40 years, and did not hold a management position. In Model 6, we treated this group as the reference group and entered a categorical variable indicating the 14 other strata. Furthermore, we entered an interaction term combining these strata with the client parenthood variable. If researchers were to rely either on voluntary cooperation or on non-stratified sampling, it is very likely that some of these strata would be absent in the actual sample. We show in Column 5 of Table 1-3 that this might change the results. In the table, in the interest of space, we present only the significant interaction terms. The male respondents that were one decade older than the reference group differ significantly from the female reference group: they predicted significantly less sanctioning. The bottom part of Table 1-3 shows several additional significant interaction terms indicating that the parenthood variable has an impact that differs across the professional strata. If a researcher were to miss these specific groups of professionals, due to incomplete sampling methods, the results would not be generalizable. Because these additional interactions are positive (= decreasing the negative main effect), the effect of having sick children would be overestimated.

## **5. Conclusion**

In this paper, we argue that a customised use of the factorial survey approach is a promising method for studying social policy implementation. As this implementation is affected by the country, region, organisation, social worker and client, as well as by the cultural, political, economic, demographic and social environments, it needs to be studied as a multidimensional topic. Existing research suffers from a lack of generalisability, mostly due to a qualitative or single-region approach. Large-scale research, mostly conducted using administrative data, suffers from a selection bias. In reality, clients are not distributed randomly over municipalities, which makes it difficult to detect what might affect treatment: the client's characteristics or the agency where the client is assisted. The factorial survey helps to overcome these problems by disseminating vignettes across a large number of respondents in multiple agencies, municipalities and various countries, using modern, inexpensive communication channels. Furthermore, the D-efficient sampling of the vignette universe and assignment to the respondents ensures that the analysis of client characteristics is impacted as little as possible by differences between respondents.

The proposed method for studying social policy implementation exhibits similarities with the recent approach to investigating professional judgment. The latter focuses on detecting the typical treatment of clients with certain characteristics. We go beyond this focus, as we are also interested in the typical treatment of clients in specific types of



municipalities or by respondents with certain characteristics. This means that, for our kind of study, a questionnaire about the respondent and the organisation she works for is necessary, in addition to a set of vignettes to rate. As shown in the example concerning the likelihood of sanctioning social assistance clients who refuse a job offer starting early in the morning, not only the client's parenthood itself mattered to explain the reduced likelihood of being sanctioned in the event the client had children. When entering respondent characteristics such as the professional's parenthood or her opinion on sanctioning in general, these characteristics proved to be significant and provided valuable information about the rationale behind the implementation of sanction measures.

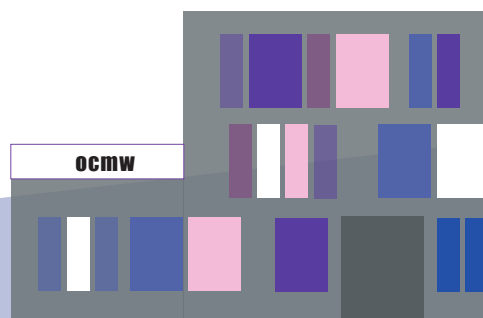
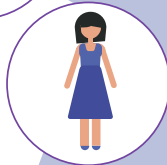
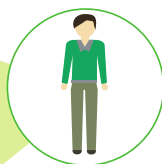
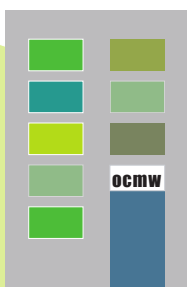
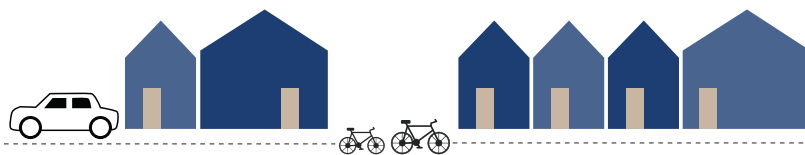
Furthermore, the sampling of respondents is of major importance. To study social policy implementation, which accounts for local variation, respondents should be distributed evenly over municipalities with specific characteristics and selected at random or stratified from the group of employees in each municipality. In our example, we show that if researchers focus on large municipalities or on the voluntary contribution of the professionals, it is very likely that the results will be biased. Some large Flemish municipalities displayed very specific results, concerning both sanctioning in general and sanctioning clients who do or do not have children. Furthermore, for the group of professionals that is overrepresented in the population (female professionals aged between 30 and 40 years), the effect of having children on the likelihood of being sanctioned is larger than for other groups of professionals. With non-random or non-stratified sampling, some of these other small groups would not appear in the sample, which would bias the results.

A final recommendation formulated in this paper was to use multi-level regression methods in its full capacity. Researchers studying social policy implementation should take into account the clustered nature of the data collected by surveying professionals (vignettes nested within respondents) from several agencies (respondents clustered into agencies). Multi-level methods allow for explanatory variables from different levels to be included simultaneously in the regression model: client characteristics (= vignette dimensions), respondent characteristics and agency characteristics. Some additional results can be obtained from these analyses. We showed that the large variation in the responses could be explained mainly by differences between professionals from the same municipality and between the different clients rated by a given professional. Entering the dummy variables from some municipalities showed that certain large municipalities displayed particular results (see the previous paragraph), but, overall, differences between municipalities (79) were by far not the most important explanation

for the variation in the responses. This is a valuable result as it contradicts the literature on decentralisation, which expects to see large differences among local organisations, and confirms – quantitatively – the street level bureaucracy research emphasizing the extensive policy-making power of local professionals (McDonald & Marston, 2005). Second, we were able to calculate the explanatory effect of specific fixed effects on the unexplained variance. We observed that the client variable parenthood reduced the unexplained variance by 7%-point, while the respondent and agency variables entered led to a decrease of 5%-point in unexplained variance. This means that the local context variables have an almost equal effect on the final results as the client characteristics. Third, by entering random slopes in the model, we showed that respondents not only differ significantly in their general predictions of sanctioning, but also in their predictions of the effect of parenthood on the likelihood of being sanctioned. The variation among respondents in predicting sanctions was higher when evaluating clients with children.

# In a category of their own?

A MULTIGROUP SEM COMPARISON OF THE WELFARE STATE ATTITUDES OF SOCIAL WORKERS AND THE GENERAL PUBLIC



Chapter 2. In a category of their own? A multigroup SEM comparison of the welfare state attitudes of social workers and the general public

---

*Marjolijn De Wilde, Bart Meuleman and Koen Abts*

*Submitted to the British Journal of Social Work*

**Abstract**

Social workers' welfare state attitudes are of prime interest, as it is expected that these attitudes may potentially influence client treatments. In this paper, we compare social workers' welfare state attitudes with those of the general public in Flanders in order to test two competing hypotheses. On the one hand, the professional identification hypothesis assumes that social workers have a more positive outlook towards the welfare state as a result of socialization, self-interest and professional identification. On the other hand, the selection hypothesis conversely assumes that only citizens with specific socio-demographic characteristics and ideological preferences choose to study social work and therefore stresses the similarities of social workers to citizens with similar characteristics and preferences. Using a multigroup structural equation modelling approach, we compare three dimensions of welfare state attitudes: (1) the perceived economic and moral consequences of the welfare state (*welfare state criticism*), (2) the call for control of benefit users (*welfare state sanctioning*) and (3) the perceived overuse of welfare state benefits (*welfare state overuse*). Our results show that the welfare state attitudes of Flemish social workers differ considerably from their fellow citizens. Although the attitudinal discrepancy decreases when controlling for socio-demographic characteristics and political preferences, the latent mean differences between social workers and the general public remain significant. Social workers are indeed more positive about the moral consequences of the welfare state, less in favour of more control and punishment of benefit users, and suspect less benefit abuse compared to the general population. Both the professional identification and selection hypotheses are partly confirmed.

## 1. Introduction

Social work professionals and street-level bureaucrats are actors of crucial importance in translating social policy into social practice. In several policy domains, they are granted a substantial degree of discretion to choose between various treatment options based on official legislation, local regulations, contextual possibilities and the capacity and needs of clients, but also based on their own preferences (De Wilde, 2016; De Wilde & Marchal, 2018; Evans, 2012; Lipsky, 1980; Soss e.a., 2011; Tummers, Bekkers, Vink, & Musheno, 2015). Several studies have shown that a policy is only 'real' if professionals implementing the policy accept the dominant ideology behind the policy (Moller & Stone, 2013). Role conflicts between professionals' own values and norms and the policy they are required to implement have a major influence on their willingness to do so (Tabin & Perriard, 2016; Tummers, Steijn, & Bekkers, 2012). As a result, it seems reasonable to expect that the general attitudes of social workers towards the welfare state will be reflected in their actual treatment decisions (Blomberg, Kroll, Kallio, & Erola, 2013; Castillo & Becerra, 2012; Cozzarelli, Wilkinson, & Tagler, 2001; De Wilde & Marchal, 2018; Keiser, 2010; Reingold & Liu, 2009).

While there is extensive literature on the general public's attitude towards welfare state provision (Blekesaune & Quadagno, 2003; Kulin & Svallfors, 2013; Larsen, 2008), only a few studies focus particularly on the welfare state attitudes of social workers, and these studies have several limitations. First, most studies have focused on poverty attributions among social workers (Cozzarelli e.a., 2001), which is only one very specific aspect of the relevant concepts identified in welfare attitude research. Second, most research does not compare social workers' attitudes to patterns among the general public, and therefore lacks a benchmark (Weiss & Gal, 2007; Weiss-Gal & Gal, 2008). For these reasons, the question concerning the extent to which the attitudes of those who implement social policies resemble general public opinion on the welfare state remains unanswered.

This paper tests the professional identification and selection theses by comparing welfare-related preferences of social workers and the general public. Specifically, we set out to determine the extent to which and why social workers deviate from general public opinion with respect to three dimensions of welfare state attitudes (Roosma, van Oorschot, & Gelissen, 2014; van Oorschot & Meuleman, 2012): 1) the perceived consequences of the welfare state on the behaviour of citizens, 2) the call for control of benefit users and 3) the perceived overuse of welfare state benefits. To address this issue, we combined and analysed two complementary datasets that use identical measures, the Belgian National Election Study 2014 (Abts e.a., 2015) and a survey of 603

social workers employed in social assistance agencies in Flanders. Moreover, to guarantee comparability, we applied state-of-the-art multigroup structural equation modelling combined with an alignment procedure to test measurement equivalence.

## 2. Literature review

### 2.1. Welfare state attitudes as a multidimensional concept

Several scholars have suggested that welfare state attitudes cannot be reduced to a single pro vs. contra welfare state stance, but are multidimensional instead (Roosma, Gelissen, & Oorschot, 2013; van Oorschot, Reeskens, & Meuleman, 2012). People can be supportive of certain aspects of welfare provision, but be quite critical of other dimensions. In this respect, Roosma, van Oorschot and Gelissen (2013) make the crucial distinction between support for the underlying principles of welfare state provision and evaluations of the concrete implementation and performance of welfare arrangements. In general, European citizens are rather positive about the goals of the welfare state (Blekesaune & Quadagno, 2003), but they are much more critical about the policy outcomes, the unintended side effects and the efficiency of certain policies (Roosma e.a., 2014; van Oorschot & Meuleman, 2014).

In this contribution, we focus on three particular dimensions of welfare state attitudes related to performance evaluations. First, *welfare state criticism* refers to perceptions that welfare state redistribution has unintended but detrimental consequences for the economy and the moral behaviour of the population. According to these critical views, the welfare state places an excessive burden on national budgets, and undercuts incentives for citizens to be self-reliant (Ervasti, 2012; van Oorschot e.a., 2012). Second, *welfare state sanctioning* refers to the conviction that welfare beneficiaries should be subjected to greater levels of administrative control, and that they should be punished when violations are discovered. A third and related dimension *welfare state overuse* – that is, beliefs that certain categories of recipients get welfare benefits that they are not entitled to – can be considered as the ‘Achilles heel of welfare legitimacy’ (Roosma e.a., 2014).

These three dimensions are particularly relevant to compare social workers and the general population, not only because they deal with contentious issues, but also because these attitudes might guide social workers in their treatment choices. The growing importance of activation and reintegration in the labour market has gone hand in hand with the introduction of behavioural conditions for eligibility (Clasen & Clegg, 2007; Standing, 2005; Van Kersbergen & Hemerijck, 2012). In this regard, successful activation requires evaluations of work willingness embedded in a deservingness

assessment (De Wilde, 2017). Although deservingness is difficult both to prescribe in legislation (Meyers e.a., 1998) and to assess in practice, the importance of the discretion used by local social assistance agencies and social workers is likely to increase as a result of the introduction of behavioural requirements rather than solely categorical conditions (e.g. lone parent) or means tests (De Wilde, 2017; De Wilde & Marchal, 2018; Moller & Stone, 2013; Van Berkel & Aa, 2012). If social workers are granted a degree of discretion to assess clients' deservingness, it seems logical that their general ideas about deservingness will influence their actual treatment decisions (Blomberg e.a., 2013; Castillo & Becerra, 2012; Cozzarelli e.a., 2001; Reingold & Liu, 2009; Weiss-Gal, 2008). However, to investigate this issue, we need a study of different types of welfare state attitudes among social workers with a benchmark – i.e. a comparison of social workers' attitudes with these of mass public.

## 2.2. Comparing social workers and the general public: professional identification or selection?

Typical social work research (Weiss & Gal, 2007) describes social workers as citizens with specific characteristics, norms and attitudes, different from those of general public. Mostly focusing on poverty attributions, studies conclude that social workers favour a structural explanation of poverty (Blomberg e.a., 2013; Bullock, 2004). Although similar research on welfare state attitudes is lacking, we expect that social workers – who think more likely in terms of structural causes of poverty – will be less critical about the negative consequences of the welfare state (*welfare state criticism*), less often suspect overuse of the welfare state system (*welfare state overuse*) and see less salvation in controlling individuals and their behaviour (*welfare state sanctioning*).

To explain these differences between social workers and general public, two competing hypotheses can be discerned: the professional identification hypothesis and the selection hypothesis. The professional identification thesis assumes that social workers have particular welfare state attitudes as a result of education, experience, self-interest and other forms of professional identification (Blomberg et al., 2013; Bullock, 2004; Guy, 2011; Weiss & Gal, 2007; Weiss, 2005). The selection approach argues that the differences between social workers and general public are essentially a result of composition effects, and rather stresses the similarities between social workers and citizens with the same socio-demographic and ideological preferences (Weiss & Gal, 2007).

The professional identification hypothesis argues that professional identity goes together with particular predispositions and beliefs that differentiate social workers from general population (Weiss & Gal, 2007). The formation of a professional identity

should be seen as an ongoing process of actively identifying with a particular context, knowledge and views (e.g., Harrison & Healy, 2016; Limb & Organista, 2006). This identification might stem from a shared educational background (Barretti, 2004; Terum & Heggen, 2016; Weiss, Gal, & Cnaan, 2004) or socialization within a particular workplace with specific experiences, ethical guidelines and a common professional culture (Blomberg e.a., 2013; Harrison & Healy, 2016), as well as being related to a typical form of self-interest. In this regard, academics traditionally assume that working with people in poverty fosters a belief in the structural causes of poverty. Other research, however, suggests that social workers who work with social assistance clients favour individual explanations more than do those working in child welfare, in which they are confronted with people of all layers of the population (Blomberg e.a., 2013; Weiss & Gal, 2007). In this case, experience seems to have the opposite effect to what would be expected. Further, self-interest – as understood in this context – explains differences between social workers and their peers, because people feel inclined to support the system they work in. Self-interest is, then, not support for institution with a direct financial or other advantage, but as support for self-formed policies or self-produced institutions (Jæger, 2006). As a result, the professional identification hypothesis argues that the expected lower levels of *welfare state criticism*, *welfare state sanctioning* and perceptions of *welfare state overuse* for social workers compared to the general population can be explained by profession specific factors.

However, the selection hypothesis argues that all attitudinal differences are due to the fact that only specific citizens choose to study social work and only specific social workers choose to work with people in poverty. As social workers are more often female, young, highly qualified, autochthon and left leaning than other citizens, they are – in line with the attitudes of similar citizens in the general population – expected to have a more positive view on the performance of the welfare state. With regard to the general population is often shown that: females, higher educated citizens and more left-leaning people are often seen as more positive about the welfare state (Bean & Papadakis, 1998; Blomberg e.a., 2013; Castillo & Becerra, 2012; Costa & Dias, 2015; Cozzarelli e.a., 2001; Edlund, 1999; Elkins & Simeon, 1979; Feldman & Steenbergen, 2001; Hasenfeld & Rafferty, 1989, 1989; Weiss & Gal, 2007). With regard to age differences, findings are mixed (Blomberg e.a., 2013; Costa & Dias, 2015; Cozzarelli e.a., 2001; Hasenfeld & Rafferty, 1989). Other socio-demographic features which might be influential are ethnicity, personal experience with poverty and economic situation (Costa & Dias, 2015; Hasenfeld & Rafferty, 1989; Lepianka, Gelissen, & van Oorschot, 2010; Limb & Organista, 2006). The differences with regard to the general population would, thus, essentially be composition effects, and will disappear by controlling for personal



characteristics. In this sense, the particularity of social work attitudes is doubted since the world view of social workers is quite similar to citizens with similar socio-demographic and ideological features.

The crucial question of this paper is how the professional identification hypothesis and the selection hypothesis interact. As we expect that both hypotheses explain part of reality, our final hypotheses are:

- (1) Social workers' attitudes with regard to the performance of the welfare state differ significantly from these of the general public. They show lower levels of *welfare state criticism*, *welfare state overuse* and *welfare state sanctioning*.
- (2) Parts of these differences can be explained by selection, i.e. similar socio-demographics and political preferences.
- (3) Other parts will not be explained by personal characteristics, and might be seen as a result of professional identification.

### **3. Data and methods**

#### **3.1. Data**

To test the hypotheses, we combine two complementary datasets purposefully designed to allow a detailed comparison of welfare state attitudes among social workers and the general public. First, the Belgian National Election Study (BNES) 2014 (Abts e.a., 2015) contains a register-based probability sample of Belgians entitled to vote in the 2014 elections. Computer Assisted Personal Interviews (CAPI) were conducted by trained interviewers, covering a wide range of topics, including welfare state attitudes. For our study, we only selected the sample of Flemish respondents (N=871; response rate: 55%). Additionally, in 2015, a second survey was carried out among Flemish social workers who were in contact with clients at social assistance agencies. By means of a web survey, a probability-based sample of 603 social workers (response rate: 72%) from 90 Flemish social assistance agencies was approached. Both surveys include the same measurements, which makes a unique comparison of both populations possible.

#### **3.2. Variables**

Our analyses focus on three attitudinal dimensions: perceived negative consequences of the welfare state (*welfare state criticism*); opinions on the extent to which uncooperative social assistance recipients should be monitored or punished (*welfare state sanctioning*); and perceptions of benefit overuse (*welfare state overuse*). To model these dimensions, we used structural equation modelling. This statistical method models latent constructs based on correlations between observable variables (items).

These observed variables are expected to be related to each other and to be explained by the latent construct.

*Welfare state criticism* combines attitudes – expressed in four items – about the unintended but detrimental consequences of welfare state redistribution for the economy (*Costs*) and the moral behaviour of the population (*Dependent, Lazy, No self* – see Table 2-1 for exact wording of the variables). *Welfare state sanctioning* focuses on the perceived need for more administrative control of citizens and beneficiaries (*Job searching, Illicit work*) and for more punishment as a response to violations (*Punishment*). The fourth item under this latent variable is an overall evaluation of the welfare state as overly strict (*Too strict*). The final latent variable, *Welfare state overuse* combines perceptions concerning recipients receiving benefits without being eligible: sickness allowances (*Sick*), unemployment benefits (*Unemployment*) and social assistance benefits (*Welfare*). An overview of the items (5-point Likert scale), including the wording and descriptive statistics for both groups, can be found in Table 2-1.

Table 2-1. Overview of latent variables and descriptive statistics of multiple indicators

Welfare state criticism (5-point Likert scale)		Mean on SW mean (se)	Mean on GP mean (se)	SW comp to GP
<i>Dependent</i>	The welfare state is too much of a safety net that people become dependent on	2.7 (0.93)	3.1 (0.86)	***
<i>Lazy</i>	The welfare state makes people irresponsible and lazy	2.2 (0.8)	2.8 (0.92)	***
<i>No self</i>	The welfare state causes people to no longer be able to take care of themselves	2.3 (0.78)	2.7 (0.8)	***
<i>Costs</i>	The welfare state costs too much money compared to what it yields	2.4 (0.93)	2.9 (0.93)	***
Welfare state sanctioning				
<i>Job searching</i>	The government should check more closely whether the unemployed are applying for jobs sufficiently	3.5 (0.97)	4.0 (0.68)	***
<i>Punishment</i>	Social benefit beneficiaries who do not do what is required of them should be punished more harshly	3.2 (0.95)	3.9 (0.72)	***
<i>Illicit work</i>	The government should check more closely whether the unemployed do additional illicit work	3.9 (0.88)	4.0 (0.76)	
<i>Too strict</i>	The government is too strict on social benefit recipients	2.5 (0.8)	2.4 (0.77)	***
Welfare state overuse				
<i>Sick</i>	People use their health insurance although they are not sick	3.0 (0.73)	3.5 (0.81)	***
<i>Unemployment</i>	People receive unemployment benefits although they could get a job if they wanted	3.4 (0.71)	3.8 (0.82)	***
<i>Welfare</i>	People receive a social assistance benefit (minimum income) although they are not actually poor	2.4 (0.65)	3.2 (0.84)	***
<i>N</i>		603	871	

\* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

Note: SW = social workers; GP = general population

To test the selection hypothesis, we included respondent variables. First, we included several socio-demographic characteristics: gender (female or male), age (in years), level

of education and national background of the parents of the respondent (Belgium, Europe, outside of Europe). Second, we included variables indicating the respondents' ideology. One variable indicated a self-positioned political orientation on a 10-point scale ranging from left (0) to right (10). The other variable measured attitudes to state responsibility, with a scale ranging from 'The state should take more responsibility to ensure that everyone is provided for' (0) to 'Individuals should take more responsibility for providing for themselves' (10). See Table 2-2. for the descriptive results for the two research populations. With regard to education, we could only differentiate between Bachelor's and Master's levels degrees, as all social workers in Belgium are required to have a Bachelor of Social Work.

### 3.3. Modelling strategy

To answer our research questions, we estimated and interpreted a series of different structural equation models. In the first step, we used the multigroup alignment model (Asparouhov & Muthén 2014) – without control variables – to test measurement equivalence and to evaluate the gross mean differences between the two groups (Model 1). In the second step, we regressed the latent variables on several relevant individual characteristics – namely gender, age, national background, political orientation and perceived responsibility for people's wellbeing – to control for compositional differences between both groups (Model 2). For this purpose, we used the 'alignment within CFA' model that was developed recently by Marsh and colleagues (2017). This model uses the results of the alignment model (see step one) as starting values for a model that includes the explanatory exogenous variables. We centred all variables at the mean value for the social workers. In the third and final step, we excluded from the general population all respondents who did not have a Bachelor's or Master's degree. The procedure was similar to that described above: we fitted the alignment analysis first (Model 3), followed by the model with the socio-demographic and ideology features (Model 4).

## 4. Results

### 4.1. Descriptive results

In terms of socio-demographic characteristics, both study populations differ in several respects. The statistical significance of the differences was investigated by the t-test for the means and by the chi-square test for the categorical variables, both at the 5% level: social workers are younger, are more often female and are less likely to have a parent born outside Belgium. As mentioned in the method section, their educational level is never below Bachelor's level (see Table 2-2.).

In attitudinal terms, our study reveals major differences between social workers and the general population. Social workers tend to place themselves as more left-leaning (see Table 2-2.) and differ significantly from the general population with regard to welfare state attitudes. The differences in attitude indicators are as expected: social workers present themselves as having a more positive attitude to the welfare state and its users on almost all items. The two research groups only do not significantly differ with regard to the extent they see wellbeing as the state’s or an individual responsibility (see Table 2-2.) and with regard to the attitude-item ‘the perceived need for more control of illicit work by unemployment benefit receivers’ (see Table 2-1).

Table 2-2. Descriptive results for the independent variables

		SW	GP	SW comp to GP	GP	SW comp to GP
		Full sample	Full sample	Full sample	Bach & mast	Bach & mast
		Mean/%	Mean/%		Mean/%	
Age		38.23	52.97	***	49.21	***
Gender (ref = female)	Male	19.90%	49.7%	***	49.04%	***
Country of birth father (ref = Belgium)	Europe	1.99%	4.94%	**	4.13%	*
	Outside of Europe	1.16%	2.64%	*	1.38%	
Country of birth mother (ref = Belgium)	Europe	1.99%	4.48%	*	3.58%	
	Outside of Europe	0.83%	2.53%	*	1.38%	
Educational level (ref = Bachelor)	Lower than secon education	0%	24.80%			
	Secon education	0%	33.52%			
	Master	10.95%	13.89%	***	33.33%	***
Left-right		3.61	5.39	***	5.29	***
Responsibility well-being (state – individual)		5.05	5.21		5.39	**
<i>N</i>		603	871		363	

\* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

Note: SW = social workers; GP = general population

#### 4.2. Measurement equivalence and mean comparison: multigroup SEM with alignment

To assess whether the latent variables of *welfare state criticism*, *welfare state sanctioning* and *welfare state overuse* – rather than the separate indicators (see Table 2-1) – differed across the groups, a structural equation modelling approach was warranted. To test the equivalence of the measurement instruments across both groups (Davidov, Meuleman, Cieciuch, Schmidt, & Billiet, 2014), we used the alignment method. The basic idea of this method is that a measurement model is estimated for each group, and that one subsequently evaluates the extent to which the measurement parameters – namely factor loadings and item intercepts – are similar across the two groups. If the latter is the case, this indicates that meaningful comparisons can be made between

social workers and the general population.

Recently, Asparouhov and Muthén (2014) developed the alignment method as a flexible way to identify the model that provides maximal cross-group comparability.<sup>6</sup> If and only if a sufficiently high level of measurement equivalence is established, can meaningful latent mean comparisons be made (which is the ultimate purpose of this analysis). The final model (Model 1) is presented in Figure 2-1. The adapted model has a good fit: the root mean square error of approximation (RMSEA) equals 0.067 and the comparative fit index (CFI = 0.923) is sufficiently close to 1. All items load significantly on the latent variables (see Appendix 7). The alignment method does not indicate loading differences between the two groups, meaning that the latent variables measure a similar general concept in the both groups.

In this paper we were primarily interested in similarities in three attitude constructs that concern evaluations of the performance of the welfare state. Social workers and the general population differ significantly from each other with regard to all of the three constructs (latent means (Im)). First, social workers are less inclined to see unintended, negative effects of welfare state provisions on the behaviour of citizens (0.70 difference). After taking account of this overall attitude difference, the singular items (*dependent, lazy, no self, costs*) no longer significantly differ across the research groups. This means that differences in ideas about the effects of the welfare state on citizens are all explained by a general difference in evaluating the performance of the welfare state.

Second, with regard to the controlling function of the welfare state, the social workers less often think that there is a need for more monitoring and punishment of benefit users (0.65 difference). Differences in items such as the welfare state being too strict, the need for punishment, or for more control of job searching, are explained by this overall difference in the appreciation of the strictness of the welfare state. However, for the item 'The government should check more closely whether the unemployed do additional illicit work', it is not only the general difference in the latent means that determine the inter-group difference. Contrary to our expectations, social workers who have similar general views on monitoring to other citizens are actually more likely to

---

<sup>6</sup> The alignment method resembles rotation in explanatory factor analyses. It starts from the configural invariant model (i.e. identical factor structures across groups, but no constraints on the parameter estimates) and subsequently uses Monte Carlo Markov Chain estimation to detect a model with the same fit as the configural model, in which there are as few as possible differences between measurement parameters across groups. For more technical details, see Asparouhov and Muthén (2014) or Marsh et al. (2017).

believe that the government should check more closely whether the unemployed do additional illicit work (see Table 2-2.).

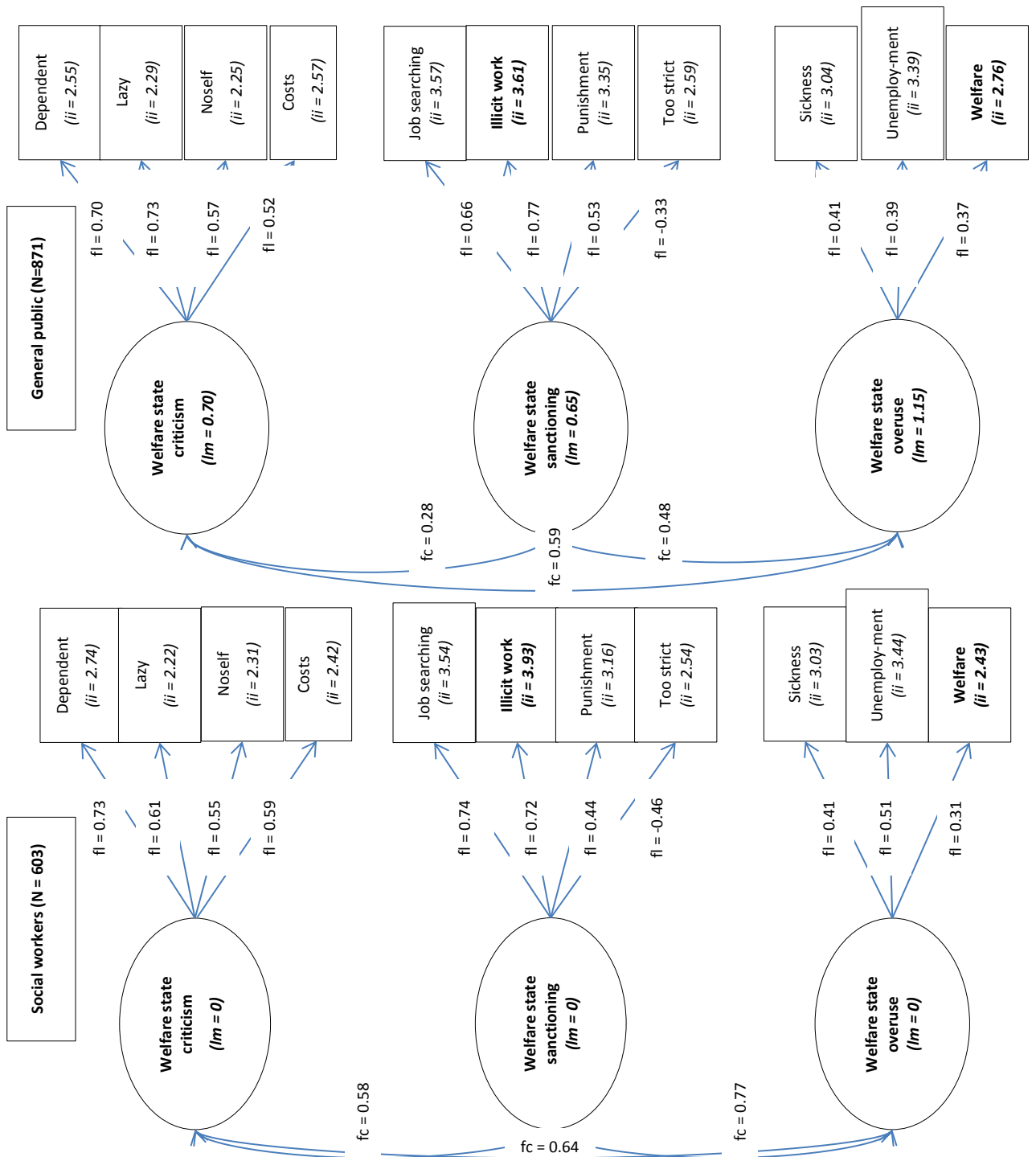
Third, the largest difference concerned the beliefs about benefit receivers obtaining a benefit without being entitled to it (1.15 difference). Social workers generally perceive less overuse, and after controlling for this general difference, there is no difference in the specific items with regard to unemployment and sickness benefit receivers. However, as the item intercept of the overuse of welfare benefits remains significantly different across groups, it can be stated that a social worker who has similar beliefs about benefit overuse to a respondent from the general population (= controlled for in the latent means) is less likely to suspect that welfare benefits are being abused.

#### 4.3. Controlling for relevant covariates

The previous model identified substantial differences in the welfare state attitudes of social workers and the general population. However, at this stage, it cannot be ruled out that these differences are merely the result of the differential composition of both groups. To evaluate the selection hypothesis potential, we added several explanatory variables – socio-demographic characteristics as well as ideological disposition – to the model to determine whether the differences between the latent means in the two groups would diminish or even disappear. In this regard, we performed an ‘alignment within CFA’ model (Marsh e.a., 2017). Model 2 has a good fit, with an acceptable RMSEA (0.038) and a CFI index (0.953) sufficiently close to 1 (TLI = 0.937).

According to the results in Table 2-3, the age of the respondents in the general population has no influence on their views about how the welfare state performs. However, the older the social workers, the more positively they evaluate the consequences of the welfare state, the less they call for more control and the less they suspect beneficiaries of overuse. In both populations, gender has – with one exception (*welfare state overuse* – GP) – no significant effect on the latent attitude constructs. Furthermore, national background has no strong predicting power. Having a non-EU-mother negatively influence opinions about overuse of benefits, but all other nationality categories remain non-significant. It should be noted that the shares of respondents with foreign backgrounds are small in both representative research samples, which means that possible population effects might not become visible in the results based on the samples.

Figure 2-1: Alignment model without explanatory variables (Model 1)



Note: model fit indices of Model1:  $\chi^2 = 422.5$ ,  $df = 98$ ,  $RMSEA = 0.067$ ,  $SRMR = 0.064$ ,  $CFI = 0.923$ ,  $TLI = 0.914$  – All parameters are significant at the  $p < 0.001$  level. – lm = latent mean, ii = item intercept, fl = factor loading, fc = factor correlation – In bold: indicators that differ between the two research populations

Comparing the Z-scores, the political preference variable (left vs right) and views on responsibility for wellbeing (individual vs societal) appear to be the strongest predictors of welfare state attitudes. People who are more right-leaning and see the individual as more responsible score higher on the three latent variables. This effect is stronger among the social workers than among the general population. As shown in the presentation of the descriptive statistics at the beginning of this section, social workers are more left-leaning than the general public (see Table 2-2.). However, being a little more right-leaning (e.g. a 1-point difference on the 11-point scale) has a stronger effect on the attitude constructs in the group of social workers than the same small difference has for the general public.

Table 2-3. Estimates of the structural part of the model (socio-demographic and ideological characteristics), latent means and intercept differences

	Model 1		Model 2		Model 3		Model 4	
	SW	GP	SW	GP	SW		SW	GP
STRUCTURAL MODEL			Coef.(Z-score)	Coef.(Z-score)			Coef.(Z-score)	Coef.(Z-score)
<i>WS criticism &lt;</i>								
Male (ref = female)			-0.01 (-0.1)	0.14 (2.1)			-0.01 (-0.1)	0.13 (1.2)
Age_ref			-0.02 (-4.5)***	0.00 (0.4)			-0.02 (-4.5)***	-0.01 (-1.7)
Father EU (ref=B)			0.45 (0.4)	-0.41 (-2.1)*			0.50 (1.3)	-0.57 (-1.9)
Father non-EU (ref=B)			0.39 (0.9)	1.14 (1.2)			0.39 (0.9)	0.34 (0.0)
Mother EU (ref=B)			-0.46 (-1.3)	0.35 (1.7)			-0.46 (-1.3)	0.43 (1.3)
Mother non-EU (ref=B)			-0.40 (-0.8)	-1.07 (-1.1)			-0.40 (-0.8)	0.34 (0.0)
leftrigh_ref			0.22 (8.3)***	0.12 (7.1)***			0.22 (8.3)***	0.17 (6.6)***
responsi_ref			0.13 (5.2)***	0.07 (4.7)***			0.13 (5.2)***	0.11 (4.2)***
<i>WS sanctioning &lt;</i>								
Male (ref=female)			-0.02 (-0.7)	0.03 (0.5)			-0.02 (-0.7)	-0.01 (-0.1)
Age_ref			-0.01 (-3.0)**	0.00 (1.5)			-0.01 (-3.0)**	-0.00 (-1.4)
Father EU (ref=B)			0.42 (1.1)	0.03 (0.2)			0.42 (1.1)	0.06 (0.3)
Father non-EU (ref=B)			0.40 (0.9)	0.37 (0.5)			0.40 (0.9)	-0.17 (-0.0)
Mother EU (ref=B)			-0.56 (-1.5)	-0.04 (-0.2)			-0.56 (-1.5)	-0.17 (-0.7)
Mother non-EU (ref=B)			-0.10 (-0.2)	-0.67 (-0.8)			-0.10 (-0.2)	-0.17 (-0.0)
Leftrigh_ref			0.24 (8.4)***	0.08 (5.6)***			0.24 (8.4)***	0.15 (7.0)***
Responsi_ref			0.14 (5.2)***	0.04 (3.4)***			0.14 (5.2)***	0.05 (2.4)*
<i>WS overuse &lt;</i>								
Male (ref = female)			-0.04 (-0.9)	0.01 (0.1)			-0.05 (-0.9)	-0.32 (-2.0)*
Age_ref			-0.02 (-4.3)***	-0.01 (-1.4)			-0.02 (-4.2)***	-0.01 (-2.7)**
Father EU (ref=B)			0.04 (0.1)	-0.16 (-0.5)			0.04 (0.1)	0.80 (1.7)
Father non-EU (ref=B)			0.91 (1.8)	3.04 (1.8)			0.91 (1.8)	-0.37 (-0.0)
Mother EU (ref=B)			0.30 (0.7)	0.06 (0.2)			0.30 (0.7)	-0.27 (-0.5)
Mother non-EU (ref=B)			-1.39 (-2.3)*	-3.66 (-2.1)*			-1.39 (-2.3)*	-0.37 (-0.0)



Leftright_ref			0.16 (4.8)***	0.12 (4.1)***			0.16 (4.8)***	0.29 (6.9)***
Responsi_ref			0.15 (4.6)***	0.09 (3.6)***			0.15 (4.8)***	0.15 (3.4)***
MEASUREMENT MODEL								
Item intercepts								
Illicit work	3.93***	3.61***	3.94***	3.61***	3.93***	3.60***	3.94***	3.62***
Welfare	2.43***	2.76***	2.44***	2.76***	2.43***	2.71***	2.44***	2.70***
Latent means								
WS criticism	0	0.69***	0	0.25**	0	0.55***	0	0.09
WS sanctioning	0	0.65***	0	0.43***	0	0.60***	0	0.39**
WS overuse	0	1.15***	0	1.00***	0	0.79***	0	0.87**
N	603	871	603	871	603	363	603	363

\* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

Note: SW = social workers; GP = general population

Model 1 = alignment model without exogenous variables, full sample; Model 2 = alignment within CFA (AwC) with exogenous variables, full sample; Model 3 = alignment model without exogenous variables, only Bachelors and Masters; Model 4 = alignment within CFA (AwC) with exogenous variables, only Bachelors and Masters

Only the item intercepts that proved invariant in Model 1 and Model 3 are shown - for the full Models see Appendix 7

More important, however, is whether the inclusion of these socio-demographic and ideological variables reduces the overall differences in welfare state attitudes between social workers and the general population. Model 2 (see Table 2-3 – compared to Model 1) indeed indicates that the differences in latent means substantially decrease after adding explanatory variables. The decrease is sharpest with regard to *welfare state criticism*. Controlling for socio-demographic and ideological characteristics, the general population's latent mean of this latent variable decreases to 0.25 (compared to zero for the social workers, who served as the reference category). Without these controls, the gap between the general population and social workers was 0.69. For *Welfare state sanctioning*, the latent mean drops from 0.65 to 0.43. For *Welfare state overuse*, the decrease, from 1.15 to 1.00, is the smallest. However, the latent means remain significantly different across the two groups. This indicates that compositional differences between both groups are partly but not completely responsible for the observed discrepancies in latent means. These results are a first confirmation of our hypotheses that both professional identification and selection should be taken into account when comparing the general population to social work professionals.

#### 4.4. Restricted sample: only Bachelors and Masters

Previous analysis was not able to include educational level in the model, although it is one of the most influential predictors of welfare state attitudes (Blomberg e.a., 2013; Hasenfeld & Rafferty, 1989; Weiss & Gal, 2007). Since all Belgian social workers in social assistance agencies need to have at least a Bachelor's degree in social work, there is not enough variation in educational degree for this particular group. To gain insight into whether welfare attitude differences are driven by the composition of both groups in

terms of educational level, we also restricted the general population sample to respondents who obtained at least a Bachelor's degree. The re-estimated Model 3 (see Table 2-3) has an acceptable fit (RSMEA=0.071; CFI=0.926; TLI=0.917). The differences between the latent means for the alignment model for the restricted sample (Table 2-3 Model 3) are indeed smaller compared to the same model for the full sample (Table 2-3 Model 1). For *welfare state criticism* there is a decrease from 0.69 to 0.55, for *welfare state sanctioning* from 0.65 to 0.60 and for *welfare state overuse* from 1.15 to 0.78. This means that the welfare state attitudes of social workers differ less from those of other citizens with higher levels of education than from those of the general public as a whole. More specifically, the *welfare state overuse* latent mean decreases more by restricting the sample to respondents with higher levels of education in Model 3 than by adding covariates (socio-demographic and ideological differences), as we did in Model 2 (see Table 2-2.).

As expected from the selection hypothesis, the differences in latent means between the two groups decrease further by adding socio-demographic and ideological characteristics (gender, age, national background, political preferences – Model 4). The *welfare state criticism* latent mean for the general population decreases from 0.55 to 0.09 and is, as a result, no longer differ significantly from 0. As the social worker's latent mean is fixed at 0 this means that, after controlling for socio-demographics and ideology, the two groups of higher educated citizens do not differ significantly with regard to general ideas about the effects of the welfare state on the behaviour of people. For *welfare state sanctioning* the decrease is from 0.60 to 0.39 and the difference between the two populations becomes less significant. For *welfare state overuse* there was a slight increase from 0.79 to 0.87. The alignment within CFA method did not allow us to conclude whether this is a significant increase or not. Overall, the effect of the exogenous variables of age, gender and political preferences remained similar when we reduced the sample to only Bachelor's and Master's (Table 2-3). The age effect for the general population became significant for one of the latent variables (*welfare state overuse*).

## 5. Discussion and conclusion

Social work and social policy literature has shown that policies are only effective in practice if the professionals who must implement the policy accept the ideology behind it. Looking at social assistance in Europe, we can observe a policy shift towards increased conditionality, activity requirements, monitoring and sanctioning of benefit users since the 1970s (Ditch, 1999; Immervoll, 2009). The question in this paper was whether social workers have followed this ideological shift (Moller & Stone, 2013). The first conclusion

is that the views of social workers about the welfare state differ considerably from those of the general public. This is in line with research on students in different Bachelor's programmes (Blomberg e.a., 2013; Bullock, 2004; Guy, 2011; Weiss, 2005) and with the research on the comparison of social workers and citizens in Scandinavian countries (Blomberg e.a., 2013). Our study, however, extended the current knowledge to a wider range of dimensions of welfare state attitudes, and controlled for more diverse factors. We focused on three crucial dimensions of attitudes on welfare state performance: (1) the perceived negative economic and moral consequences (*welfare state criticism*); (2) support for more control and punishment of welfare beneficiaries (*welfare state sanctioning*); and (3) perceived overuse of welfare benefits (*welfare state overuse*). Social workers are less critical concerning the consequences of the welfare state and less in favour of more control over citizens. It seems that they do not completely agree with the ideological shift that perceives the existing welfare state as a threat to the economy and detrimental to the behaviour of citizens, an ideological shift that is apparent in the attitude of the general public and in the rationale behind several policy decisions in Europe.

Second, we found evidence for both the selection and the professional identification hypotheses as explanations for the attitude differences. Only citizens with particular characteristics seem to choose to become social workers or to work with people in poverty (Weiss & Gal, 2007). However, even after taking these composition effects into account, the latent attitude variables remained significantly different across the two research groups (Blomberg e.a., 2013; Bullock, 2004). Our research setup ultimately made it difficult to pinpoint exactly why this was the case. Omitted variables could be an explanation. Another explanation might concern the process of professional identification through education, experience and self-interest, and should be investigated in greater depth in future research. We are inclined to follow the latter explanation, as, with regard to one of the attitude dimensions – *welfare state criticism* (see further) – the explanatory variables included do were sufficient to eliminate the difference between social workers and the general public.

Third, the partial acceptance of the selection hypothesis has two consequences for social work training and social policy practice. The first concerns the fact that only citizens with particular characteristics and preferences choose to become professionals who implement highly contested policies, and this makes these policies fragile. Current social work research suggests that the use of discretion – partly guided by attitudes (De Wilde & Marchal, 2018) – is difficult to manage through legislation (Evans, 2011; Tabin & Perriard, 2016; Tummers et al., 2012). Policymakers should be aware of the potential for their policies to encounter opposing forces in practice. The second consequence

concerns the fact that, to a certain extent, the effect of both secondary education and on the job training seem limited in orienting social workers to a common set of attitudes and values, which is the express intention of social work education (Sewpaul & Jones, 2005). In our study, the largest part of the attitude gap between social workers and the general public was explained by characteristics of the respondent. Controlling for having a higher education, without specifying which one, led to a big decrease in differences between the two populations. This means that the specific education 'social work' or on the job training only have a limited effect.

This brings us to one limitation of our study. By modelling political orientation (right vs left) and perceived responsibility for wellbeing (state vs individual) as selection variables, we consider them as characteristics on which education and on the job training had no influence. This is, of course, incorrect. Unlike variables such as age, gender and country of origin, ideological preferences are subject to change as a result of the context in which people live, study and work. However, there is an established body of literature on the finding that students who choose to study social work are more left-leaning than other students (Csikai & Rozensky, 1997; Hackett, Kuronen, Matthies, & Kresal, 2003).

The differences between the two groups that cannot be explained by composition effects leads us to a fourth conclusion: some of the results suggest that when the dimension covered by the attitudes concerned the daily practice of social workers, they disagreed more about it with the general public. After controlling for educational level, age and political preferences we no longer found differences with regard to the attitude dimension concerning the overall consequences of the welfare state (*welfare state criticism*). However, the two other attitude dimensions remained significantly different. Not surprisingly, these were attitudes concerning concrete practices, such as offering benefits and monitoring clients, which are explicit tasks of social assistance workers. Furthermore, the tendency of social workers to be less suspicious about the overuse of welfare benefits accords with this reasoning. As they are the civil servants who are responsible in part for allocating benefits, it seems only reasonable that they are less likely to believe that beneficiaries abuse the system. This may be due to experience (they know the system), but also to self-interest: they may wish to protect the social assistance system in which they work. Future research should look for more evidence to support this statement and focus on the question of whether experience or self-interest guides this positive view on aspects of the welfare state that are more closely tied to professional practice. If experience, for example, is the general explanatory factor, it seems reasonable to increase the involvement of social workers in policymaking.

Finally, it is fair to say that the attitudes of social workers about the system they work in are to a large extent the product of the characteristics and preferences of the individual social worker. However, with respect to some specific attitude constructs, social workers can be considered to be in a category of their own.



## Weighing up work willingness in social assistance:

A BALANCING ACT ON MULTIPLE LEVELS



### Chapter 3. Weighing up work willingness in social assistance: a balancing act on multiple levels

---

*Marjolijn De Wilde and Sarah Marchal*

*Working on a major revision for the European Sociological Review*

#### **Abstract**

In all European countries, social assistance receipt is conditional upon the willingness to work. Yet despite the harsh consequences of losing social assistance, we know surprisingly little about how social assistance agencies and social workers implement this policy in day-to-day practice. In this paper, we focus on three important questions regarding the implementation of work willingness as a condition for continued social assistance benefit receipt. First, how does the actual implementation of the work willingness condition take place in light of specific client characteristics, circumstances and behaviour? Second, is the interpretation of such behaviour similar across professionals and municipalities, or does the combination of vague work willingness legislation and a decentralized organisation lead to variation in implementation? Third, can such variation be seen as the express objective of decentralization and personalized work willingness assessments? We build on an innovative and purpose-designed factorial survey of social workers in Belgium. We identified the determinants of 582 social workers' sanction decisions upon a job refusal, clustered in 89 municipalities, on almost 5000 experimentally varied client cases. These unique data allow to distinguish between the effects of individual client characteristics, characteristics of the social workers assessing the individual cases and the characteristics of the local welfare agency and municipality in which she operates. Moreover, we assess how characteristics within and between these levels interact. In line with the literature, we find substantial variation in sanctions related to work unwillingness at the client level, that can be explained by individual client characteristics. Variation between municipalities is relatively limited, and can be fully explained by municipality characteristics. Surprisingly, we find the largest variation at the social worker level. Whereas some of this variation is random, a substantial part can be explained by the characteristics of the social worker. This finding raises concerns about the unintended consequences of the large discretion awarded to social workers within contemporary social assistance schemes.



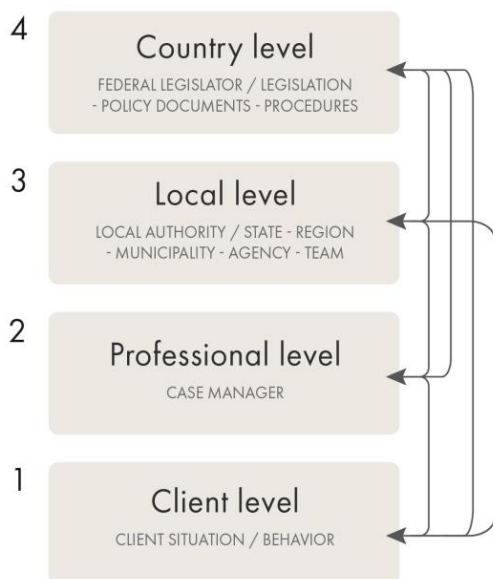
## 1. Introduction

In all European countries, social assistance receipt is conditional upon the willingness to work. Legislation allows or even mandates to withdraw or reduce benefits from beneficiaries who fail to demonstrate a motivation to work, for instance by refusing a reasonable job offer (Marchal & van Mechelen, 2017; MISSOC, 2015). Yet despite the harsh consequences of losing social assistance, we know surprisingly little about how social assistance agencies and social workers implement this policy in day-to-day practice (Pavetti, Derr, & Hesketh, 2003; Scott, 1997).

This lack of research is unsurprising as work willingness is not an objective criterion that can be implemented and investigated in a straightforward manner (Heidenreich & Aurich-Beerheide, 2014; van Berkel, 2010a, 2010b; Watkins-Hayes, 2009). Although some guidelines, such as suitable job offer specifications can be listed in legislation, it is impossible to completely prescribe behavioural conditions. Its implementation necessarily hinges on decisions of social assistance beneficiary's professionals, who are trained to translate legislation to individual cases (Meyers e.a., 1998). In addition, social assistance is notoriously decentralized, both for historical reasons and to facilitate taking account of local circumstances (Sabatinelli, 2010). More recently, scholars and policy makers alike have started to see decentralization as an incubator for innovative policies (Kazepov, 2010). In sum, it is fair to expect that the specification and the assessment of behaviour that can be considered as work (un)willingness are to a large extent subject to both the organisation and the social worker that is following the client.

In this paper, we focus on three important questions regarding the implementation of work willingness in social assistance legislation that have so far remained under investigated. First, how does the actual implementation of the work willingness condition take place in light of specific client characteristics, circumstances and behaviour? Second, is the interpretation of such behaviour similar across professionals and municipalities, or does the combination of vague work willingness legislation and a decentralized organisation lead to variation in implementation? Third, can such variation be seen as the express objective of decentralization and personalized work willingness assessments? To open up this black box of social assistance activation policies' implementation, we bring together academic research streams on conditionality, decentralization, discretion and professionalization (De Wilde & Goos, 2017; Priem e.a., 2011; Rice, 2013).

Figure 3-1. Overview of the levels determining social policy implementation



This paper is the first study to quantitatively assess these questions while expressly taking account of the different relevant implementation levels (see Figure 3-1). Past research on the topic generally focused on only one or two levels of decision-making. Examples include descriptions of national or local legislation (e.g., Eleveld, 2016), studies on large-scale register data (Soss e.a., 2011) and evidence in qualitative case studies or in case law on the specific considerations concerning particular cases (e.g. Brodtkin, 2011). A detailed examination of the determinants of work conditionality assessments at different levels is thus far lacking.

We build on an innovative and purpose-designed factorial survey of social workers in Belgium. We identified the determinants of 582 social workers' sanction decisions upon a job refusal, clustered in 89 municipalities, on almost 5000 experimentally varied client cases. These unique data allow to distinguish between the effects of individual client characteristics, characteristics of the social workers assessing the individual cases and the characteristics of the local welfare agency and municipality in which she operates. Moreover, we can assess how characteristics within and between these levels interact. Belgium is a particularly interesting case for an analysis of work willingness implementation as work conditionality is a key factor in federal social assistance legislation. Actual implementation is left to local social assistance agencies and to the professionals handling client cases.

In the following section, we identify for each of these levels likely determinants of the implementation of a work willingness related sanction. Next, we describe the Belgian social assistance scheme. We then present the data and the multi-level method we use to identify the determinants of sanctioning implementation at the client, agency and social worker level, and their interactions. After the presentation and discussion of the results, we conclude.

## **2. Literature review**

It is an open question which considerations at the client, the municipality and the professional level, may impact on the implementation of work unwillingness sanctions. Social assistance legislation is generally vague on work willingness, so that actual assessments can take individual client characteristics and circumstances into account (Meyers e.a., 1998). Certain characteristics are by most people and in most cases considered as reasons to sanction clients (van Oorschot e.a., 2017), which is what we focus on in the first section. Yet the different levels at which work willingness is assessed may give rise to variation in actual implementation. In the second section, we review the legislator's express intent and acceptance of variation by decentralising policy implementation. The last section focusses on the inevitable variation due to the human professionals applying legislation in real-world cases.

### **2.1 The assessment of client characteristics**

Client characteristics that should or should not lead to sanctions are in some countries to a certain extent detailed in legislation. However, most of the social assistance legislation is aimed at a personalized assessment of a client's individual situation. It can be expected that such an assessment is influenced by the same considerations that apply for the society at large when thinking about solidarity, as a common personal and professional understanding (Jasso, 2006; Keiser e.a., 2004; Taylor, 2006; Wallander, 2012b). Such an understanding would be guided by the assessment of 'deservingness' (van Oorschot, 2000). We expect that the need to demonstrate work willingness or to be exempt originates from the traditional deservingness criteria, namely reciprocity, attitude, control, need and identity (De Wilde, 2017; van Oorschot e.a., 2017).

There is only little quantitative evidence on how deservingness characteristics are weighted against activity-related infringements and possible exemptions for work willingness in treatment reality. Studies based on administrative data only provide insight into which groups are more often sanctioned. These studies show that sanctioned recipients are often foreign, young, never married, poorly educated, parents, sick, caring for a sick child, addicted, experiencing domestic violence, long-term welfare

recipients, experiencing human capital deficits such as limited education, or have a poor work history and transportation problems (Cherlin, Bogen, Quane, & Burton, 2002; Fording e.a., 2007; Hasenfeld, Ghose, & Larson, 2004; Keiser e.a., 2004; Monnat, 2010; Pavetti e.a., 2003; Soss e.a., 2011; Tabin & Perriard, 2016). Groups that are less often sanctioned are people who are pregnant, (young) parents, disabled, participating in mental health or substance abuse treatment, experiencing domestic violence, have unstable housing, difficulties in finding childcare, are non-native language speakers or enrolled in language courses (Cherlin e.a., 2002; Hasenfeld, 2010; Keiser e.a., 2004; Maloy, Pavetti, Shin, Darnell, & Scarpulla-Nolan, 1998; Pavetti e.a., 2003; Rehwald, Rosholm, & Rouland, 2016; Tabin & Perriard, 2016).

Some of these sanction categories are in line with deservingness criteria, as some of the groups that are more often sanctioned can be seen as having a lower need (e.g. the young), to be more responsible for their situation of need (e.g. addicted), or seen as different (e.g. of foreign background). Also when considering groups that are less often sanctioned, a higher need and less control over one's own situation appear to be characteristics that are taken into account as mitigating factors when deciding upon the implementation of a sanction. Yet for some of the groups with higher sanction rates we would rather expect a more lenient treatment based on the deservingness criteria. The administrative datasets on which these studies build however only include static information, and do not provide information on the activity infringements that gave rise to these sanctions (Keiser e.a., 2004). It is conceivable that groups confronted with numerous problems will have more trouble to adhere to activity requirements than others. This indicates the need to assess client characteristics and transgressions in combination when looking into the implementation of work willingness criteria.

Considerations based on client characteristics can be summarised in the following hypotheses:

*H1: Sanction treatment in case of work unwillingness depends on the clients' characteristics.*

*H1a: Client characteristics that reflect a negative (work) attitude and high control over one's own situation increase the client's likelihood to be sanctioned upon a clear manifestation of work unwillingness.*

*H1b: Elements that make unwillingness to work understandable in line with the deservingness theory will decrease the sanction likelihood.*

## 2.2 Devolution to the local level

In most countries, social assistance policy implementation (and often even policy making) is devolved to the local level. Reasons are manifold, and include increasing the legitimacy of anti-poverty policies (Fording e.a., 2007; Kazepov, 2010; Soss e.a., 2011), adapting social assistance and activation measures to the local labour market situation (Fording e.a., 2007; Kazepov, 2010; Sheely, 2012; Soss e.a., 2001; Vandenbroucke, Luigjes, & Lievens, 2016), and opening up opportunities for policy innovation (Kazepov, 2010). This literature leads us to expect that decentralization of social assistance policy implementation will result in local variation in line with these motives to decentralize.

Earlier research has indeed found local variation in social assistance outcomes that can be linked to these explicit aims. Some empirical studies have shown that sanction rates vary in line with the local socio-demographic and economic situation, although the evidence is mixed and direction of the association is unclear (Bell, 2005; Fording e.a., 2007; Keiser e.a., 2004; Monnat, 2010; Soss e.a., 2011; van Oorschot, 2000). Sanction rates do appear to be higher in more conservative regions (Fording e.a., 2007) and implementation of activation measures depends on the importance local policy makers attach to it (May & Winter, 2009).

The local level can have an impact in yet another way. Local welfare agencies' organizational structure may differ in ways that are unlikely to be neutral. Blom (2004) highlighted the likely difference between specialized and more generalist teams, arguing that the latter allow for a more holistic view of clients and to consider more aspects of a client's life simultaneously. Other researchers have stressed the importance of decision structures. Structures that allow the professional, team and team manager to be actively involved in decision making are claimed to have a decreasing impact on sanction decisions (Jessen & Tufte, 2014; Raeymaeckers & Dierckx, 2013). Raeymaeckers & Dierckx (2013) expect such an involvement to boost the professionals' creativity in finding solutions for difficult clients, which would in turn result in fewer sanctions. Other qualitative studies furthermore hint at the extent to which the importance of work is stressed by the head of the team or by colleagues (Brodkin, 2011; Fording e.a., 2007), to the effect of staff workloads and to the complexity of the service delivery system (Fording e.a., 2007; Meyers e.a., 1998). Experimenting with these elements can be considered as a local agency's quest for the best strategies to combat poverty. Furthermore, the composition of a team may indirectly influence actual implementation (Rice, 2013).

In sum, we hypothesise that:

*H2: There will be variation in work willingness implementation across municipalities.*

*H2a: This variation can be partly explained by socioeconomic and demographic characteristics of the municipality .*

*H2b: Local implementation of work willingness will be in line with local political preferences.*

*H2c: The overall specialisation of the professionals in a team will increase sanctioning.*

*H2d: The more involved professionals are in the decision-making process, the lower the implementation of sanctions.*

*H2e: The local composition of the professional team and the preferences of the team manager will have an effect.*

### 2.3 The discretion of the professional

Professionals are important actors with regard to implementation of work willingness (Evans, 2011; Lipsky, 1980). As it is a rather vague eligibility condition, that only to a certain extent can be detailed in legislation, it leaves substantial implementation discretion to social workers. Kazepov and Barberis (2013) term such discretion, that arises from a lack of concrete (or consistent) rules *interpretative discretion*. Whereas leaving important discretion in the hands of individual caseworkers may give rise to substantial variation in work willingness treatment, its main aim is to translate existing legislation to local circumstances in line with client needs. We would therefore expect a tendency towards a general treatment that mainly depends on client characteristics, and on local circumstances. Remaining variation should then ideally be random.

Qualitative research, however, gives some evidence of explicators that put doubt on the random character of variation at the professional level. First, several social work scholars acknowledge the likely importance of socio-demographic characteristics of the professional (Dubois, 2010; Rice, 2013; Taylor, 2012), yet they do not provide clear indications that could inform hypotheses. Based on preparatory talks with stake holders and on generation literature, we expect older professionals to sanction less. For one, the more life and work experience professionals have, the more they are aware of the difficulties people in poverty face and the more they reframe non-compliant behaviour. In addition, younger people are more easily influenced by new ideas, so they might be more influenced by the shift towards deservingness considerations in general public opinion (Blomberg e.a., 2013). This contradicts with findings for the general population, where older people are usually more strict (van Oorschot, 2006). Socio-demographic characteristics may also matter as they impact on feelings of identity with the client. The identity hypothesis formulated by Van Oorschot (De Wilde, 2017; van Oorschot e.a., 2017) state that we are more likely to help those that are similar to us or have

experienced similar problems, as we can more easily imagine ourselves to be in their situation (Krumer-Nevo & Lev-Wiesel, 2005; Watkins-Hayes, 2011). Second, job related characteristics, such as caseload, work regime, the extent of specialisation, and possibilities for counselling with colleagues or experts are often assumed to have an effect on the preference and behaviour of professional and, hence, impact on treatment (Bell, 2005; Fording e.a., 2007; Godfrey & Yoshikawa, 2012; Hasenfeld e.a., 2004; Lipsky, 1980; Meyers e.a., 1998; van Berkel & Knies, 2016). Finally, psychology literature consistently finds that behaviour is to some extent predicted by attitudes linked to this behaviour (Kraus, 1995). When social workers are granted a degree of discretion to assess clients' deservingness in line with legislation and local circumstances, it then only seems logical that their general ideas about the goals and the functioning of the welfare state will influence their actual treatment decisions (Bell, 2005; Blomberg e.a., 2013; Brodtkin, 2011; Castillo & Becerra, 2012; Cozzarelli e.a., 2001; De Wilde, 2016; Reingold & Liu, 2009).

In sum, we expect professionals' characteristics to impact on sanction implementation in the following ways:

*H3: There will be variation in work willingness implementation across professionals.*

*H3a: We expect older professionals to sanction less. Also other socio-demographic characteristics of the professionals may impact on the sanction probability.*

*H3b: Professionals will sanction clients with characteristics similar to their own less frequently, in line with the identity hypothesis.*

*H3c: Job characteristics will impact the sanctioning behaviour of professionals.*

*H3d: Negative welfare state attitudes held by professionals will be associated with more sanctions.*

### **3. The Belgian situation**

We test our hypotheses using a purpose-designed survey conducted in Belgium. Social assistance in Belgium is regulated by federal law and implemented by local welfare agencies. In short, federal law guarantees financial assistance, access to a project of social integration, or both, to all adult citizens with the Belgian nationality or a legal residence status who i) are willing to work , ii) have an income that is below a certain level (Law 'Recht op Maatschappelijke Integratie', 2002, henceforth RMI law). These conditions are assessed by professional social workers at the local (municipal) level. Financing is divided between the federal and the local level, with the federal level responsible for 60% to 100% of the awarded benefits, depending on the local caseload burden.

The RMI law is a framework law in the sense that it does not detail how social workers should interpret the willingness to work is not detailed in the RMI law, (Stranz, Karlsson, & Wiklund, 2016). It states that sanctions for activation related infringements can amount to a suspension or a temporary withdrawal of the benefit. It further refers to reasons for exemption from activation as 'health reasons' or 'fairness reasons', but no clearly described categories are included. The only detailed condition involves enrolment in full-time studies that will improve the client's chances on the labour market. In documents accompanying the RMI-law some examples are added. Health reasons might be invoked in the case of a young drug-addicted person receiving medical care or a pregnant woman exempted from doing physical labour. A possible 'fairness' reason might involve a single mother with several (perhaps disabled) children and difficulties in transportation (Vande Lanotte, 2002a). Clearly, the RMI law leaves important leeway to local regulations and preferences and the professional opinion and assessment of social workers.

At the local level, each Belgian municipality has a social welfare agency headed by a politically composed non-professional board. Decisions about whether claimants are eligible for financial assistance are made by this board – within the framework of the law – based on client files put together by the social workers who are in contact with the claimants. Even though the federal law stipulates a number of fixed elements in the social investigation, municipalities still have considerable freedom in how exactly to organize the process leading up to these files (which are essentially social workers' recommendations on individual cases to the decision board) and the files themselves.

The professionals' tasks may differ between municipalities. In some municipalities the same professional will guide clients through their entire period at the welfare agency. Elsewhere, tasks are divided, with some professionals being responsible for activation measures, others only for intake, etc. Nonetheless, it is always a qualified social worker (minimally a Bachelor's degree in social work) who prepares the file (with recommendations) for the board. In the preparation of this file, the social worker can use her own discretionary judgement to highlight certain options to move forward with a specific client and to assess the eligibility conditions, including the willingness to work or possible 'health or fairness' reasons.

In sum, both the decentralization and the explicit recognition of the professional as a qualified professional with important interpretative discretion within the RMI law make Belgium an interesting case to look at the different determinants (and their interaction) of willingness to work implementation.



## 4. Data and method

### 4.1 Data

We use a factorial survey carried out in 89 Belgian<sup>7</sup> municipalities among 582 social workers involved in eligibility decision-making. Factorial survey experiments require each respondent to read stories (vignettes) about individuals or situations before rating these individuals or situations. Every vignette consists of several factors (e.g. gender), each with their own levels or categories (e.g. male and female). We asked the respondents to rate experimentally varied vignettes on hypothetical social assistance clients according to the likelihood that they would lose their social assistance benefits if they refused a job offer. Each respondent rated nine vignettes that varied across 14 client characteristics. We selected the vignette attributes on the basis of the existing literature and on interviews with professionals, team leaders and academics (Taylor, 2012; Wallander, 2009). The characteristics kept constant in all vignettes were age (22 years), marital status (single), income (none), debts (none), contact with parents (none) and physical health (healthy). The 14 client characteristics that varied concerned socio-demographic characteristics, and characteristics that according to the deservingness literature will impact on professionals' assessment of individual situations, i.e. characteristics pointing at work willingness and possible exemption criteria.

These experimentally varied vignettes were D-efficient<sup>8</sup> distributed over the respondents. The respondents themselves were selected in a two stage (first municipalities, then respondents) stratified probability sample. We invited 105 municipalities, 15 of which declined to participate mainly due to time constraints and one that did not provide enough responses to be included in the research. In the second stage, 839 social workers were invited to fill out the online survey, of which 582 provided a completed questionnaire (see Appendix 3 for more information on the sampling procedure). The survey ran from March to November 2015. The resulting survey is particularly innovative as it not only includes the respondent's assessments of vignettes, but also information on these respondents and the organisations they work in. This provides us with three different levels of analysis: the experimentally varied

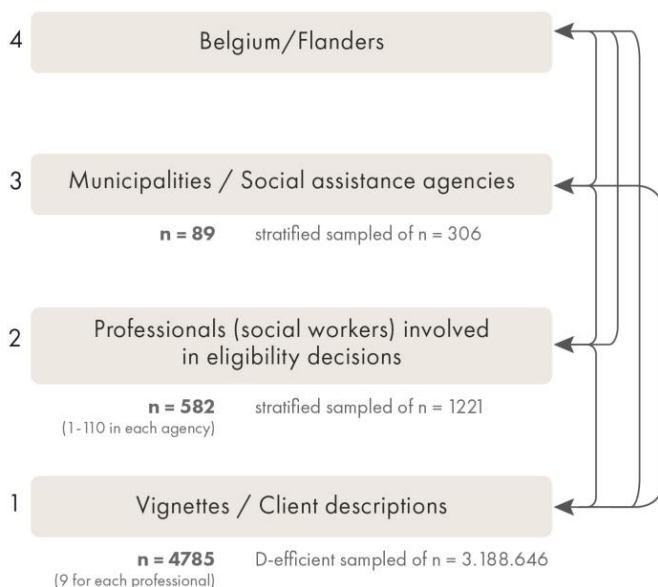
---

<sup>7</sup> All municipalities are located in the Dutch-speaking part of Belgium (Flanders). Results hence refer to the Flemish municipalities. In the text, we discuss the Belgian situation, as all policy levers were at the Belgian or the local level (and not at the regional level) at the time of the survey.

<sup>8</sup> The D-efficient sample consists of an orthogonal design in which, for each factor (e.g. gender), the categories (e.g. male/female) occur as equally often as possible (Atzmüller & Steiner, 2010; Auspurg & Hinz, 2014; De Wilde & Goos, 2017; Dülmer, 2007, 2016).

client characteristics, the characteristics of the social workers themselves, and those of the local social assistance agencies and municipalities they operate in.

Figure 3-2. Study design and sample



Self-evidently, the vignette method is not without flaws. Questions have been raised about the external validity of the method<sup>9</sup> (Auspurg & Hinz, 2014; Taylor, 2006; Wallander, 2009). This criticism equally applies to our factorial survey, with the added risk that respondents' answers may lack realism as they were asked to predict the final decision of the political, non-professional board on each specific case rather than their own preferences. Whereas professionals in Belgium do have an impact on this final decision through the case file and case recommendations they formulate (see above), ultimately this decision is not in their hands, which increases the risk of response error. However, we should not overstate this risk. Respondents indicated that they expected the board to follow their legally required advice in 80% of all files. This means that even if the results in the survey are based on what the professional would advise (rather than on a prediction), this would be close to any practical decision in reality. Even more importantly, a sanction is typically initiated by the professional: she decides when the

<sup>9</sup> Its internal validity is high since the method operates on an experimental basis (Taylor, 2006; Wallander, 2012b) which ensures that no systematic errors occur.

work willingness requirement is sufficiently violated to propose a sanction to the decision board (Fording e.a., 2007; Pavetti e.a., 2003).

## 4.2 Method

We estimate a multi-level model<sup>10</sup> of the determinants of the respondent's assessment of the likelihood of a sanction for 4700 hypothetical clients with different experimentally varied characteristics after declining the offer of a job that starts early in the morning (5 am)<sup>11</sup>. We include three levels: the client level, the social worker level and the municipal (welfare agency) level.

In

$$y = \alpha + \beta_1 x_{1ijk} + \dots + \beta_n x_{nijk} + \varepsilon_1 x_{1ijk} x_{2ijk} + \dots + \varepsilon_n x_{n-1ijk} x_{nijk} + \gamma_1 x_{1jk} + \dots + \gamma_m x_{mjk} + \zeta_1 x_{1ijk} x_{1jk} + \dots + \zeta_n x_{nijk} x_{mjk} + \delta_1 x_{1k} + \dots + \delta_p x_{pk} + z_k + u_{jk} + e_{ijk},$$

i indexes individual clients, j the social workers, and k the municipal welfare agencies. y then is the likelihood of a sanction upon refusal of the job offer (as perceived by the respondents, see section 4.1),  $x_{ijk}$  are the n individual (experimentally varied) characteristics of the hypothetical clients,  $x_{jk}$  are the m characteristics of the respondents (the professionals), and  $x_k$  are the p characteristics of the municipality.  $z_k, u_{jk}$  and  $e_{ijk}$  are the error terms at each level. We estimate the coefficients of the different independent variables x ( $\beta$  at the individual level,  $\gamma$  for the professionals' level and  $\delta$  for the municipal level). In addition, we include specific interaction terms between different client characteristics ( $\varepsilon$ ) in order to assess whether certain characteristics have a different impact on the likelihood of sanction upon a job refusal when they occur in combination. Finally, we also include interaction terms between the characteristics of the professional and the hypothetical client, in order to assess the impact of professionals recognizing themselves in clients ( $\zeta$ ).

The model was built in the following way: We first estimated a model without explanatory variables, in order to determine the unexplained variance at each level (Model 1). In a next step, we only included explanatory variables at the client level

---

<sup>10</sup> All analyses in this paper were done with the stata software package, runmlwin command.

<sup>11</sup> The analyses were also done for two other dependent variables that indicate the implementation of work willingness requirements: the assessed likelihood of a sanction upon the refusal of a job that only lasts one day, and of a job that is not in line with the client's qualifications. Results were broadly in line with those for a job that starts early. The results of these robustness checks are provided in the appendix.

(Model 2). This second model describes the effect of the perceived unwillingness to work and its interaction with possible exemption criteria over all professionals and municipalities included in the survey. Hence we added all client characteristics (vignette attributes) and all interaction terms between the number of work refusals and the exemption criteria to the original empty three-level Model 1. All client characteristics, including those that were non-significant, remained in the model. We re-excluded the interaction terms that proved not significant.

The final model is the result of a stepwise exclusion of variables until we obtained an optimal model fit. In a first step, we entered the independent variables at the second (professional) and third level (municipality). Next, we step by step excluded variables. For each new model we tested the model fit reduction with a likelihood ratio test. We put the alpha-to-remove significance at 0.1, which means that we re-entered a variable if the model fit worsened with  $p < 0.1$  by removing it. This stepwise re-exclusion of variables means that only the highly significant explaining variables are included in our final Model 3. Furthermore, some variables may have an influence, but do not turn out significant, due to low frequency in the respondent sample (e.g. almost no professionals with a foreign background; see Appendix 4 for independent variables with low frequencies). We use z-scores to compare the importance of different variables over the three levels

Table 3-1 shows the full list of variables that are included in our model, with their expected effect on the dependent variable, i.e. the likelihood of a sanction upon a clear manifestation of work unwillingness (the refusal of a job offer), as predicted by the interviewed social workers. We assess how specific client characteristics are weighted in the implementation of this sanction. These characteristics are selected based on the deservingness criteria, and either show a clear additional (on top of the job refusal) violation of deservingness criteria control, attitude and reciprocity, or hint at characteristics that might excuse the client for his or her job refusal (for instance because the situation of need is deemed beyond the client's control, or the need is very high due to adverse life experiences<sup>12</sup>). Aggravating characteristics when deciding upon a willingness to work sanction can be whether it is the first refusal, or the second or third, whether the client had negative work or activation experiences in the past, and a lack of diligence on one or more occasions without a real explanation for this behaviour.

---

<sup>12</sup> Vignettes should be as realistic as possible. Therefore, the adverse life experiences refer to rather detailed circumstances. A full list is provided in Appendix 2

At the level of the professional, we assess the impact of socio-demographic characteristics, the impact of job characteristics, and the impact of the case workers' own convictions. Job characteristics include among other things the level of specialization of the job. The specialization variable reflects the number of tasks (out of four: intake, treatment, activation and management). We also included two latent variables that reflected respondents' views regarding their impact on the decision-making processes in their social assistance agencies. *Decision\_eligibility* reflects to what extent professionals, think that they themselves, the team managers and the team are involved in the actual decision concerning eligibility made by the board. *Decision\_activation* reflects professionals' view on their impact on the decisions regarding clients' activation trajectories. Finally, we include professionals' opinions on the welfare state. We include straightforward opinion statements (developed by the University of Leuven's Institute for Social and Political Opinion Research), as well as three latent variables, that reflect professionals' views on the consequences of the welfare state (*Welfare State Criticism*, higher values reflect negative opinion about the consequences), on the need for more monitoring of welfare state claimants (*Welfare State Strictness*), and on the likelihood of over-use of different types of benefits (*Welfare State Overuse*) (for a full list of the variables see Appendix 4 – for the construction of the latent attitude variables see Appendix 5).

Table 3-1. Independent variables and expected effects with regard to sanctioning likelihood

<b>Level 1 = Client level (n=4785)</b>	
Number of work refusals (1/2/3)	↑
Motivation ( <b>willing to work</b> /wants to study/not motivated to work or study)	↓/↑
Diligence ( <b>yes</b> / missed one appointment / missed several appointments)	↑
Work experience ( <b>none</b> /positive/negative)	↓/↑
Activation experience ( <b>none</b> /one negative/two negative)	↑
Parenthood ( <b>no</b> /healthy child/sick child)	↓
Housing situation ( <b>stable</b> /unstable/homeless)	↓
Mental health ( <b>healthy</b> /undiagnosed depression/ undiagnosed intellectually disabled)	↓
Addiction ( <b>none</b> /slightly/severely)	↓
Life experiences ( <b>none</b> /child loss/youth with poverty and abuse)	↓
<b>Interactions between aggravating and mitigating characteristics</b>	?
<b>Level 2 = Professional level (n=582)</b>	
<b>Socio-demographic characteristics</b>	
Age	↓
Other: Gender ( <b>female</b> /male); Parenthood ( <b>no</b> /yes); Ethnicity ( <b>both parents born in Belgium</b> /both born in Europe/at least one born outside Europe); owner ( <b>n</b> /y); Qualification ( <b>Bachelor's degree</b> /Master's)	?
<b>Identity hypothesis</b>	
Interaction terms client and respondent on parenthood, ethnicity, gender and housing	↓
Experience with long-term unemployment ( <b>no</b> /yes)	↓
<b>Job characteristics</b>	

Seniority	↓
Work regime in percentages	?
<b>Organisational characteristics – measured on the professional's level</b>	
Specialisation (number of tasks) ( <b>one</b> /two/three or four)	↓
Decision_eligibility	↓
Decision_activation	↓
Board follows advice of professional (1-7)	↓
Involvement of board in decision making (1-7)	↑
<b>Ideology</b>	
Primary cause of unemployment ( <b>no control</b> /laziness)	↑
Political orientation (1: left – 10 : right)	↑
Responsibility of welfare of citizens (1: individual – 10 : state)	↓
WS Criticism / WS Strictness / WS Overuse	↑
<hr/>	
<b>Level 3 = municipality/social assistance agency (n=89)</b>	
<b>Socio-demographic characteristics</b>	
Percentages of social assistance receivers	↓
Size municipality	↑
Poverty	↑
Other: Unemployment; welfare index; percentage inhabitants with foreign background; job centre ( <b>no</b> /yes); subsidised psycho-social organisation ( <b>no</b> /yes)	?
<b>Ideology</b>	
Political party of the chair of the board of the agency ( <b>N-VA</b> /CD&V/Open VLD/Sp.a/other)	↓
WS Criticism_munic / WS Strictness_munic / WS Overuse_munic	↑
<b>Organisational characteristics</b>	
Eligibility_munic	↓
Activation_munic	↓
Specialisation (average number of tasks) ( <b>1</b> /2/3/no clear specialisation pattern)	↓
<b>Team composition</b>	
Average age social workers	↓
Other: Percentage female professionals; Percentage professionals with children, average seniority	?

Notes: NVA: nationalist right-wing party ; CD&V: Centrist Christian Democratic party; Open VLD: Centre-right liberal party; Sp.a Centre-left socialist party. Specialization: tasks: intake, activation, treatment, management

At the local level, we include the socio-demographic characteristics of the municipality, the local political views (reflected both by the political party of the chair of the board, and as an additional proxy, the average of the professionals' view on the welfare state), and organizational characteristics of the local welfare agency, including the socio-demographic composition of the team. The organizational characteristics of the local welfare agency include the averages of professionals perception of the decision-making process in the organization (*Eligibility\_munic* and *Activation\_munic*). Overall specialization in the organization is constructed based on the scores on individual specialization: If most professionals in a municipality fulfilled only one tasks the new variable was scored 1, in the reverse case 3. If there was a lot of variation in the individual specialization variable (1, 2 and 3), the category 'other' was chosen. We furthermore included the averages of the socio-demographic characteristics of the professionals in each local welfare agency.

## 5. Results

The respondents were asked to estimate the likelihood that hypothetical clients, refusing a job offer that started early in the morning, would be sanctioned in their office. The predicted treatment variation was substantial. The respondents rated the sanction probability on a 7 point scale. On this scale, all answer categories were chosen almost equally often over the entire population of experimental client cases. We first established at which level most of this variation was situated, by estimating a model without independent variables (Model 1). Unexpectedly, the unexplained variance at the municipality level was remarkably small, with only 6% of all variance situated at that level (Table 3-2, Model 1, Row 4). The unexplained variance at the professional level was more substantive than expected, lying around 49% (Row 5). The variation at the client level was 45%<sup>13</sup> (Row 6).

In a second model we included several client characteristics that could be interpreted by the social worker to assess whether the manifestation of work unwillingness (the job refusal) is a specific circumstance, or part of a more general attitude (see aggravating characteristics in Table 3-1). Furthermore, this model includes deservingness characteristics that might weaken the effect of apparent work unwillingness.

Respondents predicted that it is not certain that a client will keep his or her benefit if a first job offer is refused. For clients without special characteristics, the likelihood of losing the benefit in such a situation was around 52% (Table 3-2, Model 2, Row 10), with an increasing likelihood upon repeated refusals. The likelihood of being sanctioned increased to 70% if the client refused for a third time (Table 3-2, Model 2, Row 22). Not being motivated to work or not being diligent had no significant independent effect (Table 3-2, Model 2, Row 14-20).

We tested the effect of four potential mitigating factors on sanction likelihood upon work unwillingness. Having a child (-13%) or a child that was often and unexpectedly sick (-19%) strongly decreased the sanction likelihood, but it did not make refusing a job completely acceptable. Neither did the presence of a child mitigate the increasing sanction likelihood upon repeated refusals. Other significant, but small effects result from being homeless (-1%), being depressed (-2%), having (undiagnosed) mental health problems (-4%), and having lost a child in the previous year (-4%). The effect of being

---

<sup>13</sup> The two last percentages may be respectively an overestimation and an underestimation of up to 11 ppt due to the set-up of the survey, as some of the client characteristics had to be fixed at the professionals' level.

addicted is not significant as such, but the interaction term is, which means that the decreasing addiction effect became stronger with the frequency of work refusal.

The next step in the analysis aimed to explain the variation in sanction likelihood among professionals and municipalities by including variables at these levels. Whereas all independent variables listed in Table 3-1 were initially included in the analyses, for the final model only the variables that could not be excluded without decreasing the model fit were kept (i.e. the final Model 3 – see method section).

Socio-demographic features, be it of the individual professionals, aggregated at the team level, or in interaction with similar client characteristics (in line with the identity hypothesis) had no significant effect on predicted sanction likelihood and are hence not included in Table 3-2<sup>14</sup>. The demographic and socio-economic situation of the municipality did have some predicting power. First, the larger the town, the less likely it is that its professionals predict sanctioning (Table 3-2, Model 3, Row 53). Second, the higher the unemployment rate in a municipality, the more often professionals predict sanctioning upon a job refusal (Row 54). Third, the more beneficiaries are participating in an activation measure, the more professionals predict sanctioning (Row 55). Other characteristics, such as the percentage of social assistance claimants or the socio-economic welfare index of a municipality, did not prove significant. A strong predicting variable was the presence of a job centre in the municipality. Professionals in municipalities with a job centre predicted the sanctioning likelihood to be up to 10% lower than their colleagues in other municipalities (Row 56).

Two of the variables connected to the work situation at the professionals' level proved significant. First, a higher seniority decreased the likelihood to predict sanctioning (Row 44). Second, professionals with two types of tasks (intake, treatment, activation or management) predicted more sanctioning than their colleagues with only one type of task (Row 45). The aggregated variable at the municipal level (average specialisation of the professionals team) was also relevant, but in the opposite sense. The less specialised the agency is organised, the less often its professionals predict sanctioning (Row 58 and 59).

The attitudes of the professionals and the political affiliations of the chair of the agency boards were the variables with overall the largest predictive power. Professionals who perceive the welfare state as a system that has a negative influence on citizens (Row 48) or who think that benefit receivers should be more intensively monitored and

---

<sup>14</sup> In the robustness checks based on alternative job offers (Tables in Appendix 8) some socio-demographic variables did turn out significant.



sanctioned (Row 49) predict more sanctioning. The mean attitudes across all professionals per agency and the attitudes of the head(s) of the professionals' teams (as a proxy for the agency's culture) were not significant. In addition, the political affiliation of the chair of the local welfare agency turned out significant. Sanction predictions were higher in social assistance organizations headed by a representative of the nationalist right-wing party, than those with social-democratic chair persons (Table 3-2, Model 3 Row 64)<sup>15</sup>. The municipalities with chairs from parties other than the nationalist right-wing party did not significantly differ from each other (not shown in the Table 3-2).

As a model fit test we calculated the decrease of the unexplained variance discussed in Table 3-2. The unexplained variance at the professional level decreased 12 percentage points by including professional and organisation characteristics, compared to the model with only client characteristics (Table 3-2, Model 3, Row 5). Moreover, the unexplained variance at the municipality level disappeared after including the independent variables at that level (Row 4). In total, around 26% (Row 7 – 12.74% + 12.79%) of the unexplained variance was explained by adding client, professional and organisation characteristics.

Table 3-2. Multi-level results for professional and agency/municipality characteristics with regard to sanction likelihood

LIKELIHOOD TO BE SANCTIONED	Model1		Model2		Model3	
	Coef	% of total	Coef	comp to M1	Coef	comp to M2
<b>Variiances</b>						
Level 3 = municipality	54.47*	6.0%	51.64*	-5.2%	0	-100%
Level 2 = professional	442.62***	48.9%	405.72***	-8.3%	356.42***	-12.2%
Level 1 = client	407.76***	45.1%	332.21***	-18.5%	332.17***	-0.0%
Total	904.85	100%	789.57	-12.7%	688.59	-12.8%
	Coef	z-score	Coef	z-score	Coef	z-score
Intercept	46.24***	31.69	52.36***	20.77	74.42***	15.33
<b>Level 1 = client characteristics</b>						
<b>Aggravating characteristics</b>						
Not motivated to work (ref = willing to work)			-0.77	-1.14	-0.78	-1.16
Missed one appointment (ref = diligent)			-0.26	-0.22	-0.24	-0.2
Missed several appointments (ref = diligent)			-0.85	-0.73	-0.79	-0.68
Negative work experience (ref = none)			0.13	0.18	0.12	0.18
Positive work experience (ref = none)			0.61	0.88	0.64	0.92
One negative activation experience (ref = none)			-0.39	-0.57	-0.34	-0.5
Several negative activation exp (ref = none)			-0.69	-1.01	-0.68	-0.99
2 work refusals (ref = 1)			11.06***	4.57	11.14	4.87
3 work refusals (ref = 1)			17.21***	7.12	17.17	7.45
<b>Mitigating characteristics</b>						

<sup>15</sup> This effect was even stronger with regard to the two variables used as robustness checks (see first table in Appendix 8).

One healthy child (ref = none)			-13.3***	-19.52	-13.33***	-19.56
One sick child (ref = none)			-18.65***	-26.88	-18.69***	-26.96
Unstable housing (ref = stable housing)			0.35	0.52	0.33	0.5
Homeless (ref = stable housing)			-1.46*	-2.12	-1.529*	-2.22
Depressed (ref = no mental health problems)			-2.43***	-3.54	-2.4***	-3.5
Intellectually disabled (ref = no mental health problems)			-4.05***	-5.58	-4.09***	-5.63
Beginning addiction (ref = none)			-0.41	-0.27	-0.43	-0.28
Severe addiction (ref = none)			-1.71	-1.07	-1.81	-1.14
Lost a child (ref = no special life exp mentioned)			-2.05**	-2.96	-2.12**	-3.05
Violent & abusive upbringing (ref = no special life exp)			-0.41	-0.6	-0.44	-0.64
Interaction between work willingness and exemption criteria						
Several missed appointments * beginning addiction			3.34*	2.01	3.33*	2.0
Several missed appointments * severe addiction			1.3	0.77	1.31	0.78
2 work refusals * beginning addiction			-1.22	-0.73	-1.19	-0.71
2 work refusals * severe addiction			-3.45*	-2.01	-3.39*	-1.98
3 work refusals * beginning addiction			0.33	0.2	0.34	0.21
3 work refusals * severe addiction			-3.0	-1.74	-2.9	-1.69
<hr/>						
Level 2 = Professional level						
Job characteristics						
Seniority					-0.29**	-3.17
2 tasks of intake, treatment, activation or management (ref = 1)					7.19***	3.43
3-4 tasks of intake, treatment, activation or management (ref = 1)					2.49	1.08
Ideology						
WS Criticism					5.11**	2.86
WS Strictness					9.97***	4.21
<hr/>						
Level 3 = Municipalities						
Socio-demographic characteristics						
Number of inhabitants					-	-2.95
Unemployment rate					0.00004**	
% beneficiaries in activation					0.02***	3.96
Job centre (VDAB) (ref. = no)					0.29*	2.38
Job centre (VDAB) (ref. = no)					-10.31***	-3.97
Organisational characteristics						
Most cm's two tasks (ref = 1)					-5.24*	-1.98
Most cm's three or four tasks (ref = 1)					-18.82***	-3.92
No clear specialisation pattern (ref = 1)					-6.63*	-2.0
Ideology / party of the chair of the board						
Centrist Christian Democratic party (CD&V) (ref = nationalist right-wing party, N-VA)					-2.09	-0.67
Centre-right liberal party (Open VLD) (ref = N-VA)					-1.74	-0.41
Centre-left socialist party (Sp.a) (ref = N-VA)					-9.9***	-3.45
Other					-10.25*	-2.33
<hr/>						
N	Municipalities	89	89	89	89	
	Professionals	582	582	582	582	
	Vignettes	4785	4785	4785	4785	

## 6. Discussion and conclusion

In this paper we used a factorial survey to assess how rather vague work willingness criteria in the Belgian social assistance legislation are implemented. We investigated which client behaviour or characteristics aggravated or mitigated the implementation of a sanction upon the a clear violation of the work willingness requirement. Law makers in Belgium and Europe explicitly aim for variation with regard to this assessment, as the express purpose of activation within social assistance is to be tailored to a clients' needs and characteristics. For such a tailor-made assessment, professionals should use their discretionary power (Evans, 2011; Kazepov & Barberis, 2013; Lipsky, 1980), within a local welfare agency that takes account of the local context (Fording e.a., 2007; Kazepov, 2010; Soss e.a., 2001; Vandenbroucke e.a., 2016). This paper is the first to quantitatively assess the determinants of implementation at these three different levels, allowing us to identify and compare the impact of each level vis-à-vis one another.

We expected to find variation on all three levels included in our multilevel model. We hypothesized this variation to be substantial at client and municipality level, and in line with client and municipality characteristics, and present, but mainly random, at the professional level. In line with our hypothesis (H1), we did find substantial variance at the client level. Sanction treatment is clearly tailored treatment and depends on both the needs and the responsibilities of the client. We found a strong effect of client's behaviour that clearly indicated work unwillingness, namely the refusal of job offers. The likelihood to be sanctioned after a first work refusal is considerable (50%) and it only increases with further work refusals. Exemption criteria, selected in line with the deservingness criteria, have a modest mitigating impact. The largest effect stems from having sick children, but even then the predicted sanction likelihood after one refusal remains around 33%. Other characteristics indicating need or absence of control had almost no effect. The combination of the strong effects of aggravating behaviour, with only a weak mitigating impact of possible exemption criteria - and a quasi-absence of interaction effects between both - might explain the higher sanction rates among vulnerable groups found in administrative registers (Cherlin e.a., 2002; Fording e.a., 2007; Hasenfeld e.a., 2004; Keiser e.a., 2004; Monnat, 2010; Pavetti e.a., 2003). For these individuals activation infringements might go hand in hand with their higher deservingness. Our study showed that characteristics such as need and lack of control only partly overrule demonstrated work unwillingness in (predicted) sanction decisions.

The variance at the client level that can be explained by specific client characteristics is an indication of a deservingness assessment that is similar across agencies and professionals. As expected, our results also show variation in work willingness

assessment at the municipal and the professional level. At the municipal level, this variation is surprisingly small. Despite a quite strong decentralization in Belgium that should allow for establishing local policies (Vandenbroucke e.a., 2016), municipalities do not appear to develop different sanction treatments in practice. Furthermore, we fully explained the variation found by the hypotheses put forward: local socioeconomic context (H2a), political party of the board of the organisation (H2b) and organisational characteristics of the social assistance agency (H2c-d). Socioeconomic characteristics of the municipality clearly have an impact. Both the unemployment rate and the presence of a job centre in the municipality are influential. Clearly, as unemployment is high, social assistance clients are expected to take the rare available job opportunities. The decreasing effect on sanction predictions of the availability of a job centre is harder to interpret. In future research we should investigate whether this effect remains after inclusion of potentially omitted variables. A second reason to decentralize policies is to increase the legitimacy of these policies for local governors and citizens (Fording e.a., 2007; Kazepov, 2010; Soss e.a., 2011). This led us to expect that political preferences influence treatment decisions. This hypothesis was confirmed in our study, as professionals in social assistance agencies governed by the Flemish right wing party predicted sanctions to be more likely. Finally, we hypothesized that the organizational characteristics are unlikely to be neutral. We did find that the specialization choices of the agency appeared to matter: the more specialized the teams, the more the professionals tended to predict sanctioning (Blom, 2004). These organisational differences can be seen as ways to experiment with new local settings aimed at a more efficient poverty reduction (Kazepov, 2010).

The most surprising finding of this study is the substantial variation at the professional's level. Given the lack of clear work willingness criteria set out in the social assistance legislation, professionals have an important degree of discretion to interpret legislation while tailoring it to the needs of the client. Self-evidently, this should result in treatment differences between colleagues. However, the magnitude of this variance, that is as substantial as the variation found at the client level, raises the crucial question whether this variation is indeed random and unavoidable in a context of professional discretion, or whether there are specific explanatory factors. An important finding of this paper is that it shows that this variance is not completely random. The professionals' attitudes have a high influence on sanction predictions. Their perception of the need to monitor and sanction welfare state recipients in general strongly influences their predictions concerning sanctioning the hypothetical clients in our study. This observation confirms the previously claimed but seldom empirically investigated idea that attitudes and beliefs may influence treatment choices (Blomberg e.a., 2013; Castillo & Becerra, 2012;

Cozzarelli e.a., 2001; Reingold & Liu, 2009). This personal attitude effect remained relevant also after controlling for the general attitude in a social assistance agency and the attitudes of a team manager. The hypotheses on socio-demographic characteristics (H3a), identity (H3b) and job characteristics (H3c) were not, or only partially confirmed.

A second observation is that a large part of the variation at the professional level remained unexplained. It needs to be said, that part of this unexplained variation among professionals may be caused by methodological shortcomings, in particular the external validity of the method. Vignettes remain descriptions, not real life observations (see method section). Further, even though we took great efforts in constructing realistic vignettes, the method leaves important aspects, such as body language and presence, to the imagination of the social worker, which might inflate variance to some extent. Second, despite our thorough literature review, we might miss important explaining factors not integrated in our analyses, possibly decreasing the randomness of the variation. However, if these factors are not determined, the policy question is then whether the high unexplained variance at the professional's level is the result of an inevitable random variation that allows professionals to use their creativity in treating clients or that this variation is actually too high to call it a natural and desirable variation. Both issues represent an important future research agenda, as the policy and theoretical implications are large.

Table 3-3. Overview of affirmed and not affirmed hypotheses

	Hypothesis	Affirmed
Variation at which level?	H1: Sanction treatment depends on the clients' characteristics	Yes
	H2: There will be variation in work willingness implementation across region or municipalities.	Limited
	H3: There will be variation in work willingness implementation across professionals (in the same organisation).	Yes
Explaining factors at client level	H1a: Client characteristics that are a sign of unwillingness to work increase the client's likelihood to be sanctioned.	Yes
	H1b: Elements that make this unwillingness understandable (in line with the deservingness theory) will decrease or eliminate the sanction likelihood.	Limited
Explaining factors at Municipality level	H2a: Socioeconomic and demographic characteristics of the municipality will have an effect on the assessment of work willingness	Yes
	H2b: Local implementation of work willingness will be in line with local political preferences.	Yes
	H2c: The specialisation of teams will increase sanctioning.	No
	H2d: The more involved professionals are in the decision-making process, the lower the predicted implementation of sanctions.	No
	H2e: The local composition of the professional team and the preferences of the team manager will have an effect.	No
Explaining factors at professional's level	H3a: We expect older professionals to predict sanctioning less and other socio-demographic characteristics to have an effect on the predicted sanction probability.	Limited
	H3b: Professionals will predict to sanction clients with characteristics similar to their own less frequently (identity hypothesis).	No
	H3c: Job characteristics will impact the predicted sanctioning behaviour of professionals.	Limited
	H3d: Professionals own beliefs on social policy will influence their predictions of implemented sanctions.	Yes

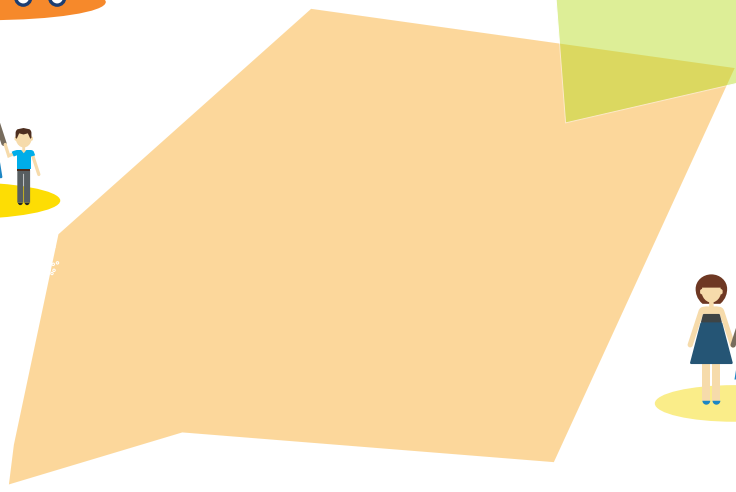
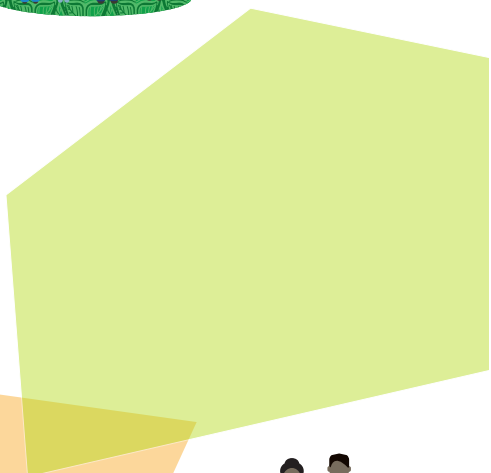
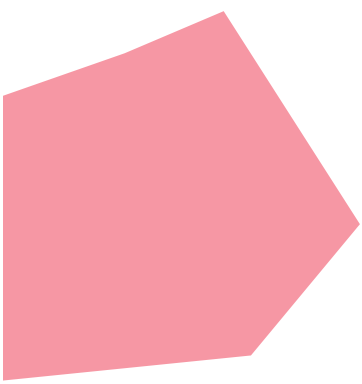
In sum, the findings of this study shed new light on the implementation of activation related sanctions, and show how case managers balance different deservingness criteria against one another in making sanction decisions upon manifestations of work unwillingness. In addition, the findings of this study contribute to important discussions on the merits of discretion and decentralization. In both cases, our findings caution against an overoptimistic view. Whereas we do find that differences at the municipal level are driven by the local socioeconomic context and political preferences, in line with the expected benefits of decentralization, the actual variation between municipalities is very limited. This is surprising in a context where municipalities have large power in the implementation of activation within social assistance, and suggests that other expected benefits of decentralization, such as policy experimentation, may not naturally manifest. As for the expected benefits of discretionary power of social workers, we do find a lot of

variance at the level of the professional, even for exactly the same clients in our experimental set up. Despite a clear choice for professional social workers in Belgium, discretion does lead to a lot of variance. Insofar as this variation is random, it may be an unfortunate but inevitable side-effect of an otherwise beneficial aspect. Yet this study found that a substantial part of this variation can be traced back to the personal attitudes and ideological beliefs of the social workers. This clearly is, although partly natural, a cause of concern. Self-evidently, further research is needed to further substantiate these findings. In future analyses we will focus more closely on the variance at the professional's level, the magnitude of the explaining factors and the organisational elements that in- or decrease the inter professional variation. Further, we are convinced that a repeat of this study in different institutional contexts, for instance with more and less thorough decentralization and where professionals are also the final decision makers, will bring new insights.





# Deservingness in social assistance administrative practice



*Marjolijn De Wilde*

*Customised version of* De Wilde, M (2017). Deservingness in social assistance administrative practice. In van Oorschot, W., Meuleman, B., Roosma, F., Reeskens, T., *The Social Legitimacy of Targeted Welfare*. Cheltenham : Edward Elgar Publishing,

**Abstract**

This chapter focuses on the link between deservingness assessment and the predicted way social assistance clients are dealt with. Social assistance legislation is partly a product of the deservingness opinions of policymakers. The final implementation of this legislation is, however, subject to the discretion of street-level bureaucrats and their organization. Using a vignette experiment conducted among 90 Flemish agencies (municipalities) and 640 professionals, I tested whether eligibility for welfare benefits could be explained by deservingness assessments. The analyses consist of multi-level regression models (first level = vignette, second level = respondent, third level = municipality). Characteristics linked to the attitude, control and reciprocity criteria are best able to predict eligibility. However, the effects are relatively small. As expected, some characteristics linked to the need criterion (adverse life experiences) decrease or cancel out the effects of reciprocity, control or attitude. This study is one of the first to investigate the link between deservingness assessment and the predicted treatment of deserving and non-deserving clients. Reciprocity, control and attitude considerations do determine treatment, but in a fairly limited way. The financial need of the client seems to be the most important determining factor. Although the legislation presents the means test (need) and work willingness (reciprocity, control, attitude) criteria as equal conditions for social assistance eligibility, professionals and municipalities use their discretion to give much more weight to the need criterion.

## **1. Introduction**

The interaction between public policy and public opinion is a recurrent issue in the study of welfare deservingness (Swaan, Manor, Øyen, & Reis, 2000). However, very few studies link opinions about deservingness to the actual treatment of welfare claimants. Several studies provide insights into the treatment of claimants and their characteristics, which determine their eligibility for benefits. These analyses are nevertheless mostly based on official registration data (Bargain e.a., 2012; Carpentier e.a., 2014) that include static information, such as gender, age, nationality background, parenthood, housing situation, level of education and employment history. Important elements including attitude, motivation, mental health and addiction are seldom recorded officially, in spite of the fact that they align with the traditional CARIN deservingness criteria. Furthermore, registration data does not generally allow for analyses at the professional level.

By contrast, studies that focus on such managers – or more broadly, ‘street-level bureaucrats’ – mostly investigate their overall attitudes towards welfare redistribution (Blomberg, Kallio, Kroll, & Niemela, 2015; Blomberg e.a., 2013; Bullock, 2004; Kallio & Kouvo, 2015; Swaan e.a., 2000; Weiss-Gal, Benyamini, Ginzburg, Savaya, & Peled, 2009). The conclusions are that compared with individuals from other occupational backgrounds, social workers think more often that poverty is caused by structural factors, instead of individual factors, are more in favour of the welfare state. Furthermore, their opinions have become stricter over time and opinions differ from country to country. However, the question of whether street-level bureaucrats’ deservingness opinions influence their treatment of welfare claimants remains unanswered.

The aim in this chapter is to study whether traditional CARIN criteria guide the eligibility predictions made by social assistance professionals. To achieve this, I surveyed professionals using experimentally-designed client cases. Applying an experimental set-up allows the focus to be on the professionals’ treatment choices and to include all relevant client characteristics that align with the CARIN model and that are expected to inform professionals about their clients’ welfare deservingness.

## **2. Theory and hypotheses**

The implementation of social policy legislation develops in two interrelated steps: the legislation itself and the assessment of the situations in which this legislation will be applied. These steps are interrelated in the sense that they influence each other. Legislation is partly influenced by existing practice and practice depends, among other

things, on legislation. Deservingness opinions play a role in both steps. First, existing legislation is a product of various factors, including deservingness opinions. Policymakers put in place rules to assure that citizens who are considered deserving receive what they should receive. Second, there is a certain degree of professional agreement about the interpretation of legislation with regard to specific cases or problems. This is a deservingness assessment that is shared between professionals. Third, the assessment of deservingness also differs between professionals, because street-level bureaucrats interpret cases – and the link between them and existing legislation – in their own way. This personal view of what is good for a client and how this relates to national and local legislation is influenced by several factors, including experience (both professional and personal), the education and training of the professional and personal, political and welfare state related opinions (see Chapter 3 and Chapter 5). In the remainder of this section, I first demonstrate the role played by deservingness opinions in Belgian social assistance legislation. I then discuss the concept of discretion, which is the degree of flexibility that street-level bureaucrats have in assessing situations and choosing the way in which clients are treated, followed by linking deservingness opinions to client treatment.

### 2.1. Deservingness opinions in social assistance legislation

Out of all the CARIN criteria<sup>16</sup>, ‘need’ is a key characteristic of social assistance schemes. To understand its role, we can compare social assistance with social security schemes, such as unemployment benefits or sickness allowances. The two types of programmes are open only to those who have no income (from work). With regard to social security benefits, the need criterion goes hand in hand with the reciprocity criterion: sick or unemployed claimants can only receive benefits if they contributed to social security for a long enough time before losing their job or becoming ill. Social assistance, however, does not involve a reciprocity criterion. Instead, it is a means-tested scheme. All citizens of a given country whose income is below a certain level are eligible. In Belgium, the right to social assistance was created in 1974<sup>17</sup> by policymakers concerned about the fact that certain groups of people who had no income were not entitled to benefits such

---

<sup>16</sup> CARIN criteria = deservingness criteria as defined in the book on deservingness in social policy of van Oorschot, Roosma, Meuleman and Reeskens (2017): Control, Attitude, Reciprocity, Identity and Need

<sup>17</sup> The law established in 1974 is *Wet tot instelling van het bestaansminimum* (Act establishing the subsistence level). In 2002 it was replaced by *Wet betreffende het recht op maatschappelijke integratie* (Act on the right to social integration).

as unemployment or sickness allowances. Need, rather than prior contributions, is the key element guiding social assistance (OECD, 2013).

Of course, this does not mean that need is the only criterion incorporated in social assistance legislation. In most European and OECD countries, social assistance regulations include a variant of the work willingness criterion (Ditch, 1999; W. Eichhorst, Kaufmann, Konle-Seidl, & Reinhard, 2008; Marchal & van Mechelen, 2017). The need to demonstrate work willingness seems to cover three of the CARIN criteria. First, reciprocity operates in two directions. A claimant who has contributed to society is considered more deserving of benefits. However, reciprocity also refers to the intention to give something back to the society claimants are appealing to for help. In that sense, work willingness is considered as the intention to contribute in the future, and thus increases deservingness (Moller & Stone, 2013). Second and third, work willingness could also be seen as an attitude and a control requirement. A negative attitude, inappropriate behaviour, difficulties concerning co-operation between the client and the professional, and so on, might be seen as signs of unwillingness to work. Furthermore, a lack of motivation to work could be considered as a cause of poverty. In that regard, claimants who appear unwilling to work can be seen as being in control of their own situation and can therefore be excluded from the benefit scheme.

However, Belgian social assistance law, provides for exemptions from the work willingness criterion. These exemptions are termed 'fairness or health reasons', without further specification (*Wet betreffende het recht op maatschappelijke integratie*, 2002, art. 3, 5e).<sup>18</sup> The inclusion of fairness or health reasons in legislation can be understood as deservingness considerations. Policymakers believe that the work willingness criterion can be waived if there is a particular need. Need is, in this respect, broader than the need for money (means test). It also involves giving the client the benefit of the doubt, because she might be in need of patience (and thus of a benefit), rather than stringent activation. Fairness and health reasons can also be linked to the control criterion. Certain life circumstances make it impossible for clients to work or to look for work, in which case their situation is then considered to be out of their own control.

## 2.2. Discretion in deservingness assessment

Knowing more about the legislation does not mean that we know what happens in reality. The implementation of legislation is – to a certain extent – subject to the

---

<sup>18</sup> The only fairness reason mentioned in the legislation is full-time education. A student in regular full-time education (secondary or tertiary) is exempted from activation (*Wet betreffende het recht op maatschappelijke integratie*, 2002, §2 art.11).

discretionary freedom of professionals, their teams (Evans & Harris, 2004; Hupe & Hill, 2007; Lipsky, 1980; Meyers e.a., 1998) and the organizations they work for (Rice, 2013). In this respect, Evans (2012) introduces an important distinction between *de jura* discretion and *de facto* discretion. *De jura* discretion can be understood as the flexibility that is provided by the legislator, for example through the inclusion of vague regulations in law. *De facto* discretion is the discretion that professionals or organizations actually make use of in practice, regardless of whether it is provided by legislators or managers. It might be that relatively detailed legislation is available to professionals, but they nevertheless attempt to circumvent it or bend the rules because they believe it is to the benefit of their clients and/or society.

With regard to *de jura* discretion, it is fair to say that Belgium's social assistance law is relatively decentralized (Carpentier e.a., 2014). At the local level, each Belgian municipality has a social assistance agency (OCMW) headed by a politically composed, non-professional<sup>19</sup> board. Decisions about whether claimants are eligible for financial assistance are made by this board, based on client files put together by professionals (qualified social workers) who are in contact with the claimants. These files are presented to the board by the professional or by the head of a team of professionals. In practice, the majority of the decisions made by the boards are identical to the advice that professionals include in the client files. Because some of the eligibility criteria – particularly work willingness and health and fairness reasons – are formulated somewhat generally, and with almost no specifications, both municipalities and individual professionals in most OCMWs are guaranteed a certain degree of freedom in implementation.

The Belgian situation is in line with a growing body of literature that states social assistance given to clients should depend on their individual needs. Targeting and adjusting the policy to a client increases the usefulness and meaningfulness of this policy for that client (Tummers and Bekkers, 2014). Social assistance delivery is more than just 'giving money' based on guidelines prescribed in legislation. Financial support is accompanied by administrative, psycho-social support and support aimed at re-integration, all of which involve decisions made by professionals. Literature on professionalism suggests that employees can only be considered professionals if they are assigned at least some decision-making power (Evetts, 2002; Handler, 2003; Hupe & Hill, 2007; Lipsky, 1980). Otherwise, they are simply executors of prescribed rules, much like clerks, and this decreases their willingness to implement such rules (Thomann, van

---

<sup>19</sup> Each board is appointed by the municipality's town council. The president of the board is typically a member of one of the political parties in the majority.

Engen, & Tummers, 2018; Tummers & Bekkers, 2014). Furthermore, the situations in which clients ask for assistance are often so heterogeneous that it is impossible to provide specific legislation for all cases (Thoren, 2008b), making some discretionary authority necessary.

In order to shed light on how deservingness assessments at the professional level might influence actual treatment, I refer to the only related study conducted to date, by Jenny Nybom (2013) in Sweden. Nybom followed almost 400 clients from four municipalities for one year and studied their activation trajectories. Two of the decision logics used in that research, namely neediness logic and work test logic, are similar to the deservingness criteria described above (van Oorschot, 2000). Work test logic, for example, can be compared with the control and attitude criteria: clients who exhibit problematic behaviour are directed towards specific types of activation in order to test their willingness to work (Nybom, 2013). Clients whom professionals believe are not really prepared to work are given a task that could be seen as work. If they accept the task and perform it in an appropriate way, they pass the work willingness test.

Nybom's study finds evidence for both types of logic. She points out that clients who are unwilling to work are sanctioned more often – due to attitude and control criterion – but are also exempted from activation more often (Nybom, 2013). This might indicate that the absence of work motivation is sometimes accepted by professionals, and that other deservingness criteria – such as neediness – happen to be more compelling. This hypothesis could have been checked by combining several characteristics that represent other deservingness logics into interaction terms.

Nybom also describes one other criterium: gender. She notes that activation policies are influenced by labour market phenomena, in which gender roles are deeply imbedded. In Nybom's opinion, it is impossible to post the gender effect to any other criterion (Nybom, 2013).

In conclusion, social assistance legislation incorporates deservingness opinions in several ways. First, social assistance is clearly based on the need principle (see the means testing). Second, the 'work willing' criterion is an appeal for reciprocity, for taking control of resources and for a positive attitude. Third, some claimants are exempted from the work willingness criterion for fairness and health reasons, in recognition of their individual needs and the lack of personal control afforded by particular situations. In the experiment described below in the method section, I operationalize these criteria as specific claimant characteristics. It is not possible to interpret these characteristics directly from the legislation, because some of the relatively vague characteristics require municipalities and professionals to carry out a deservingness assessment. This is the

reason for my particular interest in the predictions of professionals who work with clients in practice, because they are the people who use their discretion in implementing the legislation.

### **3. Data and Methods**

#### **3.1. Factorial survey and research setup**

In this section, I describe the working of a factorial survey in general and my own research in detail. Factorial survey experiments require each respondent to read and rate stories about individuals or situations (vignettes). Each vignette consists of several factors (for example gender) that correspond to certain levels (for example female and male). In a factorial experiment, the vignette population is made up of all possible combinations of all levels over all factors. To avoid excessive complication, a D-efficient sample is taken from this vignette population and presented to the respondents (Atzmüller & Steiner, 2010; Dülmer, 2007; Goos & Jones, 2011). After viewing the vignettes, respondents are typically asked to answer a set of rating questions.

In order to study the link between deservingness opinions and actual treatment in Flanders (the northern region of Belgium), I used a factorial survey carried out in 90 stratified-sampled<sup>20</sup> Flemish municipalities among 640 stratified-sampled<sup>21</sup> employees involved in eligibility decision-making. I asked them to rate nine vignettes concerning hypothetical clients, with regard to the likelihood that these clients would receive a social assistance benefit (with responses on a seven-point scale).

The characteristics of the hypothetical clients included in the vignettes are of particular interest in predicting treatment. I selected the vignette attributes on the basis of existing literature and interviews with professionals, team leaders and academics, keeping the deservingness criteria in mind (Nybom, 2013; Taylor, 2012; van Oorschot, 2000; Wallander, 2009). All the respondents rated nine vignettes that varied across 14 client characteristics (see Table 4-1). The characteristics kept constant in all vignettes were: age (22 years), marital status (single), income (none), debts (none), contact with parents (none) and physical health (healthy<sup>22</sup>).

---

<sup>20</sup> The strata were based on number of inhabitants, relative number of benefit recipients and a socioeconomic index taking into account industrialization, types of building, prosperity, demographic features, and so on (Belfius, 2007).

<sup>21</sup> The strata were based on gender, age and management level

<sup>22</sup> As described in the theory section, health reasons can lead to exemption from activation. Due to variation limits in the experimental design, I chose to operationalize these health reasons as



Since none of the vignette clients had any financial resources and all were above 18 years of age and legally resident in Belgium, strictly speaking they would have been eligible for social assistance benefits in the country. This means that any variation found in the rating of this variable would be due either to the general reluctance of some professionals or municipalities to offer benefits, or to the final condition for receiving these benefits, namely the client's willingness to work.

It is important to note that the respondents were asked not to state what they would personally prefer or do, but to consider all the elements that might influence the decision, such as their own contact with the client, federal and local legislation, team consultation and the decision of the board. Asking for predictions in this way might lead to validity problems. Theoretically, there is no guarantee that such predictions correspond in any way to actual decisions. However, one element in the survey does contribute to validity: the respondents (professionals) were asked about the likelihood that the board would follow the advice provided in the client files. Almost 80 per cent of the respondents indicated that this likelihood was high to very high (6 or 7 on a scale of 1 to 7). This means that even if the respondents were reporting their own advice – as they do on a daily basis in their agencies – rather than the expected final decisions, we can assume that this advice would be relatively close to the actual decisions taken by the board.

### 3.2. Analyses and hypotheses

For the analyses, I randomized the respondent-ID and municipality-ID variables<sup>23</sup>, which resulted in a three-level regression model. Level 1 includes the vignettes (client cases). As all the respondents rated nine client cases, in Level 2, the first level was clustered by respondent (professional). These professionals were then clustered in 90 municipalities (Level 3). The model was run using the software package MLwiN (used via runmlwin in Stata). Subsequently, I included two blocks of variables to test the deservingness perceptions associated with specific client characteristics. First, I entered all the client characteristics (the vignette attributes in Table 4-1) into the model to test the effect of

---

mental health and the health of the client's child, rather than the client's physical health (see Table 4-1).

<sup>23</sup> The multi-level nature of reality (almost every organization employs several professionals and each professional handles several clients) was taken into account in the multi-level model. This chapter presents the average treatment recommended by an average professional in an average organization. Other types of analyses could focus on differences across professionals in a single organization and variation across organizations.

the clients' need and work willingness. Second, I entered interaction terms between several client characteristics to test the effects of fairness and health reasons.

One or more CARIN expectations were linked to each of the vignette variables (see the last column in Table 4-1).<sup>24</sup> One attribute relates to the identity criterion: the background of the client in terms of nationality (Fording, Soss, & Schram, 2011; Terum & Heggen, 2016). I included an interaction term between 'the national background of the client' and 'the language abilities of the client', as one can expect that a nationality effect is different for clients with foreign roots that master the national language and client who do not.

The criterion of need, the central aspect of social assistance legislation (see above), is inherent to some of the attributes that were kept constant, namely the claimant's lack of income or savings. Furthermore, elements such as having children, being homeless, being mentally ill, having an addiction and/or encountering difficult life experiences can be considered as increasing the client's need in the broad sense, as described in the theory section. Poor language skills, addiction problems, lack of intellectual ability and a low level of education might be interpreted as higher need, because clients with these characteristics have fewer opportunities in the labour market and are therefore more reliant on social assistance. Being educated to bachelor's degree level or higher might decrease the need for assistance.

Work willingness – as a violation of reciprocity and attitude, and as a sign of personal responsibility – is operationalized in the vignettes as (a lack of) motivation to work. The vignette attribute 'positive work experience' may be seen as a sign of reciprocity and a good attitude. Furthermore, the loss of a job due to cutbacks – the way in which one type of job loss was operationalized in the vignettes – decreases the client's responsibility for the situation. Because clients with negative activation experiences were modelled in the vignettes as those who had difficulties complying with agreements, and clients with negative work experience were those who had experienced conflict with their superiors, these attributes hinted at the personal responsibility of the clients. Lastly, I include in the vignettes clients who had not been diligent on one or more occasions and who had no real explanation for this behaviour. This can be interpreted as a sign of a bad attitude and failure to take responsibility.

The last section in the Belgian law that is of interest here is the reference to health or fairness reasons for exempting clients from work willingness duty. Therefore, it was

---

<sup>24</sup> The attachment of deservingness opinions to specific client characteristics is less straightforward than it appears in Table 4-1. For the purposes of this paper, I opted for simplicity.

relevant to test – by using interaction terms – the relationship between characteristics linked to work willingness and characteristics that could be seen as fairness and health reasons. These reasons link to increased need in the broad sense as described above (for example having children, mental problems, adverse life experiences and housing problems) and to lack of control over a person’s own situation. The sense of a lack of control was enhanced in several ways in the vignettes. For example, a child was ill with an immune disorder – meaning that the problems could not have been caused by the client’s choices. The symptoms of the illness were unpredictable and no sufficient medical assistance was available. Furthermore, a deceased child (adverse life experience) had been born with heart problems. Therefore, the child’s death could not have been caused by the client. In this second step, I combined multiple characteristics, selecting the variables that had proven to be significant in the first models and constructing interaction terms between the significant but ‘contradictory’ criteria.

Drug-addicted vignette clients could be considered both unwilling to work and as having a fairness reason to receive social assistance, despite not being willing to co-operate. The drug-addicted clients presented in the vignettes used in my study refused help and might – in that sense – have been considered to blame for their own situation. However, some professionals might have considered addiction to be an uncontrollable factor leading to a higher likelihood of eligibility.

A last element that was included in the analysis was gender, as Nybom (2013) predicts women to be treated as more deserving than men.

Table 4-1: Deservingness predictions based on the client descriptions

Attribute	Category1	Category2	Category2	CARIN criteria
Gender	Female	Male		Gender
Nationality background	Client and parents born in Belgium	Client born in Belgium – parents born in Morocco	Syrian refugee – one year in Belgium. Registered	Identity
Command of the national language	Good	Limited	Poor	Need & control
Level of education	Primary education	Secondary education	Bachelor’s degree	
Parenthood	No children	One healthy child aged two years	One sick child aged two years (immune disorder)	Need & control
Housing situation	Stable housing in a subsidized apartment	Lives with a friend – lots of quarrels	Homeless – staying with various friends	Need
Mental health	Good	Depressed – no treatment	Lack of intellectual ability – not diagnosed by a doctor	Need & control
Drug addiction	None	Slightly addicted – refuses treatment	Severely addicted – refuses treatment	Need & control
Adverse life experiences	<i>No information included</i>	Youth with violence and abuse	Lost a baby one year ago	Need & control
Criminal history	<i>No information included</i>	Been in prison for shoplifting	Has been prostituting her/himself	Control & attitude
Aspirations	Wants to work	Wants to study	Is not motivated to work or study	Reciprocity, control & attitude
Employment experience	No experience	Positive experience – made redundant	Negative experience – quarrels with superior	Reciprocity, control & attitude
Activation experience	No experience	One negative experience – not diligent	Several negative experiences – not diligent	Reciprocity, control & attitude
Attitude	Diligent	Missed one appointment	Missed several appointments	Attitude

#### 4. Results

In order to test the validity of the hypotheses concerning the deservingness criteria, in the first step all client attributes were included as fixed effects in Model 1. The client characteristics linked to work willingness influenced the likelihood of receiving benefits significantly (see Table 4-2). Not being motivated, (often) missing appointments, negative work experiences and unwillingness to receive treatment for addiction decreased this likelihood. Some of these effects can also be explained as control effects, for example refusal to seek help for drug addiction and job loss because of conflicts with superiors (negative work experience). However, the effects were quite small: I observed a

maximum decrease of 4.5 per cent%-point in the likelihood of receiving benefits (client not motivated).

The effects referring to increased need are weaker, but nevertheless apparent. Clients with a bachelor's degree were predicted to be less likely to receive benefits. Those with children (either healthy or sick) and who had suffered adverse life experiences were more likely to be considered eligible. Surprisingly, being homeless or in an unstable housing situation resulted in a decrease in the likelihood of eligibility. I hypothesized that housing problems would indicate a need for more support, but the professionals appear to have problematized the uncertainty. In Belgium, clients are eligible in the municipality where they are residing (not necessarily the municipality in which they are registered; in the study, all hypothetical clients were described as residing in the municipality where the respondent worked). Perhaps the professionals thought that clients with uncertain housing situations could move to another municipality at any moment. They seemed to question the usefulness of starting a trajectory with the client, if she might move in the near future. Actually, this doubt goes against the legislation. Yet, it should be stressed that the effect was small (-2%-point).

The identity criterion was visible, but not directly. Only second-generation migrants who spoke good or limited Dutch were considered less eligible. Second-generation migrants with a poor command of the national language and refugees were considered to be as eligible as 'native' Belgians (see Table 4-2). I also found proof of the gender effect. The negative impact on eligibility of being male was – again – very small (-2%-point) , however.

Table 4-2. Multi-level regression results of social assistance benefit deservingness

Likelihood of receiving a benefit		Model 1	Model 2
Attributes	Levels	Coef. (Std. Err)	Coef. (Std. Err)
Intercept		84.18 (1.39)***	85.5 (1.49)***
<b>Client attributes</b>			
Gender	Male (ref = female)	-1.96 (0.47)***	-1.94 (0.47)***
Nationality background	2-generation immigrant (ref = B)	-1.76 (0.82)*	-1.79 (0.82)*
	Refugee (ref = B)	0.28 (0.78)	0.25 (0.78)
Command of national language	Limited (ref = good)	0.05 (0.77)	0.04 (0.77)
	Poor (ref = good)	-1.42 (0.77)	-1.46 (0.77)
Nationality background	2-generation immigrant * limited	0.76 (1.09)	0.77 (1.09)
* command of national language	2-generation immigrant * bad	2.62 (1.09)*	2.62 (1.09)*
	Refugee * limited	0.25 (1.1)	0.2 (1.1)
	Refugee * bad	1.06 (1.09)	1.12 (1.08)
Level of education	Secondary education (ref = primary)	-0.22 (0.41)	-0.23 (0.41)
	Bachelor's degree (ref = primary)	-2.34 (0.52)***	-2.32 (0.52)***
Parenthood	1 healthy child (ref = no)	1.58 (0.44)***	1.56 (0.44)***
	One sick child (ref = no)	2.03 (0.45)***	2.02 (0.45)***
Housing situation	Unstable (ref = stable)	-1.75 (0.44)***	-1.76 (0.44)***
	Homeless (ref = stable)	-2.32 (0.44)***	-2.31 (0.44)***
Mental health	Depressed (ref = healthy)	-0.47 (0.45)	-0.39 (0.77)
	Intellectually disabled (ref = healthy)	-0.98 (0.47)*	-0.06 (0.78)
Drug addiction	Slightly addicted (ref = no)	-0.83 (0.44)	-0.84 (0.44)
Adverse life experiences	Severely addicted (ref = no)	-1.47 (0.45)***	-1.49 (0.44)***
	History of violence (ref = none)	1.99 (0.44)***	0.98 (0.78)
	Loss of a child (ref = none)	1.72 (0.45)***	0.84 (0.76)
Aspirations	Wants to study (ref = wants to work)	-2.0 (0.44)***	-3.1 (0.99)**
	Not motivated (ref = wants to work)	-4.46 (0.44)***	-7.19 (0.99)***
Work experience	Positive experience (ref = none)	0.76 (0.45)	0.75 (0.45)
	Negative experience (ref = none)	-1.22 (0.45)**	-1.22 (0.45)**
Activation experience	1 negative experience (ref = none)	-0.13 (0.45)	-0.12 (0.45)
	Several negative experiences (ref = none)	-0.6 (0.44)	-0.61 (0.44)
Attitude	Missed 1 appointment (ref = diligent)	-1.8 (0.45)***	-1.8 (0.45)***
	Missed several appointments (ref = diligent)	-3.72 (0.44)***	-3.73 (0.44)***
<b>Interactions: client characteristics</b>			
Adverse life experiences	History of violence * wants to study		0.2 (1.08)
* Aspirations	History of violence * not motivated		3.29 (1.08)**
	Loss of a child * wants to study		-0.67 (1.09)
	Loss of a child * not motivated		2.95 (1.09)**
<b>Variations</b>			
	Level 3 (municipality)	31.8 (10.91)*	31.61 (10.87)*
	Level 2 (respondent)	152.89 (10.2)***	152.96 (10.2)***
	Level 1 (vignette – client)	165.92 (3.31)***	165.25 (3.3)***
<b>N</b>	Level 3 (municipality)	90	90
	Level 2 (respondent)	645	645
	Level 1 (vignette – client)	5655	5655

Model 1 = including the client attributes as fixed effects; Model 2 = adding interaction effects between work willingness characteristics and health or fairness reasons

\* p<0.05; \*\* p<0.01; \*\*\* p<0.001.

In the second step (Model 2 in Table 4-2), I combined several client characteristics in interaction terms to test how they interacted. I was primarily interested in the interplay between characteristics that might be interpreted as the work willingness of the client

and characteristics possibly considered as being linked to fairness or health reasons. As I selected only the significant terms in these groups, the following interactions were tested: parenthood and adverse life experiences, each combined with aspirations, attitude and addiction problems. Only one interaction term proved to be significant. Among clients with no adverse life experiences, the effect of a lack of motivation meant a 7.2%-point decrease in the likelihood of receiving benefits. However, as expected, characteristics linked to the need criterion (adverse life experiences) decreased or cancelled out the control or attitude effect. This indicates that a lack of motivation only influenced eligibility negatively if the client had not had any adverse life experiences. This is clearly an example of the relationship between work willingness and fairness reasons.

## **5. Discussion and conclusion**

This chapter focuses on the link between deservingness assessment and the predicted way social assistance clients are dealt with. Social assistance legislation is partly a product of the deservingness opinions of policymakers. The final implementation of this legislation is, however, subject to the discretion of street-level bureaucrats and their organization. Using a vignette experiment conducted among 90 Flemish agencies (municipalities) and 640 professionals, I tested whether eligibility for welfare benefits could be explained by deservingness assessments. The client descriptions in the study were designed in such a way that, first, all the basic information was available, second, the primary condition for eligibility for social assistance – namely need – was fulfilled (none of the ‘clients’ had any money), third, work willingness assessments could be made (the control, attitude and reciprocity criteria) and fourth, several criteria which could be considered fairness reasons for exemption from the obligation to be work willing (need/lack of control) could be used to check for the interrelation between two contradicting deservingness attributes.

The analyses consist of multi-level regression models (first level = vignette, second level = respondent, third level = municipality). Including the client characteristics as fixed effects reveals that the characteristics linked to the attitude, control and reciprocity criteria are best able to predict eligibility. However, the effects are relatively small. This means that although willingness to work is one of the major conditions for securing social assistance benefits in Belgium, professionals and municipalities are also likely to award these benefits to clients who are not motivated. Furthermore, the fact that being late once and/or not being motivated to accept treatment for drug addiction also seem to be interpreted as unwillingness to work, clearly shows that the use of discretion is at play here. These characteristics are not mentioned explicitly in legislation. On the

contrary, the jurisdiction states that not being punctual once may not result in the withdrawal of benefits (Versailles, 2008).

As expected, some characteristics linked to the need criterion (adverse life experiences) decrease or cancel out the effects of reciprocity, control or attitude. It seems that municipalities and/or professionals are better able to accept undesirable behaviour from clients who have troubled histories. In this regard, the findings support the conclusions of Nybom (2013), who found that undesirable behaviour led either to exclusion from social assistance or to exemption from work duties. This study indicates that the latter finding could probably have been explained by other client characteristics that reflected a higher degree of need.

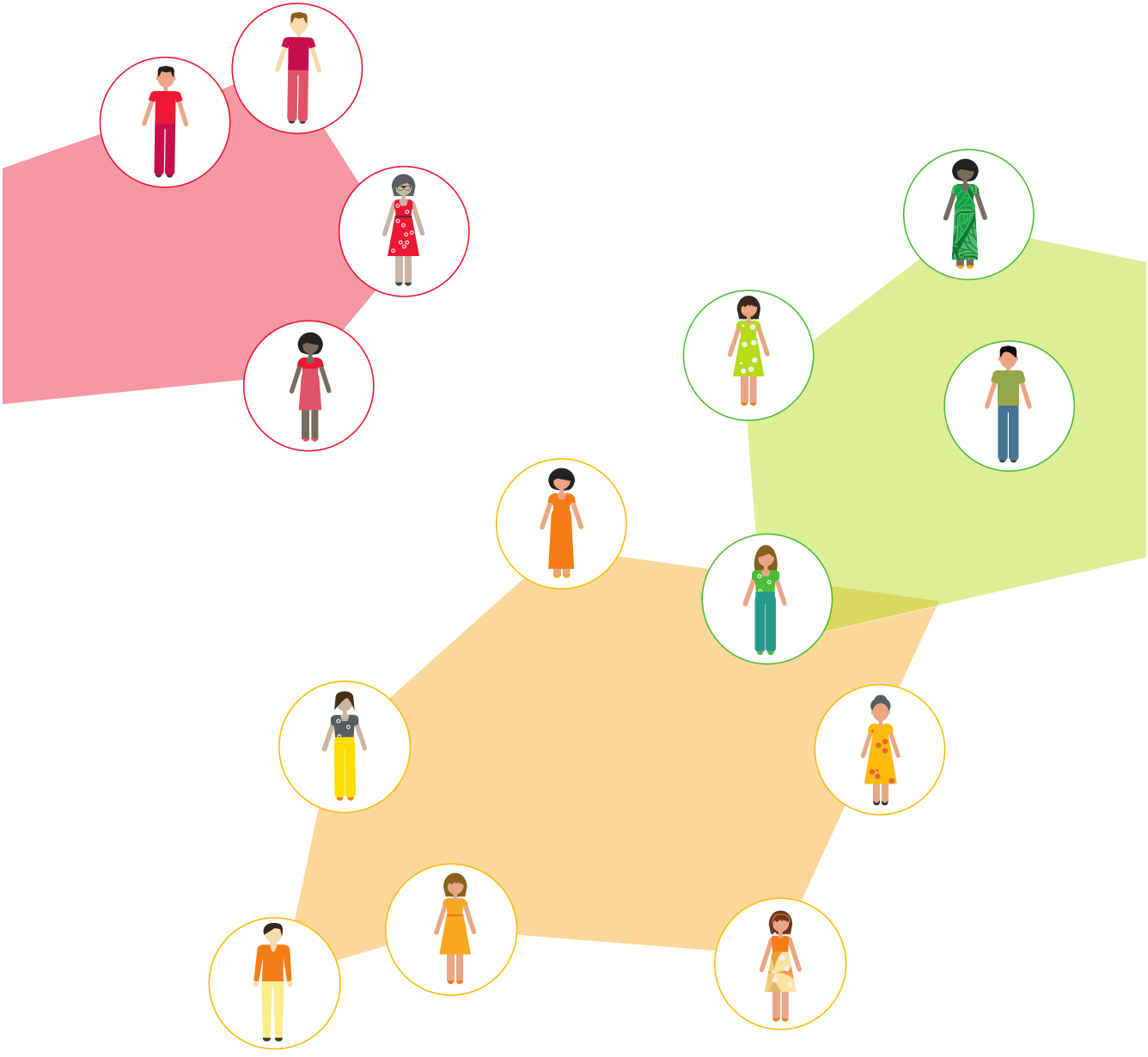
One criterion that has been relatively important in deservingness studies conducted among the general population (van Oorschot e.a., 2017), namely identity, did not prove to be significant directly. Only second-generation migrants with a good command of the national language were predicted to be less eligible. It is difficult to say whether this result is a direct manifestation of the identity criterion. What is certain is that it contradicts the legislation. The same applies to the gender effect and the decrease in eligibility for clients with housing problems. A possible explanation for these results which contradict direct prescriptions in legislation could be that because professionals are constantly assessing difficult situations, they also apply their discretion to situations for which it was not intended.

This study is one of the first to investigate the link between deservingness assessment and the predicted treatment of deserving and non-deserving clients. Reciprocity, control and attitude considerations do determine treatment, but in a fairly limited way. The financial need of the client seems to be the most important determining factor. Although the legislation presents the means test (need) and work willingness (reciprocity, control, attitude) criteria as equal conditions for social assistance eligibility, professionals and municipalities use their discretion to give much more weight to the need criterion. Kroeger (1975) found something similar. While he was testing the hypothesis that more assertive clients received more support, he found that differences were small and that the financial hardship of the client explained most of eligibility variation. Further discretionary freedom is evident in the criteria that were interpreted as work willingness (motivation to work and diligence) and those that were not: negative work and activation experiences and in the criteria that became significant without any basis for this in legislation (gender, nationality, housing situation). The study clearly shows that deservingness assessment plays an important role in the implementation of legislation.



# The use of discretion by social assistance professionals:

THE QUANTITATIVE VERIFICATION



Chapter 5. The use of discretion by social assistance professionals: a quantitative verification

---

*Marjolijn De Wilde*

*Planned submission to Journal of Public Administration Research and Theory*

**Abstract**

Since the seminal work of Lipsky on street-level bureaucrats, the idea that local professionals shape policy has established itself on the research agenda. However, the question of whether the use of discretion by street-level professionals is high or low has not been answered. It is often taken for granted, based on qualitative research, but a quantitative verification is lacking due to the scarcity of data on treatment that can be related to professionals and their organizations.

Here, I aim to address this research gap using data from a richly documented experimental vignette study in which 572 social assistance workers from 76 Belgian agencies were each surveyed concerning the social assistance eligibility and sanction treatment of nine unique client cases. Belgium has federal social assistance legislation that is implemented by professional social workers in municipality-based local agencies. Besides citizen criteria, this legislation provides little detail on eligibility and sanction criteria.

The data, analysed using multi-level techniques, show that eligibility decisions vary moderately among social workers in the same agencies. Inter-professional treatment variation is considerably higher with regard to sanction treatment, however. The disagreement among professionals was highest with regard to client characteristics that deviate from what could be considered 'clear' profiles, for example having children, homelessness, health problems and lack of diligence. Little evidence was found to support the notion that agency-specific arrangements lead to more or less use of discretion. In municipalities in which professionals reported feeling that their recommendations were mostly followed in clients' treatment, more inter-professional variation was found. However, in agencies which had internal guidelines on the interpretation of the work willingness condition, I found more variation.

I conclude that qualitative statements about the room for and use of discretion should more often be substantiated by quantitative evidence.

## 1. Introduction

Since the seminal work of Lipsky (Lipsky, 1980) on street-level bureaucrats, the idea that local professionals shape policy has established itself on the research agenda. Social work, social policy and social administration literature refers to this creative process as discretion. Much debate has centred around the assumed decrease in discretion due to the focus on managerial control, which has followed evolutions in information and communication technology (ICT) (Moller & Stone, 2013). These evolutions make it possible to monitor more closely the extent to which legislation is put into practice (Bovens & Zouridis, 2002; Buffat, 2015; Dunleavy, Margetts, Bastow, & Tinkler, 2006). However, other authors have suggested that discretion is inevitable and highly necessary to translate policies into effective and client-tailored practices (Evans & Harris, 2004; Evetts, 2002; Handler, 2003; Hupe & Hill, 2007; Nothdurfter & Hermans, 2018; Thomann e.a., 2018; Thoren, 2008b; Tummers & Bekkers, 2014) and thus remains stable, even if regulations increasingly attempt to limit it (Evans, 2011). Others state that the use of discretion is in fact increasing due to changes in policy focusses (Heidenreich & Aurich-Beerheide, 2014; Jessen & Tufte, 2014; Künzel, 2012; van Berkel, 2010a, 2010b; Watkins-Hayes, 2009).

Given this debate, it is highly surprising that so little is known about the actual variation among professionals in the implementation of social policies (Stranz e.a., 2016). Discretion is mostly studied qualitatively, based on statements by professionals or clients (Sandfort, 2000; Watkins-Hayes, 2011). Quantitative investigations predominantly rely on the degree of freedom that professionals feel they have when making treatment decisions. This subjective variable is then used as a dependent or independent variable to better understand evolutions in social work and social administration reality (Tummers & Bekkers, 2014). The question of whether actual client treatment variation is high or low among professionals has not been answered, as comparable quantitative data about treatment that can be related to professionals and their organizations are lacking. However, studies of discretion and policy proposals which draw on these studies are only relevant if, besides deep, qualitative knowledge, we have some insight into its actual use.

In this paper, I draw on the results of an innovative and purpose-designed survey of social workers in Belgium to address the question of actual variation in the implementation of social assistance eligibility legislation. Using an experimental vignette study (described below), professionals were surveyed not about their own files, but about hypothetical cases. This method allowed for a large-scale, quantitative investigation of the use of discretion (see 4. Methods).

In the next section, I elaborate on current understanding of the use of discretion, formulate a research question and list hypotheses based on existing empirical research. Section 3 explains why Belgium, with its rather vague social assistance legislation and professional interpretation of social assistance work, is an interesting case for studying the actual use of discretion by professionals. Section 4 introduces the experimental method used, the two dependent variables (social assistance eligibility and sanction treatment), the analysis method (multi-level with residual variance modelling) and the independent variables at the client, professional and agency levels. The results in Section 5 are divided into a description of the actual variation in treatment and the factors that make the variation interpretable. I conclude that inter-professional variation is low with regard to the eligibility variable and high with regard to the sanction variable. Analysing this variation quantitatively enriches the current understanding of factors influencing the use of discretion.

## **2. Literature review**

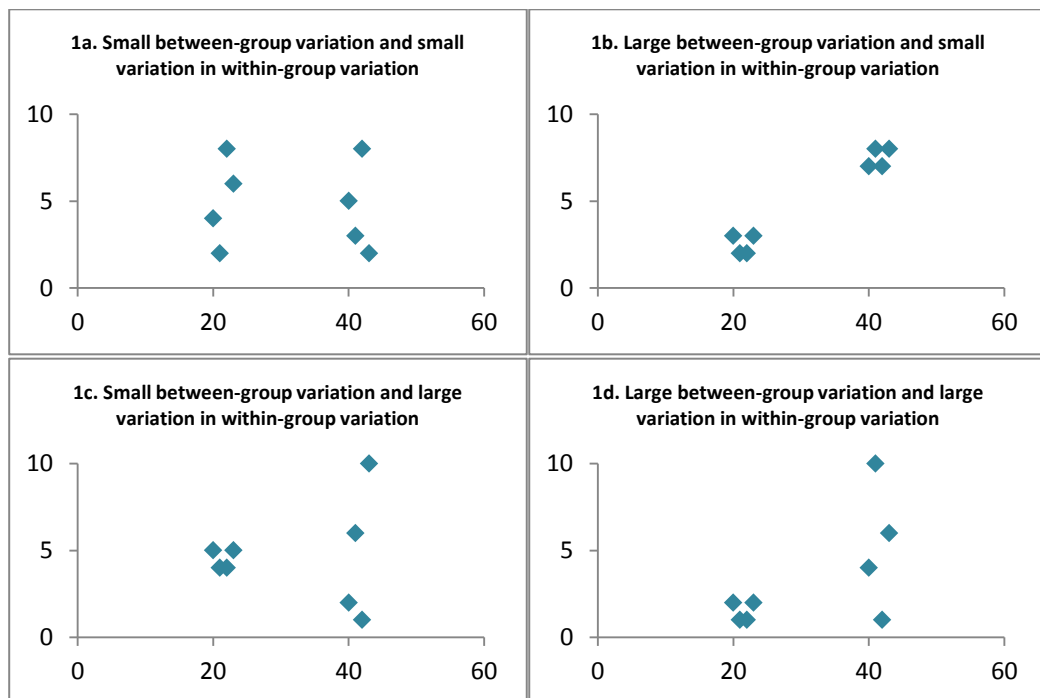
### **2.1. Manifestation of *de facto* discretion**

Evans (2012) draws an important distinction between *de jura* discretion and *de facto* discretion. *De jura* discretion concerns the national, local or organizational rules that govern the room afforded to professionals to be creative in treatment. *De facto* discretion is the degree of discretion they use in practice, which could be in line with *de jura* discretion but could also be lower (professionals follow the guidelines quite uniformly and straightforwardly, even if there is room to interpret them creatively) or higher (professionals interpret existing guidelines in their own way or break these guidelines) (Evans, 2012). For this study, I am particularly interested in the study of *de facto* discretion, both in terms of its magnitude and the influencing factors.

The use of discretion might manifest itself in three ways: first, in professionals with specific characteristics who have specific treatment preferences; second, in treatment variation among professionals with specific characteristics and in specific settings; and third, in factors that increase or decrease the professionals' tendency to deviate from prescribed legislation or from standard treatment. The difference between the first and second manifestations is actually the distinction between between-group and within-group variation. If, for example, the age of a professional has an influence on the use of discretion, this might mean that older professionals choose different types of treatment. Here, differences in the group means are studied (see Figure 5-1b and Figure 5-1d). Another interpretation is that certain age groups exhibit more within-group variation than other age groups. Older professionals might, for example, disagree more than younger professionals in their treatment proposals. Here, the differences in group

variance are the focal point of interest (see Figure 5-1c and Figure 5-1d). Between-group variation was the focus of the study in Chapter 3 in which Sarah Marchal, where was concluded that more senior professionals who were less critical of the functioning of the welfare state tend to sanction their clients less. In this chapter, I focus on the within-group variation among groups of professionals.

Figure 5-1. Hypothetical illustrations of within-group and between-group variation.



Note: dependent variable = treatment X (0-10) for two age groups: 20-25 and 40-45

The third manifestation of discretion is closely related to the other two; however, it not only investigates differences themselves, but also relates this investigation to the question of whether professionals deviate from standard interpretations of legislation or from what is expected of them. This is referred to as ‘bending the rules’ (Kazepov & Barberis, 2013) or to ‘(pro-socially) breaking the rules’ (Kazepov & Barberis, 2013; Morrison, 2006). Although this is an interesting angle, I wish to focus on the second manifestation of discretion, which leads to the following research question: *What is the overall level of inter-professional treatment variation and which characteristics increase or decrease this inter-professional variation?*

## 2.2. Factors influencing inter-professional variation

Quantitative evidence on treatment variation among professionals is scarce. Opinions about its magnitude range from 'decreasing' to 'inevitable' (see introduction). These statements are based primarily on observations in single agencies or on interviews with professionals or clients. To the best of my knowledge, actual large-scale treatment or outcome variables have never been studied with a focus on inter-professional variation. Nevertheless, for the first hypothesis I mainly follow the professionalization argument, which stresses the importance and inevitability of discretion. Thus, an employee can only be considered a professional provided she is assigned at least some decision-making power (Evetts, 2002; Handler, 2003; Hupe & Hill, 2007; Lipsky, 1980). Other authors suggest that professionals will only be willing to implement social policies with regard to complex social problems if they are allowed a certain amount of discretion to do so (Thomann e.a., 2018). Furthermore, the situations in which clients ask for assistance are often so heterogeneous that professionals routinely exercise creativity to answer the question they encounter. The use of discretion is part of their profession by default (Thoren, 2008b). The first hypothesis is then as follows:

*H1: Treatment variation and thus the use of discretion is high among professionals who work in settings with heterogeneous problems and client circumstances.*

The remainder of this description of the hypotheses, based on evidence from the literature, is structured as follows: the influence of legislation, of client characteristics, of professional characteristics and of local characteristics.

According to Kazepov and Barberis (2013)<sup>25</sup>, the type of *de facto* discretion used in an organization is the result of institutional constraints and specific relationship structures. These constraints and structures lead to differences in procedures, rights and duties, which result in different degrees of freedom and resources for different local professionals. Based on this understanding of the causes and results of discretion, Kazepov and Barberis (2013) distinguish between three types of discretion: discretion within standard rule (not a lot of variation), interpretative discretion and contrasting discretion (high variation). In this chapter I mainly focus on interpretative discretion, as it applies to Belgium (see Section 3. The Belgian situation). Institutions that use interpretative discretion are characterized by rules that are rather elusive and by overlap between rules and gaps in legislation. Often, there are also large numbers of

---

<sup>25</sup> This description of the discretion typology relies heavily on Kazepov and Barberis's terminology. I have simply re-ordered the terminology and interpreted certain elements in more detail.

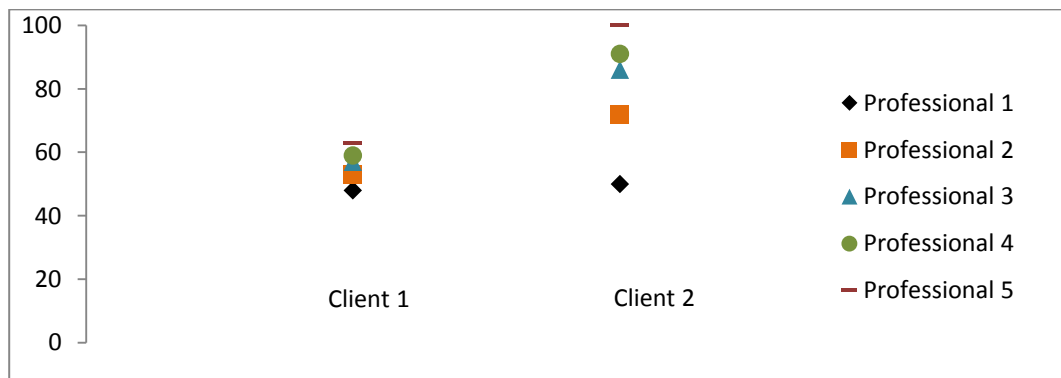
institutional actors with roles that intertwine. As a result, the actual treatment offered to claimants is rather unpredictable and it is unclear who is ultimately accountable for the treatment offered (Kazepov & Barberis, 2013).

*H2: Treatment variation is high in countries with legislation and institutional organization that allows for interpretative discretion.*

The previous hypotheses focus on the extent of variation. The following hypotheses look for factors that have an influence on this extent. The growing importance of activation and reintegration in the labour market has also gone hand in hand with the introduction of behavioural conditions for eligibility (Clasen & Clegg, 2007; Van Kersbergen & Hemerijck, 2012). Successful activation requires evaluations of work willingness, including job-seeking strategies. Conditions like these are difficult both to prescribe in legislation and to assess in practice (Meyers e.a., 1998). Professionals are forced to take a creative role in deciding how to interpret policy goals in specific situations. Therefore, implementation variation among professionals is likely to increase as a result of the introduction of behavioural conditions instead of categorical conditions (e.g. lone parent) or means tests alone (Moller & Stone, 2013).

These effects are visualized in Figure 5-2: five hypothetical professionals vary more with regard to Client 2's treatment than to Client 1's treatment.

Figure 5-2. Example of inter-professional variation related to client characteristics



*H3: Professionals will vary more in their treatment of clients in situations that are described only vaguely in legislation (e.g. work willingness).*

*H4: Treatment variation among professionals will be highest with regard to client characteristics that deviate most from direct prescriptions in legislation (e.g. variation in eligibility decisions concerning a client who is not willing to work).*

Some hypotheses can be formulated concerning the use of discretion related to professionals' characteristics. Creativity in the implementation process is expected to be higher if professionals feel their job is meaningful and if they experience autonomy and opportunities to tailor solutions to the problems they encounter (Morrison, 2006). The opposite might be true if they experience strict follow-up by higher-level managers. Thus, high perceived involvement of superiors might result in a decrease in inter-professional treatment variation. Other characteristics that may be relevant in explaining the use of discretion include job seniority and qualifications. I expect professionals with more seniority to experience more freedom to draw on their own treatment preferences or to deviate from a straightforward interpretation of the legislation. Furthermore, it is believed that the deskilling of professional roles makes professionals more rule-abiding and decreases their sense of their ability to use their discretion (Taylor, 2014). Existing research remains rather speculative, however. For a hypothetical example of this sort of variation in within-group variation, see Figure 5-1.c-d.

*H5: Older and more experienced professionals will display more inter-professional variation than younger professionals.*

*H6: Professionals with a higher educational level will display more inter-professional treatment variation than professionals with a lower educational level.*

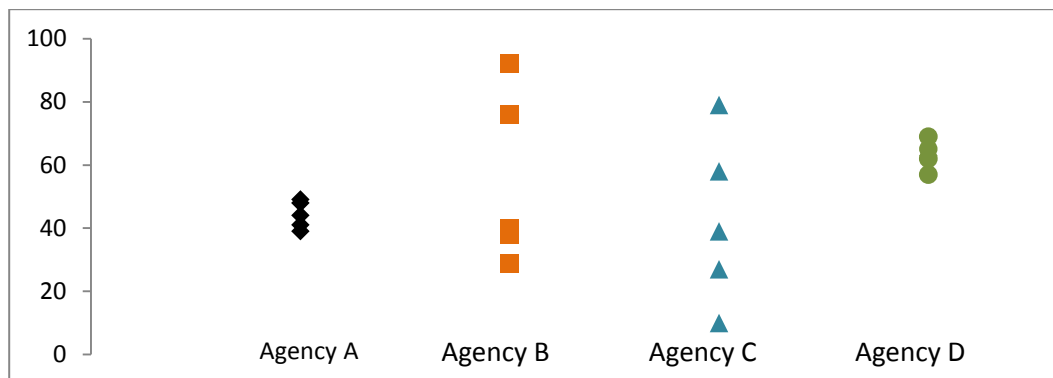
*H7: Among professionals who experience a lot of autonomy, more inter-professional treatment variation will be observed than among professionals who do not experience autonomy.*

With regard to characteristics at the agency level I expect that, in line with the previous consideration in this section, in agencies where the majority of professionals experience a large degree of autonomy, I expect to find more variation in treatment. The factor that is described most often in literature as influencing the use of discretion is a management style that aims to limiting variation. This ranges from traditional managerial techniques, such as creating standard procedures and internal guidelines, to performance management, through specific goal-setting and analysis by client outcomes per professional (Moller & Stone, 2013; Sandfort, 2000; Wastell, White, Broadhurst, Peckover, & Pithouse, 2010), to digitalization of most of the administrative processes (Dunleavy e.a., 2006). Other predictions concern the variety of attitudes in an organization. As attitudes and values influence treatment choices (De Wilde & Marchal, 2018), organizations may attempt to foster similar values among their own professionals (Keiser, 2010). The expectation is then that in agencies with less variety in attitudes, less variation in treatment will be found.



This phenomenon is illustrated in Figure 5-3, which shows the mean treatments of five hypothetical professionals in four agencies. Agencies A and D display the least inter-professional variation. If these agencies made the most use of performance management techniques, these data would confirm Hypothesis 10 (see below).

Figure 5-3. Hypothetical example of inter-professional variation related to agency characteristics



*H8: In agencies where the majority of professionals experience autonomy, inter-professional treatment variation will be higher than in agencies where autonomy is generally low.*

*H9: Managerial techniques focused on minimalizing internal variation will decrease inter-professional variation.*

*H10: In agencies with less variety in values and attitudes, less treatment variation will be found.*

### 3. The Belgian situation

We tested these hypotheses using a purpose-designed survey conducted in Belgium. In this section, I first discuss Belgian legislation concerning social assistance eligibility and sanctioning. Second, I explain the organization and responsibilities of welfare agencies and, third, the responsibilities of the professionals.

Social assistance in Belgium is regulated by federal law and implemented by social workers in local public welfare agencies. In short, federal law guarantees financial assistance, access to a project of social integration or both to all adult citizens who have Belgian nationality or legal residence status, are willing to work and – with regard to financial help – have an income that is below a certain level (Law ‘Recht op Maatschappelijke Integratie’ (‘Right to Societal Integration’), 2002, henceforth RMI law).

The RMI law does not detail how social workers in local agencies (see below) should interpret a client's willingness to work. The RMI law also refers to reasons for exemption from work willingness as 'health reasons' or 'fairness reasons', but no clearly described categories are included. The only detailed condition involves enrolment in full-time education or training that will improve the client's chances on the labour market. Documents accompanying the RMI law do include some examples. Health reasons might be invoked in the case of a young drug-addicted person receiving medical care or a pregnant woman exempted from doing physical labour. A possible 'fairness' reason might involve a single mother with several (perhaps disabled) children and transport difficulties (Vande Lanotte, 2002a). Sanctions can be given to beneficiaries for two reasons: benefit fraud and non-compliance with the agreements made. Sanctions range from temporary (1 to 3 months) suspensions of the benefit payment to definitive suspensions (very rarely). Clearly, the RMI law allows important leeway to local regulations and preferences and to the professional opinions and assessments of social workers.

At the local level, each Belgian municipality has a public social welfare agency headed by a politically composed non-professional board. These agencies are, among other things, responsible for welfare decision-making and actual payment or withdrawal of welfare benefits. They also organize the activation of welfare clients. Activation can be outsourced, but the main responsibility remains with the public social welfare agency. Decisions about whether claimants are eligible for financial assistance or should be withdrawn are made by this board – within the framework of the law – based on client files put together by the social workers who are in contact with the claimants. Even though the federal law stipulates a number of fixed elements in the social investigation, municipalities still have considerable freedom in how exactly to organize the process of producing these files (which are essentially social workers' recommendations on individual cases to the decision board) and the files themselves. The federal level reimburses the municipalities for 60% to 100% of the social assistance benefits distributed, depending on the size of the municipality and the type of client.

The professional's tasks may differ between municipalities. Nonetheless, it is always a qualified social worker (with at least a Bachelor's degree in social work) who prepares the file (and recommendations) for the board. In the preparation of this file, the social worker can use her own discretionary judgement to highlight certain options to move forward with a specific client and to assess the eligibility conditions, including willingness to work or possible 'health or fairness' reasons.

This short introduction to the organization of social assistance in Belgium should make it clear that several actors are responsible for implementing social assistance legislation. First, there is the social worker in contact with the client, preparing the client files. This social worker mostly interacts with colleagues and management figures before finalizing the file. Second, there is the agency board at the decision-making level. Third, the federal level oversees operations. Each municipality is audited by the federal administration on a regular basis. In sum, the inclusion of rather vague behavioural conditions in the RMI law combined with the country's multi-layered responsibilities make Belgium an interesting case through which to test the theoretical assumption of Kazepov and Barberis (2013) concerning *interpretative discretion* (see Hypothesis 2).

## 4. Methods

### 4.1 Data

In order to check the treatment variation among professionals, I needed treatment outcome data and indications as to which clients were treated by which professionals. Such clustered outcome data (client treatment in professionals) are not readily available. Administrative data do not usually include information on the professional. Investigations based on files collected by the agencies would be a good alternative, but very time-consuming. The use of a survey with hypothetical client descriptions provided the opportunity to collect large-scale data on treatment along with information on the professional assessing the hypothetical case. In this respect, I carried out a factorial survey in Belgian<sup>26</sup> municipalities among social workers involved in eligibility decision-making.

Factorial survey experiments require each respondent to read stories (vignettes) about individuals or situations before rating these individuals or situations. Every vignette consists of several factors (e.g. gender), each with their own categories (e.g. male and female). Social assistance professionals were asked to rate experimentally varied vignettes on hypothetical social assistance clients according to, first, the likelihood that they would receive a benefit in the municipality of the respondent (professional) and, second, the likelihood that they would keep this benefit if they refused a job or activation offer. Each respondent rated nine vignettes that varied across 14 client characteristics. The vignette attributes were selected on the basis of the existing literature and interviews with professionals, team leaders and academics (Taylor, 2012;

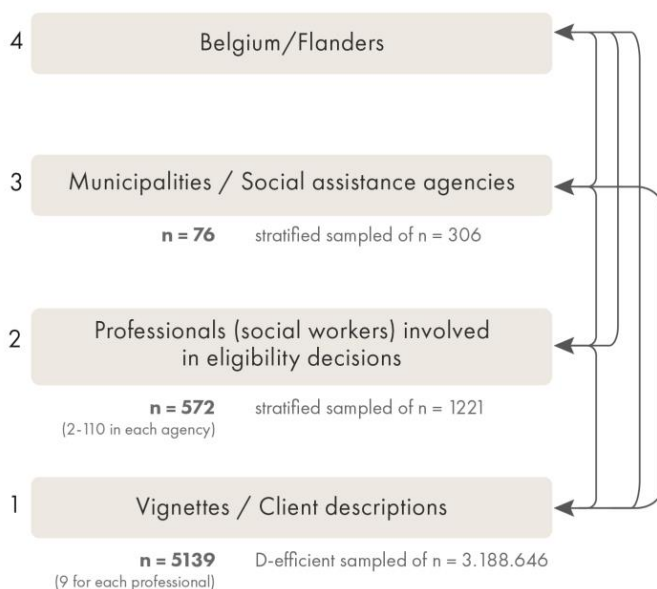
---

<sup>26</sup> All municipalities were located in the Dutch-speaking region of Belgium (Flanders). In the text, I describe the Belgian situation, as all policy levers were situated at the Belgian or local level (not the regional level) at the time of the survey.

Wallander, 2009). The characteristics kept constant in all vignettes were age (22 years), marital status (single), income and debts (none), contact with parents (none) and physical health (healthy). The 14 client characteristics that varied included socio-demographic characteristics and characteristics that would impact on professionals' assessment of individual situations according to the deservingness literature, i.e. characteristics indicating work willingness and possible exemption criteria.

These experimentally varied vignettes were D-efficient<sup>27</sup> distributed over the respondents. The respondents themselves were selected in a two-stage stratified probability sample (first agencies/municipalities, then respondents), with reselection. I invited 105 municipalities, 15 of which declined to participate mainly due to time constraints; one municipality did not provide enough responses to be included in the research. In the second stage, 839 social workers were invited to fill out the online survey, 582 of whom returned a completed questionnaire. Specific to this investigation, I excluded the 14 agencies that were participating with only one social worker. Analyses were ultimately performed on 76 agencies and 573 professionals (see Figure 5-4).

Figure 5-4. Study design and sample



<sup>27</sup> The D-efficient sample consisted of an orthogonal design in which, for each factor (e.g. gender), the categories (e.g. male/female) occur as equally often as possible (Atzmüller & Steiner, 2010; Auspurg & Hinz, 2014; De Wilde & Goos, 2017; Dülmer, 2007, 2016).

The respondents constituted a stratified sample (gender, age and task) based on lists of all professionals involved in eligibility assessment in each agency. The survey ran from March to November 2015. The resulting survey is particularly innovative as it includes not only the respondent's assessments of vignettes, but also information on these respondents and the organizations they work in. This provides us with three different levels of analysis: the experimentally varied client characteristics, the characteristics of the social workers themselves, and those of the local social assistance agencies and municipalities they operate in.

Self-evidently, the vignette method is not without flaws. Questions have been raised about the external validity of the method<sup>28</sup> (Auspurg & Hinz, 2014; Taylor, 2006; Wallander, 2009). This criticism applies equally to the factorial survey, with the added risk that respondents' answers may lack realism as they were asked to predict the final decision of the political, non-professional board on each specific case rather than their own preferences. While professionals in Belgium do have an impact on this final decision through the case file and case recommendations they formulate (see above), ultimately this decision is not in their hands, which increases the risk of response error. However, we should not overstate this risk. Respondents indicated that they expected the board to follow their legally required advice on 80% of all files. This means that even if the results in the survey are based on what the professional would advise (rather than on a valid prediction of reality), this would be close to any practical decision in reality.

## 4.2 Analyses

### 4.2.1 The models

As outlined in the literature section, in this chapter I am particularly interested in inter-professional variation in social assistance treatment. The dependent variables are:

- The likelihood, as predicted by the respondents, that the hypothetical clients would receive a benefit. The variable was rated on a 7-point scale (1 = very unlikely – 7 = very likely) and was – for the sake of clarity – recalculated to percentages in most of the analyses used in this chapter.
- The likelihood, as predicted by the respondents, that the hypothetical clients would keep their benefit, even if they refused one or more job offers. This variable is constructed using three other variables. Each of these variables measured sanction likelihood with regard to one type of job offer: a job that started early in the morning, a job that lasted only one day and a job that did

---

<sup>28</sup> Its internal validity is high since the method operates on an experimental basis (Taylor, 2006; Wallander, 2012b), which ensures that no systematic errors occur.

not match the client's qualifications. I brought these three variables together in one variable and used the type of job offer refused as an explaining variable in the analyses.

A first exploratory exercise consisted of simple frequency tables. These give a rough idea of the treatment variation. However, to study the size of the variation at the professional level, I needed to be able to distinguish the variation at this level from the variation due to other levels. In this project there were three hierarchical levels, namely the client (vignette), the professional (respondent) and the agency (municipality). To distinguish between these levels, a three-level model was estimated, by randomizing the vignette, respondent and municipality IDs. Thus, I first estimated a multi-level 0-model without explanatory variables, in order to determine the unexplained variance at the respondent/professional level (Model 0). The variance partition coefficient (= unexplained variance at a certain level divided by the total unexplained variance) gives a good idea of the importance of a certain level in explaining treatment variation (Hypotheses 1 and 2).

In the next – purely preparatory – model, fixed explanatory variables at each of the three levels were added (Model 1). The final model was the result of a stepwise exclusion of variables until an optimal model fit was obtained (see Chapter 3 for the procedure). The parameters of the final model are shown in Appendix 9. As differences in professionals' mean preferences (fixed effects) are not the focus of this chapter, the results will not be discussed in the results section. However, before being able to model the residual variance (Models 2, 3 and 4), I needed to be sure that the remaining variance was purely residual variance and could not be explained by differences in mean treatment preferences (between-group variance) (Kuppens & Yzerbyt, 2014).

Traditional random intercept models assume homoscedasticity at all levels. However, based on the literature, I expected that professionals with certain characteristics or in certain agencies would differ more than (in) others and that professionals would disagree more with regard to certain types of clients than to others. This assumption was tested by modelling the second level residual variance. As Kuppens and Yzerbyt (2014) state: "In the same way that the score on the dependent variable can be predicted by a range of variables, so too can the residual variance of the model be a function of predictor variables". Client, social worker and agency characteristics were then included as random effects at the second level. This means that inter-professional variation was not considered random (homoscedasticity), but was expected to be dependent on other variables (heteroscedasticity).

In the case of Model 2, the randomized variables were client variables, with the expectation that professionals would differ more in treatment predictions with regard to certain clients than to others (Model 2) (see Hypotheses 3 and 4 and Figure 5-2). All possible client characteristics were added one by one. These models were used as exploratory models. The variables were added one by one to Model 1 and re-excluded. The significant random slopes in one model were not retained in the next model, as this would complicate the model estimation. The model fit of each Model 2 was compared to Model 1 fit, to test whether the model fit improved by modelling the residual variance. If it did, the Level 2 residual variance for each of the client characteristics was calculated (see below for formula).

Models 3 and 4 modelled the residual variance further, by randomizing Level 2 and Level 3 variables step by step. Based on the literature, I expected that professionals with certain characteristics or in certain agencies would differ more than (in) others. This assumption was tested by allowing for heteroscedasticity. Level 2 and level 3 variables were randomized at the second level of Model 1. First, I included professional characteristics as random slopes one by one, hypothesizing that professionals with specific characteristics would display more or less treatment variation than other groups (Model 3 – Hypotheses 5-7 - Figure 5-1-c-d). In Model 4, organization characteristics were randomized, as according to the literature, professionals in organizations or regions with certain characteristics are thought to differ more in their treatment than in other organizations/regions (Hypotheses 8-10 - Figure 5-3).

The parameter of interest, namely the residual variance at the second level after fitting Model 2, 3 and 4, was calculated using the following formula:

$$\text{residual variance at the second level for } x_{ijk} = u_{0jk} + 2*\text{cov}(\alpha/x_{2ijk}) * x_{2ijk} + 2*\text{cov}(\alpha/x_{3ijk}) * x_{3ijk} + u_{1jk}^2 * x_{2ijk} + u_{2jk}^2 * x_{3ijk},$$

with  $x_{ijk}$  being a client characteristic (e.g. parenthood) with three categories, 1 (no child), 2 (healthy child), 3 (sick child).  $x_{2ijk}$  and  $x_{3ijk}$  are the dummies for these two categories. Included in the calculation are: the general second two residual variance ( $u_{0jk}$ ), the covariance between the intercept and each of the client characteristic categories  $x_{2ijk}$  (healthy child) and  $x_{3ijk}$  (sick child) and the specific variance components for the two client categories  $u_{1jk}^2$  and  $u_{2jk}^2$

Table 5-1 gives an overview of the formulas for the models, where i indexes individual clients, j the social workers, and k the municipal welfare agencies. y is the likelihood of receiving benefits (as perceived by the respondents).  $x_{ijk}$  are the n individual (experimentally varied) characteristics of the hypothetical clients,  $x_{jk}$  are the m

characteristics of the respondents (the social workers), and  $x_k$  are the p characteristics of the municipality.  $z_{0k}$ ,  $u_{0jk}$  and  $e_{0ijk}$  are the error terms at each level. I estimated the coefficients of three different independent variables x ( $\beta$  at the individual level,  $\gamma$  at the social workers' level and  $\delta$  at the municipality level). I also randomized (random slopes) client, professional and agency characteristics at the second level.  $u_{1jk} - u_{njk}$  are the n specific variance parts attached to the slopes  $x_{ijk}$ ,  $x_{jk}$  and  $x_k$ .

The parameter of interest, namely the residual variance at the second level after fitting Model 2, 3 and 4, was calculated using the following formula:

$$\text{residual variance at the second level for } x_{ijk} = u_{0jk} + 2*\text{cov}(\alpha/x_{2ijk}) * x_{2ijk} + 2*\text{cov}(\alpha/x_{3ijk}) * x_{3ijk} + u_{1jk}^2 * x_{2ijk} + u_{2jk}^2 * x_{3ijk},$$

with  $x_{ijk}$  being a client characteristic (e.g. parenthood) with three categories, 1 (no child), 2 (healthy child), 3 (sick child).  $x_{2ijk}$  and  $x_{3ijk}$  are the dummies for these two categories. Included in the calculation are: the general second two residual variance ( $u_{0jk}$ ), the covariance between the intercept and each of the client characteristic categories  $x_{2ijk}$  (healthy child) and  $x_{3ijk}$  (sick child) and the specific variance components for the two client categories  $u_{1jk}^2$  and  $u_{2jk}^2$



Table 5-1. Overview of models used in the analyses

Code	Research question	Sort of model	Model
Preparation	What is the variation?	Frequency table	
M0	Is the variance at the professional level extensive?	Multi-level null model (client – social worker – municipality) with variance partition coefficient for social worker level	$y = \alpha + z_{0k} + \mathbf{u}_{0jk} + e_{0ijk}$
M1	Modelling the best fitted fixed effects model	Client, social worker and agency/municipality characteristics as fixed effects	$y = \alpha + \beta_1 x_{1ijk} + \dots + \beta_n x_{nijk} + \gamma_1 x_{1jk} + \dots + \gamma_m x_{mjk} + \delta_1 x_{1k} + \dots + \delta_p x_{pk} + z_k + u_{jk} + e_{ijk}$
M2	Is the residual variance dependent on client characteristics?	Client characteristics as random effects at the second level	$y = \alpha + \beta_1 x_{1ijk} + \dots + \beta_n x_{nijk} + \gamma_1 x_{1jk} + \dots + \gamma_m x_{mjk} + \delta_1 x_{1k} + \dots + \delta_p x_{pk} + z_{0k} + \mathbf{u}_{0jk} + \mathbf{u}_{1jk} x_{1jk} + \dots + \mathbf{u}_{nj} x_{njk} + e_{0ijk}$
M3	Is the residual variance dependent on the professional's characteristics?	Professional characteristics as random effects at the second level	$y = \alpha + \beta_1 x_{1ijk} + \dots + \beta_n x_{nijk} + \gamma_{-1} x_{-1jk} + \dots + \gamma_{-m} x_{-mjk} + \delta_1 x_{1k} + \dots + \delta_p x_{pk} + z_k + u_{0jk} + \mathbf{u}_{1jk} x_{1jk} + \dots + \mathbf{u}_{nj} x_{njk} + e_{0ijk}$
M4	Is the residual variance dependent on agency characteristics?	Agency characteristics as random effects at the second level	$y = \alpha + \beta_1 x_{1ijk} + \dots + \beta_n x_{nijk} + \gamma_1 x_{1jk} + \dots + \gamma_m x_{mjk} + \delta_1 x_{1k} + \dots + \delta_p x_{pk} + z_k + u_{0jk} + \mathbf{u}_{1jk} x_{1k} + \dots + \mathbf{u}_{nj} x_{nk} + e_{0ijk}$

#### 4.2.2 The independent variables

In Model 2, all client characteristics were randomized at the second level. The two middle columns of Table 5-2 present the expected effects on the unexplained variance at the second level, based on Hypotheses 3 and 4. Overall, I first expected client characteristics to have an effect on the unexplained variance if they were related to the vague criteria ‘work willingness’ and ‘fairness or health reasons to be exempted from the work willingness condition’ (H3). Second, I hypothesized that the characteristics related to fairness reasons would decrease the variance on the eligibility variable and would increase the variance on the sanction variable. For the characteristics related to work willingness, the opposite reasoning applied. Signs of work unwillingness would increase the inter-professional variance on the eligibility variable and decrease this variance on the sanction variable (H4). Fairness reasons were expected to contribute to the main deservingness criterion with regard to social assistance eligibility, namely ‘need’, and thus minimize disagreement among professionals as to whether a client

should be considered eligible. Characteristics which suggested a sense of unwillingness to work, however, would make professionals doubt eligibility and this doubt was expected to manifest itself as more variation in assessment. The addiction characteristic lies somewhere between a fairness reason and a sign of unwillingness to work, as the addicted client was not willing to look for help to resolve the addiction. However, with regard to sanctioning, signs of work unwillingness contribute to the legitimacy of sanctioning and thus are expected to decrease disagreement among colleagues. Fairness reasons, in contrast, need assessment by professionals to decide whether they outweigh the signs of work unwillingness (e.g. job refusals). Thus, I expect an increase in professional disagreement with regard to these items.

Table 5-2. Overview of randomized client variables in Model 2 with expected effects

Effect on the second level residual variance	Likelihood of receiving social assistance	Likelihood of keeping social assistance after job refusal(s)	Tested hypotheses
Gender ( <b>woman</b> /man)	=	=	
Nationality background ( <b>Belgian</b> /Second generation migrant/Recognized refugee)	=	=	
Command of the national language ( <b>Good</b> /limited/poor)	=	=	
Level of education ( <b>Primary</b> /Secondary/Bachelor's degree)	=	=	
Frequency of job refusal ( <b>1</b> /2/3)		↓	
Motivation ( <b>willing to work</b> /wants to study/not motivated to work or study)	↑	↓	H3 & 4
Diligence ( <b>yes</b> /missed one appointment/missed several appointments)	↑	↓	H3 & 4
Work experience ( <b>none</b> /positive/negative)	↑	↓	H3 & 4
Activation experience ( <b>none</b> /one negative/two negative)	↑	↓	H3 & 4
Parenthood ( <b>no</b> /healthy child/sick child)	↓	↑	H3 & 4
Housing situation ( <b>stable</b> /unstable/homeless)	=	↑	H3 & 4
Mental health ( <b>healthy</b> /undiagnosed depression/undiagnosed intellectual disability)	↓	↑	H3 & 4
Addiction ( <b>none</b> /slight/severe)	?	↑	H3 & 4
Adverse life experiences ( <b>none</b> /child loss/youth with poverty and abuse)	↓	↑	H3 & 4

Note: in bold = the reference category

In Model 3, some professional characteristics were included in order to model the residual variance (see Table 5-3). This model tested Hypotheses 5 to 7. The age and seniority of the respondents were added as a continuous variable in terms of years of experience (H5). As all social workers in a social assistance agency in Belgium need to have a Bachelor's degree in social work, I was only able to differentiate between

professionals with a Bachelor’s degree and professionals with both a Bachelor’s and a Master’s degree. In fact, this is not a good operationalization of Hypothesis 6, which differentiates between employees with high and low educational levels. Hypothesis 7 predicted that the autonomy experienced by professionals would result in differences in the use of discretion. There was no variable in the study that directly operationalized autonomy. However, *Eligibility\_professional*, *Activation\_professional*, *Eligibility\_board*, *Activation\_board* and *Advice\_professional* were used as proxies. *Advice\_professional* reflects the extent to which professionals feel that their advice on eligibility is followed by the board in decision-making (self-rated on a 1-7 scale). *Eligibility\_professional* and *Activation\_professional* reflect the extent to which professionals believe that they themselves are involved in the actual decision concerning eligibility and activation (1-7 scale). This variable captures whether professionals see themselves and their colleagues as important actors in eligibility and activation processes. Lastly, *Eligibility\_board* and *Activation\_board* reflect the involvement of the board in both eligibility and activation decisions. I used these variables as proxies, as, if *Eligibility\_professional*, *Activation\_professional* and *Advice\_professional* score highly, it means the board trusts the assessment of the professional, which is likely to result in higher levels of experienced autonomy. In contrast, feeling that the board is heavily involved in the decision-making process might be a sign of control and decrease in the use of discretion. However, it is clear that these variables are only proxies, as other mechanisms may explain possible effects.

Table 5-3. Overview of randomized professional variables in Model 3 and expected effects

	Mean/percentage of non-reference category	Effect on the second level residual variance	Hypothesis tested
Age	38.21	↑	H5
Qualification (Bachelor’s degree/Master’s)	11%	↑	H6
Seniority	11.75	↑	H7
<i>Advice_professional</i> (1-7)	5.81	↑	H7
<i>Eligibility_professional</i> (1-7)	4.6	↑	H7
<i>Activation_professional</i> (1-7)	5.42	↑	H7
<i>Eligibility_board</i> (1-7)	6.3	↓	H7
<i>Activation_board</i> (1-7)	5.05	↓	H7

Finally, in Model 4, several agency and municipality characteristics were randomized at the second level. First, I randomized control variables such as the size of the municipality and the relative percentage of social assistance clients.

Hypothesis 8, namely the expectation that the use of discretion would be higher in agencies where most professionals experience autonomy, was operationalized as the

municipality mean response on the five variables outlined in the previous section (*Eligibility\_professional\_municipality*, *Activation\_professional\_municipality*, *Eligibility\_board\_municipality*, *Activation\_board\_municipality* and *Advice\_professional\_municipality*). These municipality means were calculated based on the mean of the respondents in the sample. As described in Section 4.1.Data this sample was as representative as possible.

Although three of the above variables, namely *Eligibility\_board\_municipality*, *Activation\_board\_municipality* and *Advice\_professional\_municipality*, can also be considered variables that test the effect of managerial techniques, I included some extra variables to test Hypothesis 9 more thoroughly. When agencies were asked – in a post-questionnaire in 2018 – whether they thought that their social workers knew how to implement the work willingness condition, they most often referred to team consultation as the place where homogeneity was strived for. Therefore, I checked whether more team involvement led to less use of discretion. The variable used was the importance of the team in the eligibility decision-making process (as indicated by the respondents) (*Eligibility\_team\_municipality* - (1-7 scale)). I also used a variable from the post-questionnaire that asked about internally issued guidelines concerning the work willingness requirement. The expectation was that agencies with internal guidelines would show less variation among professionals.

Finally, I included the variation in professionals' opinions on the welfare state as an explaining variable in the random part of the model. I included the agency-specific standard deviation of two of the latent variables (described extensively in Chapter 2) which reflect professionals' views on the consequences of the welfare state (*Welfare State Criticism*, higher values reflect negative opinions about the consequences) and on the need for more monitoring of welfare state claimants (*Welfare State Strictness*) (for an overview of the construction of these latent variables, see Chapter 2 and Appendix 5).

Table 5-4. Overview of randomized municipality variables in Model 4 and expected effects

	Mean/percentage of non-reference category	Effect on second level unexplained variance	Tested hypothesis
Size of municipality	30442.7	?	Control variable
Relative percentage of social assistance clients	0.57	?	Control variable
Advice_professional_municipality	5.81	↑	H8 and 9
Eligibility_professional_municipality	4.82	↑	H8
Activation_professional_municipality	5.55	↑	H8
Eligibility_board_municipality	6.26	↓	H8 and 9
Activation_board_municipality	5.1	↓	H8 and 9
Eligibility_team_municipality	4.17	↓	H9
Guideline_work_willingness (No/yes)	13.33%	↓	H9
Inter-agency variation of welfare state criticism	0.27	↑	H10
Inter-agency variation of welfare state strictness	0.16	↑	H10

## 5. Results

### 5.1 Studying the variation

First of all, I checked for the variation found in the data before performing any regressions. The scale of the dependent variable ranged from 1 (very unlikely the client would receive or keep a benefit) to 7 (highly likely that the client would receive or keep a benefit). As shown in Table 5-5, the distribution is right-skewed if the client has not refused any job offers and it is almost evenly distributed over all scores if the client has refused one or more job offers. Most respondents indicated that the likelihood of receiving a benefit is high (5 or 6) to very high (6 or 7) at the beginning of the activation pathway (no job refusals). This means that the variation in treatment proposals for this situation is rather limited. This is not surprising, as all hypothetical clients had no money and no one to rely on and possessed the citizen characteristics (age, nationality) necessary for eligibility. The variation increases if the client refuses job offers.

Table 5-5. Frequency table of likelihood of receiving or keeping social assistance benefits

	<b>Likelihood of receiving benefit</b>	<b>of receiving</b>	<b>Likelihood of keeping benefit after refusal(s)</b>	<b>of keeping benefit after</b>
1 – very unlikely	16	0.3%	1826	13.0%
2	74	1.4%	2558	18.3%
3	120	2.3%	2504	17.9%
4	456	8.9%	2516	18.0%
5	1016	19.8%	2038	14.6%
6	2012	39.2%	1695	12.1%
7 – very likely	1445	28.1%	864	6.2%
Total	5139	100%	14001	100%

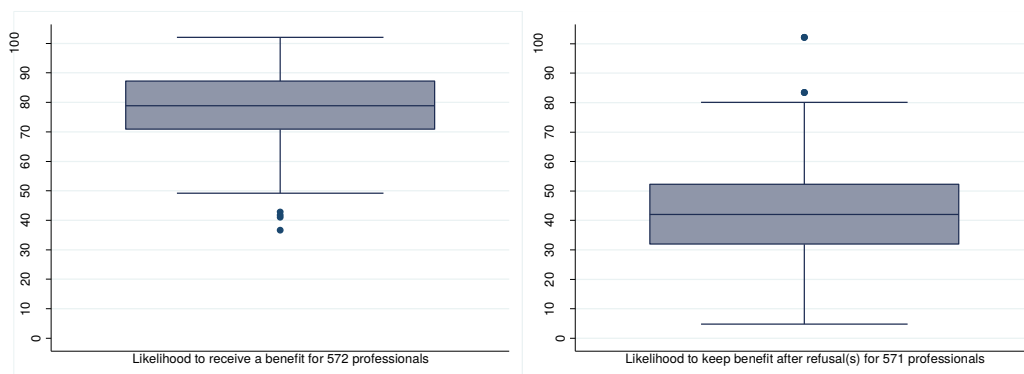
One of the first conclusions to be drawn could be that the use of discretion is not high when it comes to the first eligibility assessment, but that it might increase over the course of the client’s activation pathway. However, based on this frequency table it is impossible to distinguish variation caused by client circumstances or treatment typical of the social assistance agency from the variation we are interested in this paper: variation caused by the professional’s characteristics. Therefore, in the following step, I analysed which level the existing variation was situated at, using a three-level 0-model (M0). Based on the variance at each level randomized in the multi-level model, the variance partition coefficients were calculated.

The Level 2 variance partition coefficient was 44% for the eligibility variable and 30% for the sanction variable. Although the relative variance at the second level is higher for the eligibility variable, the absolute scores indicate that in absolute terms this variance remains higher for the sanction variable (see Table 5-6). This is also clear in Figure 5-5, which shows the distribution of mean predictions for the 571 professionals based on the multi-level 0-models. The variation between professionals is higher for the sanction variable than for the eligibility variable.

Table 5-6. Results of multi-level 0-models

		Likelihood of receiving benefit		Likelihood of keeping benefit after refusal(s)	
Intercept		78.65***		42.67***	
Variance		Absolute score	Variance partition coefficient	Absolute score	Variance partition coefficient
Level 1 = client		175.61***	49.4%	558.49***	64.8%
Level 2 = professional		154.66***	43.5%	255.94***	29.7%
Level 3 = municipality		25.16*	7.1%	47.96**	5.6%
N		5139		4001	
Level 1 = client		572		571	
Level 2 = professional		76		76	
Level 3 = municipality					

Figure 5-5. Box plots of professional-specific predictions based on multi-level 0 models



## 5.2 Modelling heteroscedasticity

In the next models, I identify which first, second and third level variables determine the second-level variance components. In preparation, the best fitted multi-level model with fixed effects at the three levels was estimated (Model 1). The results are shown in Appendix 9, but will not be described in this section, as the fixed effects represent differences in preferences (in the mean) and not differences in variation (in the variances).

In Model 2, I randomized Level 1 variables, namely client characteristics, at the second level (see 4.2.1 The models). The random slopes indicate that social workers disagree more on the effect that certain characteristics have on eligibility, such as having no children compared to having children, homelessness compared to stable or unstable housing, having a Bachelor’s degree compared to a primary education qualification, being healthy compared to having mental health problems, having no adverse life experiences compared to major life events, being not motivated to work compared to

wanting to work or wanting to study, and missing several appointments compared to being diligent (see Table 5-7). It appears, first, that the client characteristics related to vague prescriptions in legislation (work willingness, fairness reasons) are the characteristics that prove significant. Characteristics such as the client's gender or nationality are assessed with similar inter-professional variation, regardless of the category (e.g. female or male). Second, professionals disagree more on characteristics that make them doubt eligibility (e.g. unwillingness to work, homelessness, addictions – see Appendix 9 for the significant effects of these client characteristics included as fixed effects) than on characteristics that lead to higher eligibility likelihood (such as having children or being willing to work).

The results are less straightforward with regard to the sanction variable. Factors that result in higher disagreement among professionals are problematic command of the national language, a problematic housing situation, depression, a severe addiction and several missed appointments. Less variation is found with regard to having a child, having a Bachelor's degree, having an intellectual disability, a slight addiction and negative activation experiences. The client characteristics that result in the smallest variation among professionals are two job refusals compared to one and the refusal of a job that did not match the client's qualifications compared to a job that started early in the morning (see Table 5-7). The conclusion might be that disagreement as to whether or not to sanction clients increases if the client situation is something for which the client can be held personally responsible (language problems, housing, depression without being willing to seek professional help, addictions or missing appointments).



Table 5-7. Overview of the difference at Level 2 – variance based on random slopes models (randomization of Level 1 variables at Level 2)

Level 1 = client characteristics		Difference at Level 2 – variance	
		Likelihood of receiving benefit	Likelihood of keeping benefit after refusal(s)
Refusal frequency	2 (ref = 1)		-75.61
	3 (ref = 1)		-7.51
Type of refused offer	Day contract (ref = early in the morning)		-3.22
	Not matching qualifications (ref = early in the morning)		-65.07
Gender	Man (ref = woman)	n.s.	1.76
Nationality background	Second generation migrant (ref = Bel)	n.s.	-5.31
	Refugee (ref = Bel)	n.s.	3.23
Language	Speaks Dutch poorly (ref = speaks good Dutch)	n.s.	-8.13
	Speaks no Dutch (ref = speaks good Dutch)	n.s.	15.85
Parenthood	One healthy child (ref = none)	-10.55	-37,34
	One sick child (ref = none)	-18.98	0,4
Qualifications	Secondary education qualification (ref = primary education)	0	6.34
	Bachelor's degree (ref = primary education qualification)	36.25	-12.87
Housing situation	Unstable housing (ref = stable housing)	15.1	44.85
	Homeless (ref = stable housing)	53.07	29.5
Mental health	Depressed (ref = no mental health problems)	-10	19.2
	Intellectual disability (ref = no mental health problems)	-36.81	-17.54
Drug addiction	Slight addiction (ref = none)	4.86	-18.21
	Severe addiction (ref = none)	6.19	54.94
Adverse life events	Violent & abusive upbringing (ref = none)	-33.07	3.25
	Lost a child (ref = none)	-13.96	-0.03
Aspirations	Wants to study (ref = willing to work)	0	-9.18
	Not motivated to work (ref = willing to work)	47.77	8.64
Diligence	Missed one appointment (ref = diligent)	-7.41	4.58
	Missed several appointments (ref = diligent)	33.59	11.16
Work experience	Positive work experience (ref = none)	0	4.72
	Negative work experience (ref = none)	39.24	-6.73
Activation experience	One negative activation experience (ref = none)	-6.76	-6.71
	Several negative activation exp (ref = none)	-9.18	-32.83
<b>N</b>	Municipalities	76	76
	Professionals	572	571
	Vignettes	5139	14001

In Model 4, I randomized specific professional variables one by one to check whether inter-professional residual variance depended on the randomized variables. Most variables proved not to be significant. Two variables were significant at a  $p < 0.1$ -level (see Table 5-8). First, the likelihood that the board would follow the advice of the professional: Figure 5-6 presents the residual variance at the second level depending on these variables. The parabola is difficult to interpret, as the measurement '5' seems to

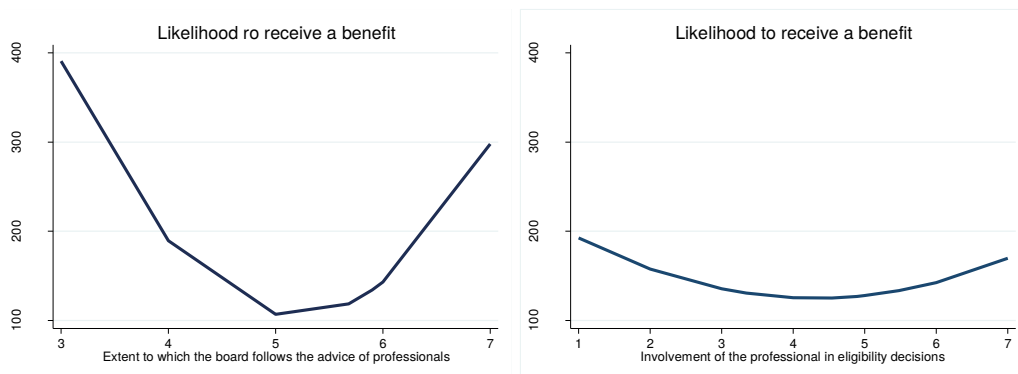
act as a turning point. However, only 2.3% of all professionals indicated that the likelihood of the board’s following the advice of the professional (on a scale from 1 to 7) would be lower than 5. As the residual variance increases after ‘5’, it seems fair to say that inter-professional variation increases with the perceived likelihood that the board would follow the advice of the professional. The effect of the perceived involvement of the professional in eligibility decisions is less strong. Here, the turning point occurs at 4 and 19.06 of the respondents indicated being less involved than 4 out of 7. Here too, the inter-professional variation increases with the perceived involvement of the professional, but less strongly than was the case for *Advice\_professional*.

Table 5-8. Overview of the significance of randomized professional variables

Level 2 = Professional characteristics		Likelihood of receiving benefit	Likelihood of keeping benefit after refusal(s)
Age		n.s.	n.s.
Qualification	Master’s degree (ref = Bachelor’s)	n.s.	n.s.
Advice_professional		**	n.s.
Eligibility_professional		*	n.s.
Activation_professional		n.s.	n.s.
Eligibility_board		n.s.	n.s.
Activation_board		n.s.	n.s.

\* p<0.1; \*\* p<0.05 .

Figure 5-6. Residual variance at the second level by the randomized variables



Model 4 studies differences in inter-professional variation in different types of organizations. To achieve this, a number of agency variables were randomized at the second level. With regard to the residual variation among the sanction intercept, none of the modelled variables became significant. The exception was the existence of an internal guideline concerning the interpretation of the work willingness conditions. Surprisingly, inter-professional variation was higher in agencies that did have such a

guideline than in agencies that did not. This, undoubtedly, confounds the intention of the guideline.

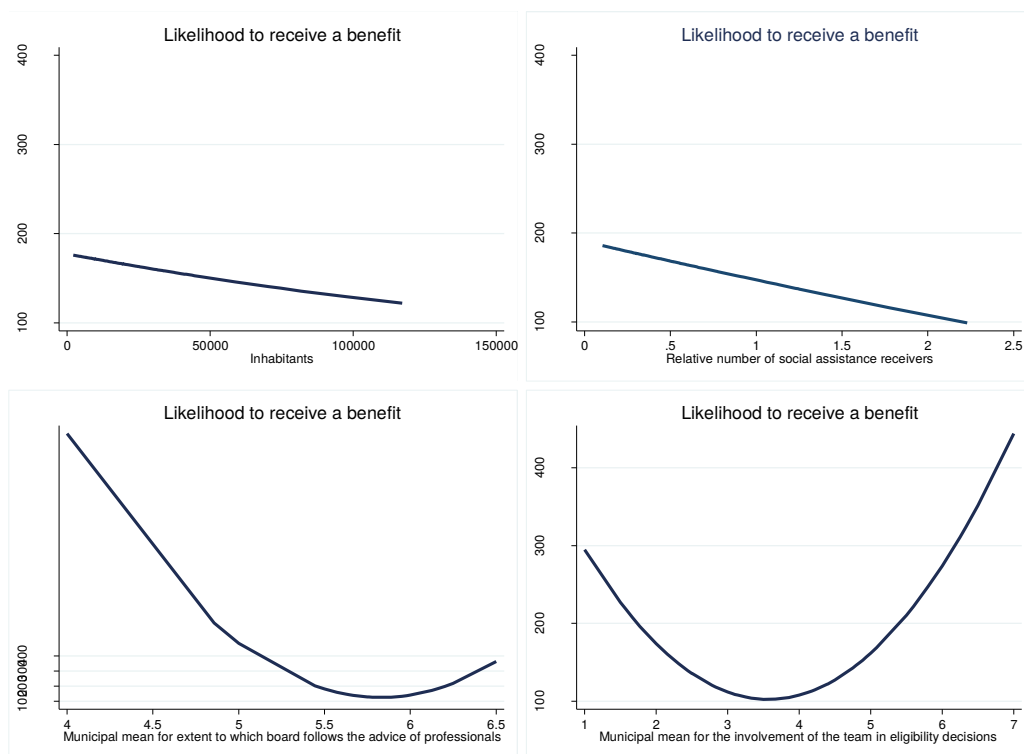
Inter-professional variation on eligibility seemed to be linked to several local variables. Treatment differences decreased with the number of inhabitants and the relative percentage of social assistance beneficiaries. Thus, perhaps counterintuitively, the predicted use of discretion was higher in small municipalities and municipalities with a low number of beneficiaries. Furthermore, agencies in which the respondents generally thought that there was a high likelihood that their advice would be followed by the board displayed higher inter-professional variation. The lowest point of the parabola is around 5.75 and only 28.89% of the 89 municipalities fell below this point. The effect of the generally perceived involvement of the team in eligibility decisions was as surprising as the effect of having internal guidelines, described above. Agencies with more perceived team involvement displayed higher residual variance parameters (one third of the municipalities had a lower involvement than 3.6, the turning point in the figure).

Table 5-9. Overview of the significance of randomized municipality variables

Level 3 = Agency/municipality characteristics	Likelihood of receiving benefit	Likelihood of keeping benefit after refusal(s)
Size of municipality	**	n.s.
Relative percentage of social assistance clients	*	n.s.
Advice_professional_munic	***	n.s.
Eligibility_professional_munic	n.s.	n.s.
Activation_professional_munic	n.s.	n.s.
Eligibility_team_munic	***	n.s.
Eligibility_board_munic	n.s.	n.s.
Activation_board_munic	n.s.	n.s.
Specialization_munic (number of tasks for most of the professionals)	n.s.	n.s.
Guideline_work_willingness Yes (ref. = no)	n.s.	65.93**
Within agency variance in <i>Welfare criticism</i>	n.s.	n.s.
Within agency variance in <i>Welfare strictness</i>	n.s.	n.s.

\* p<0.1; \*\* p<0.05 ; \*\*\* p<0.01

Figure 5-7. Residual variance at the second level by the randomized variables



## 6. Discussion

Professionals' use of discretion is a hotly debated topic in social work, social administration and social policy research (Nothdurfter & Hermans, 2018). Unfortunately, to date, no quantitative information is available on the magnitude of professionals' use of discretion. The debate about the decrease, continuation or even increase of the use discretion is mostly based on either *de jure* discretion (the discretion provided for in regulations or by management) or descriptions of *de facto* (actual use of) discretion by professionals and managers (Evans, 2012). Discretion is seldom measured in terms of variation in client treatment, which is a problematic blind spot. Several claims have been made about the influence of professionals' and agencies' characteristics on the use of discretion, but again these have rarely been verified against large-scale data.

The research presented in this chapter was an initial exploratory investigation. I studied the issue in Belgium, a country whose social assistance legislation is rather vague with regard to the concept of 'work willingness'. The expectation was that such vague descriptions would lead to variation among agencies and professionals (De Wilde & Marchal, 2018; Kazepov & Barberis, 2013). Furthermore, the requirements that each

agency have at least one professional social worker and that the agency be governed by a political board only increased this expectation.

One of the first conclusions to be drawn is that inter-professional variation was limited with regard to the eligibility of the hypothetical cases in the factorial survey used, but high with regard to the likelihood of keeping the benefit after activation refusals (see Table 5-5 and Figure 5-5). It seems that professionals assess the situations of clients who have no resources and no network quite similarly as 'in dire need of financial support'. Inter-professional disagreement increases when it comes to punitive forms of treatment. This conclusion can be read as a comment on the hypothesis of Kazepov and Barberis (2013) (Hypothesis 2). Vague legislation and multi-layered responsibilities do not necessarily lead to large treatment variation. However, this thesis should be tested in more countries and with more outcome variables.

Second, the results in this chapter add depth to the findings presented in Chapter 4 (eligibility) and Chapter 3 (sanctioning). Chapter 4 showed that potentially problematic behaviour did not have a large effect on eligibility. Here, I found that these are the client characteristics about which professionals disagree most. Thus, deservingness assessments based on control, reciprocity and attitude (van Oorschot e.a., 2017) are more important to some professionals than to others. Furthermore, I concluded in the previous chapter that it was merely 'need' that oriented professionals in their eligibility decisions. The results in this chapter add weight to this conclusion. The inter-professional differences were smallest with regard to client characteristics that could be considered as increasing this need (e.g. having children).

However, a more subtle form of deservingness assessment was revealed, based on the sanction results. The most professional consensus was found with regard to sanctioning after two refusals. This means that some professionals considered sanctioning after one refusal to be highly acceptable whilst others did not. I concluded in Chapter 3 that the sanction likelihood after one refusal was quite high, but this was clearly not the case among all professionals. The lack of consensus also increased again with the third refusal. Some professionals saw no difference between two and three refusals (32% of all cases), while others increased the sanction prediction considerably. Yet, some professionals also seemed to begin to doubt whether sanctioning was the right answer if a client repeatedly refused job offers. In 15% of all client cases, professionals predicted the sanction likelihood to be lower after three refusals than after two refusals. Furthermore, the consensus in sanction likelihood was lower with regard to client characteristics that could be considered by some professionals as things that the client had personal control over (addiction, housing situation, depression (vs. intellectual

disability), etc.). Chapter 3 showed some straightforward effects of the control deservingness criterion (van Oorschot e.a., 2017), such as the effect of the frequency of refusal on the sanction likelihood. However, the variety of interpretations of 'in control of the personal situation' adds to our understanding of the complexity of deservingness assessment (Moller & Stone, 2013). These nuances can only be highlighted by studying inter-professional variation in detail.

Apart from studying the effects of client characteristics on variation in treatment, I wanted to understand which professional and agency characteristics increase or decrease inter-professional variation. Here, the evidence is less strong than is often claimed in qualitative research. None of the possible organizational characteristics that could be included in the models and that could have an impact on inter-professional variation – such as internal guidelines concerning eligibility conditions, the importance of team involvement in decision-making or the involvement of the board in eligibility and activation considerations – proved significant. This means that, based on my data (which is admittedly limited with regard to this item), organizational structure neither eliminates or encourages the use of discretion.

However, the group of professionals who felt they were considered important actors in the decision-making process displayed more inter-professional variation. The question is whether this effect can be understood as an agency or a personal effect. If it is an agency effect, there should be an element external to the professionals which increases uniformity in treatment in agencies and which I was unable to capture through the other variables. A personal effect would mean that some professionals have the tendency to follow their own lead in treatment and, thus, view themselves as important actors. The fact that the municipality mean of one of the variables, which indicates the importance of professionals (*Advice\_professional\_municipality*), was a significant explicator of the difference in inter-professional variation between agencies could mean that it is indeed an agency effect. However, the effects are small and not easy to interpret. Furthermore, the constructed variables are only proxies of the phenomena they are intended to measure. Lastly, strict management supervision and process management are not common in Belgium. The research should be repeated in countries which have different traditions in this respect.

In conclusion, Table 5-10 shows that most of the hypotheses that were formulated based on the literature study were not or only partly confirmed. This might be due to limitations in the data. Yet it is also, at least, a call to supplement qualitative studies of organizations with a more quantitative approach that links organizational settings with

client outcomes. This might shed new light on findings that are generally taken for granted in literature.

Table 5-10. Overview of confirmed and unconfirmed hypotheses

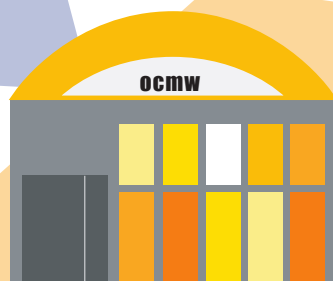
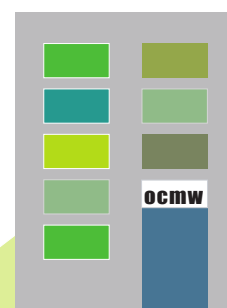
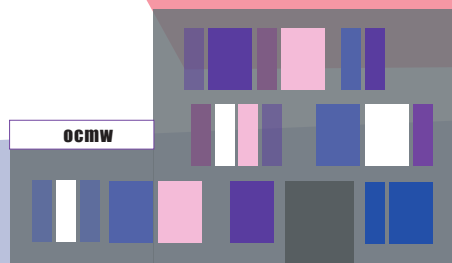
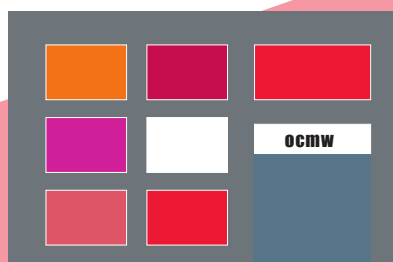
Hypothesis	Likelihood of receiving benefit	Likelihood of keeping benefit after job refusal(s)
H1: Treatment variation and thus the use of discretion is high among professionals who work in settings with heterogeneous problems and client circumstances.	Partly confirmed	Confirmed
H2: Treatment variation is high in countries with legislation and institutional organization that allows for interpretative discretion.	Not confirmed	Confirmed
H3: Professionals will vary more in their treatment of clients in situations that are described only vaguely in legislation (e.g. work willingness).	Confirmed	Partly confirmed
H4: Treatment variation among professionals will be highest with regard to client characteristics that deviate most from direct prescriptions in legislation (e.g. variation in eligibility decisions concerning a client who is not willing to work)	Confirmed	Partly confirmed
H5: Older and more experienced professionals will display more inter-professional variation than younger professionals.	Not confirmed	Not confirmed
H6: Professionals with a higher educational level will display more inter-professional treatment variation than professionals with a lower educational level.	Not confirmed	Not confirmed
H7: Among professionals who experience a lot of autonomy, more inter-professional treatment variation will be observed than among professionals who do not experience autonomy.	Partly confirmed	Partly confirmed
H8: In agencies where the majority of professionals experience autonomy, inter-professional treatment variation will be higher than in agencies where autonomy is generally low.	Partly confirmed	Not confirmed
H9: Managerial techniques focused on minimalizing internal variation will decrease inter-professional variation.	Not confirmed	Not confirmed
H10: In agencies with less variety in values and attitudes, less treatment variation will be found.	Not confirmed	Not confirmed





# Leefloon krijgen of verliezen:

HOE DECENTRAAL WORDT HET BIJSTANDSBELEID  
GEÏMPLEMENTEERD IN VLAANDEREN?



## Chapter 6. Leefloon krijgen of verliezen: hoe decentraal wordt het bijstandsbeleid geïmplementeerd in Vlaanderen?

Marjolijn De Wilde

---

*Adjusted version of De Wilde, M. (2017). Leefloon krijgen of verliezen: hoe decentraal wordt het bijstandsbeleid geïmplementeerd in Vlaanderen? In: P. Raeymaeckers, Jaarboek Armoede en Sociale uitsluiting. Tijd voor Sociaal Beleid. Acco: Leuven,*

### 1. Extensive English Summary

In 1974, municipally organized public assistance committees (Dutch: *Commissies van Openbare Onderstand*) – which would later become the Public Centres for Social Welfare (PCSWs) – were made responsible for allocating last-resort social assistance benefits (originally called the *bestaansminimum* in Dutch, now known as the *leefloon*). It is perhaps unsurprising that these existing, decentralized organizations were chosen. While the choice undoubtedly represented the continuation – and formal anchoring – of a tradition, the number of beneficiaries was also so low (approximately 10,000) that it seemed absurd to set up a separate, nationally organized service. It is also unsurprising that, for this benefit – the last possible recourse for people who have exhausted all their other rights – services were chosen which allowed ample room to tailor the level of assistance to the specific needs of the client and which could take the local context (existing services, existing problems) into account.

The situation has changed considerably since then and the number of social assistance beneficiaries has risen to more than 120,000 per month (Van Mechelen, Zamora, & Cantillon, 2016). It is no longer the case that these beneficiaries are a residual group with very diverse characteristics. The current population of social assistance recipients is made up of several large groups which, despite having clear profiles, are difficult to cater for in the social security system as it is currently organized. The first of these groups comprises women (and sometimes men), often low-skilled, who have never worked or not worked enough to be entitled to social security and who, after an increasingly common divorce or break-up, become dependent on social assistance benefits. The second group are students who are responding to the growing importance of qualifications but who cannot rely on their parents for financial support, and who, because of their age and limited work history, have not yet contributed enough to the social security system. The third group is made up of low-skilled people who have difficulties finding work at their skill level. The fourth group are refugees who have not yet worked enough in Belgium, and the fifth are young graduates who are not yet

entitled to what is known as the ‘integration allowance’ after completing their training and before starting their first job (Van Mechelen, Zamora, & Cantillon, 2016). These are just a few examples of the growing groups of people who are unable to make use of the Belgian social security system because of its basis in previous work history and contributions. In the implementation of social security benefits, we are also seeing a trend towards stricter applications of the eligibility conditions (and also sanctions), as a result of which a growing number of rejected or sanctioned clients are becoming dependent on social assistance benefits, temporarily or otherwise.

This means that financial assistance services for vulnerable groups are becoming increasingly decentralized. But this is not done in service of the ideals commonly proposed by the advocates of decentralization, namely the chance to tailor services to local capacities and needs and to encourage co-governance by local authorities (see introduction). On the contrary, the trend is driven by sociodemographic changes and by changed views on who is entitled to sickness and unemployment benefits. This unintended decentralization of social security is part of the problem of passive subsidiarity, in which groups of clients are placed under the care of local authorities though there are no clear budgets or policy goals to support them (Bonoli & Trein, 2016; Kazepov, 2005).

In light of this, it is important to verify whether substantial variation exists between municipalities and how such variation could be explained. The question, then, is whether decentralization mainly manifests as variation because of the sociodemographic context (first justification for decentralization) or because of the specific policy preferences of local governments (second justification for decentralization) or whether, on the other hand, there are other random factors related to each PCSW that explain the variation. One observation from the research presented in this chapter is that municipality variation with regard to predictions about social assistance payments and sanctions was extremely small (see Table 6-2). However, this does not negate the fact that some municipalities differed from each other by up to 20% in their predictions of the likelihood that a benefit or sanction would be issued (see Figure 6-1). These clear differences between 10 to 20 of the 90 municipalities we studied meant it was necessary to verify the explanatory factors.

As expected, we found that the context of the municipality (size, prosperity, etc.) plays a role. However, it is difficult to interpret the effects we found as conscious choices, or in other words, as choices intended to tailor a PCSW’s policy to the local context so that the target group or wider community would benefit from it. One of the effects, in particular, raises additional questions: the drop in sanction probability in municipalities

where public employment offices are located. This effect may be the result of a conscious tailoring of policy to the local situation, but it remains difficult to interpret the rationale behind it. A second argument for decentralization is the importance of giving local authorities and population groups a greater say (Fording e.a., 2007; Soss e.a., 2011). This political influence is evident in the strong effect of the board chair's political affiliation. In addition, there are a number of factors that are not directly related to the local context or policy preferences of the local authorities. To some extent, the implementation of assistance policy depends on the composition of the PCSW team (e.g. age and work experience). Yet, what appears to be even more important is the effect of the team members' and team leader's general attitudes towards the welfare state. This may indicate a municipality's sanction sensitivity (Holzner, Munz, & Büttner, 2009). There is a chance that team members are selected in part because of these attitudes and that this, too, is a result of the local authorities' policy preferences. Other elements around which PCSWs may make conscious choices include social workers' degree of specialization and workload. These organizational choices enable PCSWs to experiment and to work towards their chosen implementation policy (see Table 6-3).

The results described above pose two potential problems. First, this type of assistance – and the very intensive, dedicated follow-up it involves – was never intended for large groups of beneficiaries. Second, and despite the limited variation observed among PCSWs, the question is whether it is a good idea to make basic services for a growing and very precarious group of people dependent on factors such as the political preference of the PCSW board's chair, the age of its social workers, the organization of the team or the personal opinions of social workers. Given that many potential beneficiaries belong to clearly defined groups (see list in the second paragraph of this section), it may be possible to shift towards administrative allocation of social assistance benefits for these groups using simple criteria. This would likely benefit social workers' workloads and also speed up case handling. Social workers' true work (career counselling, activation, etc.) could then begin following allocation or refusal on the basis of administrative information. After all, it is to be expected that certain clients will fall through the cracks of the administrative allocation system. For these clients, the social assistance system could again be put to work as it was originally in 1974: largely reliant on the discretionary space allowed to municipalities and social workers, and the final safety net for people who have nowhere else to turn.

## **2. Inleiding**

Meer dan 120.000 leefloongerechtigden per maand in 2016 in België. Dat is niet wat de beleidsmakers in 1974 voor ogen hadden toen ze het bestaansminimum (voorloper van

het leefloon) introduceerden als een uitkering voor personen zonder of met onvoldoende bestaansmiddelen, die op geen enkele andere uitkering beroep konden doen. Het eerste jaar na de introductie waren er ongeveer 10.000 gerechtigden. Dat destijds gekozen werd voor een decentrale organisatie van deze financiële bijstandsverlening is niet vreemd. Het betrof enerzijds een officialisering en formalisering van de uitkeringen die de C.O.O.'s (Commissies van Openbare Onderstand – voorlopers van de OCMW's) eerder ad hoc ter beschikking stelden. Deze C.O.O.'s waren gemeentelijk georganiseerd. Het bleek logisch deze organisatievorm te bestendigen. Anderzijds kadert de lokale organisatie van bijstandsinstanties binnen een internationaal gedeeld uitgangspunt dat de noden van de kleine groep bijstandsgerechtigden het best ingeschat en gelenigd kunnen worden door lokale overheden, ambtenaren en hulpverleners. Zij kunnen daarbij rekening houden met de plaatselijke mogelijkheden (Kazepov, 2010; Minas e.a., 2014; Vandenbroucke e.a., 2016). De lokale organisatie van de bijstand in België gaat echter niet enkel om het aanpassen van centraal vastgelegde programma's aan de plaatselijke noden en mogelijkheden. Ze laat ook ruimte voor eigen gemeentelijke beleidslijnen en voor plaatselijk verankerde experimenten. Ook dit is een internationaal gedeelde reden om beleid decentraal te organiseren. Men gaat ervan uit dat de echelons die de gevolgen dragen van het beleid (bv. gemeente die invloed ondervindt van de aanwezige armoede) ook mee het beleid moeten kunnen maken (Kazepov, 2010; Vandenbroucke e.a., 2016). Deze ruimte voor eigen beleidslijnen en experimenten doet veronderstellen dat er verschillen zijn in de dienstverlening voor cliënten. Ik bestudeer in dit hoofdstuk in welke mate die diversiteit inderdaad zichtbaar wordt in de praktijk en hoe deze verklaard kan worden. De vraag is dan vooral of het overwegend de socio-demografische context of de lokale beleidsvoorkeuren zijn die de variatie verklaren.

### **3. Theoretische achtergrond**

#### **3.1 Leidt decentralisatie tot variatie in de implementatie van sociaal beleid?**

België heeft een nationaal verankerde, maar lokaal geïmplementeerde bijstandswetgeving. Bij de aanvang van de Belgische financiële bijstand in 1974 werden gemeenten verantwoordelijk gesteld voor het verstrekken van de dienstverlening. Reeds van bij aanvang, maar des te meer door de tijd heen, is er een kader voorzien dat de grenzen hiervan aangeeft. Zo ligt de hoogte van de leefloonbedragen vast, en werd voorgeschreven welke inkomsten in rekening genomen mogen worden om te bepalen of iemand al dan niet over voldoende bestaansmiddelen beschikt. Andere objectieve criteria, zoals verblijfsstatuut, nationaliteitsvoorwaarden en leeftijds grenzen werden gedetailleerd opgenomen in wetgeving en omzendbrieven. Verder zijn er een aantal

richtlijnen opgesteld rond het sociaal onderzoek dat dient vooraf te gaan aan het verstrekken van een leefloon. Dit sociaal onderzoek moet eerst en vooral door een gediplomeerd maatschappelijk werker uitgevoerd worden en sluit dingen als een huisbezoek in. Uit het sociaal onderzoek volgt een advies met betrekking tot leefloonverstrekking, dat ter beslissing voorligt op de raad of het bijzonder comité van het OCMW, beide politiek samengesteld (De Wilde, 2016, Koninklijk Besluit betreffende de minimumvoorwaarden voor het sociaal onderzoek, 2013; Wet Recht op Maatschappelijke Integratie, 2002). Ondanks deze (beperkt toenemende) reglementering, blijft er een grote lokale vrijheid met betrekking tot de implementatie van het leefloonbeleid. Dit zijn enkele van de elementen die potentieel de verscheidenheid beïnvloeden: de organisatie van de sociale dienst, het aantal beschikbare mensen om de taken uit te voeren, hun specialisaties, de procedures die gevolgd dienen te worden voor een leefloonaanvraag kan voorgelegd worden aan de raad, en niet in het minst: de interpretatie van meer vage leeflooncriteria als 'werkbereidheid'... (De Wilde, 2016)

Er is relatief weinig bekend over het effect van deze wisselwerking tussen nationale kadering en de lokale autonomie op de praktijk. Het onderzoek met betrekking tot decentralisatie bestudeert voornamelijk de structuur zelf en hoe deze tot stand komt (Minas e.a., 2012; van Berkel, 2006). Er bestaan weinig studies over het effect van deze decentralisatie op de implementatie of op de eigenlijke begeleiding van een cliënt (Carpentier, 2016; Soss e.a., 2001). Praktijkelementen die wel onderzocht werden voor België, zijn de duur van het ontvangen van een leefloon, het type uitstroom en de wijze van organiseren van activeringsprogramma's. Er blijken (ook als rekening gehouden wordt met verschillen in kenmerken van de cliënten) grote verschillen tussen de Belgische gemeenten met betrekking tot al deze elementen (Carpentier, 2016; Hermans, 2005). Dit hoofdstuk is op twee manieren aanvullend bij deze studies. Enerzijds is de focus anders en breder: ik bekijk zowel de kans op een uitkering als de kans op een sanctie bij tekenen van werkonwilligheid. Anderzijds, ben ik – door het experimentele opzet van de studie – in staat om te controleren voor zowel invloeden vanuit de cliënt, de maatschappelijk werker die de cliënt begeleidt en de gemeente waarin de cliënt geholpen wordt. Uit andere literatuur blijkt dat het niveau van de maatschappelijk werker belangrijk is (De Wilde, 2016; De Wilde & Marchal, 2018). Dit betekent dat de studies die dit niet doen wellicht een vertekend beeld van de reële verschillen tussen gemeenten geven.

### 3.2 Wat verklaart de variatie in begeleiding tussen verschillende lokale overheden?

Het meten van de variatie is één ding. Een tweede onderzoeksvraag in dit hoofdstuk peilt naar de kenmerken op gemeente of OCMW-niveau die deze verschillen verklaren. Zoals aangegeven in de inleiding, is één van de doelen van decentralisatie het aanpassen van de begeleiding aan de lokale mogelijkheden. Moest decentralisatie ook dat effect hebben, dan zouden een aantal lokale demografische kenmerken een invloed moeten hebben op het bijstandsbeleid. Hier is internationaal enige evidentie voor te vinden. De lokale arbeidsmarktsituatie kan op twee manieren het bijstandsbeleid beïnvloeden. Ten eerste, kan het dat cliënten in gebieden met een hoge werkloosheid makkelijker een uitkering krijgen en minder makkelijk gesanctioneerd worden, omdat de OCMW-medewerkers beseffen dat het niet vinden van werk in deze gebieden niet op conto van de cliënten geschreven kan worden (van Oorschoot, 2000). Ten tweede, kan het ook zijn dat OCMW's meer geneigd zijn om cliënten die zich niet bereid tonen om te werken te sanctioneren, omdat ze verwachten dat cliënten alle mogelijkheden die er zijn aangrijpen (Fording e.a., 2007; Keiser e.a., 2004; Soss e.a., 2011). Verder blijkt er meer gesanctioneerd te worden in grote gemeenten. De auteurs linken dit aan omstandigheden (hoge armoede, weinig jobs...) die het moeilijk maken om aan de verwachtingen van de bijstandsorganisaties te voldoen (Fording e.a., 2007). Dat er een hoger percentage leefloongerechtigden is in arme gebieden, lijkt logisch. Het is echter onduidelijk of eenzelfde cliënt in deze gebieden makkelijker een leefloon krijgt dan in rijkere gebieden, of net niet. Met betrekking tot het sanctioneren van cliënten in deze arme regio's zijn de resultaten tegenstrijdig. Sommige onderzoekers besluiten dat er meer gesanctioneerd wordt (Fording e.a., 2007), andere net minder (Monnat, 2010). Wel blijkt de sanctiegraad lager in gebieden met veel leefloongerechtigden (Fording e.a., 2007; Soss e.a., 2011).

Een tweede drijvende factor achter decentralisatie is de wens om lokale beleidsmakers zeggenschap te geven over domeinen waarmee zij lokaal aan de slag moeten. Dit kan betekenen dat de politieke overtuigingen van burgers en hun beleidsmakers in een bepaald gebied de eigenlijke implementatie bepalen. Het blijkt inderdaad zo te zijn dat hoe conservatiever burgers stemmen, hoe minder uitkeringen er verstrekt worden en hoe meer er gesanctioneerd wordt (Fording e.a., 2007; Soss e.a., 2011).

Ten slotte, is variatie in de implementatie wellicht niet te vermijden omdat de bijstand verstrekt wordt door bijstandsorganisaties met een eigen cultuur, een eigen mix van medewerkers met specifieke kenmerken en een eigen beslissingsstructuur. Zolang mensen een rol te spelen hebben in de verstrekking van de uitkeringen, zal er variatie

optreden. Dit bepaalt de verschillen tussen maatschappelijk werkers in één OCMW, maar zeker ook de verschillen tussen OCMW's met verschillende medewerkersteams. Voorlopig zijn er weinig aanwijzingen in de literatuur voor indicatoren die een directe invloed hebben. Elementen die in kwalitatief onderzoek aan bod komen, zijn de volgende: de team bijeenkomsten waarin het belang van activering al dan niet sterk vooropgesteld wordt of waarin mogelijke vrijstellingen van activering systematisch in vraag gesteld worden (Brodkin, 2011; Fording e.a., 2007), de sanctiegevoeligheid binnen een organisatie (Holzner e.a., 2009), de complexiteit van het traject van aanvraag tot uiteindelijke uitkering of sanctie (Fording e.a., 2007), het gebrek aan tijd of gelegenheden om complexe cases in detail te bespreken met collega's (Fording e.a., 2007). Een element dat in Vlaamse context onderzocht werd, is de mate van betrokkenheid van maatschappelijk werkers bij de uiteindelijke beslissingen omtrent uitkeringen en activering. In OCMW's waar de maatschappelijk werkers zich sterker betrokken voelden bij de uiteindelijke beslissingen, werkten maatschappelijk werkers minder met sancties om cliënten te overtuigen van het belang van activering. Ze kozen vaker voor positieve interventies (Raeymaeckers & Dierckx, 2013)

Het voorgaande betekent dat ik veronderstel dat drie types kenmerken een invloed hebben op het bijstandsbeleid in de Vlaamse gemeenten: de socio-demografische kenmerken van de gemeente, de politieke strekking in de meerderheid en de OCMW-eigen kenmerken (beslissingsstructuur, samenstelling medewerkers, attitudes maatschappelijk werkers en hoofd van de sociale dienst...).

## **4. Methodologie**

### **4.1 Opzet onderzoek**

Om de intergemeentelijke begeleidingsverschillen inzake financiële bijstand in Vlaanderen na te gaan, gebruikte ik data van een experimenteel vignet-onderzoek waaraan 585 maatschappelijk werkers uit 89 Vlaamse OCMW's deelnamen. De OCMW's vormden een representatieve steekproef van alle Vlaamse OCMW's. In die OCMW's werd volgens leeftijd, gender en functie (al dan niet hoofdmaatschappelijk werker) een representatieve steekproef genomen van maatschappelijk werkers die betrokken waren bij het beslissingsproces over het al dan niet verstrekken van een leefloon. We hielden enkel de gemeenten over waar meer dan

Aan deze maatschappelijk werkers werden negen cliëntbeschrijvingen voorgelegd. Ze moesten – via gesloten vragen – voor elk van deze cliënten een aantal begeleidingspistes uitzetten. De cliëntbeschrijvingen hadden een identieke opbouw, maar verschilden met betrekking tot een aantal cruciale kenmerken. De kenmerken die



niet varieerden waren de leeftijd van de cliënt (22 jaar), de burgerlijke staat (alleenstaand) en de financiële situatie (geen inkomsten en geen spaargeld). De andere wel variërende cliëntkenmerken verwijzen naar socio-demografische kenmerken (gender, opleiding, nationaliteit, woonsituatie...), de werkbereidheid van de cliënt, en mogelijke gezondheids- of billijkheidsredenen die een cliënt vrijstellen van werkverplichtingen. In de verhalen werden alle categorieën (bv. man – vrouw) van alle kenmerken (bv. gender) met elkaar gecombineerd. Dit betekent dat elke categorie even vaak voorkwam in alle verhalen die gepresenteerd werden aan de respondenten en dat bepaalde combinaties van kenmerken voldoende vaak voorkwamen om er statistisch verantwoorde uitspraken over te kunnen doen.

Over deze hypothetische cliënten beantwoordden de respondenten, op een schaal van 1 (zeer onwaarschijnlijk) tot 7 (zeer waarschijnlijk) een aantal vragen. Wij focussen in dit hoofdstuk op:

- de verwachte kans op een uitkering;
- de verwachte kans op een (tijdelijk) verlies van de uitkering als de cliënt een werk- of activeringsaanbod weigerde dat ...
  - o ... om vijf uur 's morgens startte;
  - o ... bestond uit een dagcontract;
  - o ... niet overeenkwam met het eigen diploma.

Er werd de respondent gevraagd om te voorspellen wat de eindbeslissing van de raad zou zijn. Dit betekent dat ze niet hun eigen voorkeur moesten weergeven, maar wat ze verwachtten dat er binnen hun OCMW beslist zou worden. Daarbij dienden ze rekening te houden met alle actoren en processen die hieromtrent werkzaam zijn in hun OCMW / gemeente. Zo zijn potentieel de OCMW-raad, de teamleider en het team invloedrijk in dit proces (De Bie & Vandebussche, 2016; De Wilde, 2016). De respondenten moesten het uiteindelijke eindoordeel inschatten. Deze inschatting kan op minstens twee manieren een vertekend beeld geven. Enerzijds kan de respondent zich vooral laten leiden door eigen voorkeuren, eerder dan door wat de raad zou beslissen. Anderzijds kan de inschatting van de respondent van de raad afwijken van hoe die in de praktijk beslist. We kozen toch voor deze methode omdat we niet geïnteresseerd zijn in het oordeel of de voorkeur van de maatschappelijk werker zonder dat zij geen rekening dient te houden met andere actoren. We zijn daarentegen geïnteresseerd in de reële praktijk. Er zijn twee factoren die de validiteit van dit onderzoek verhogen. Eén van de vragen aan de respondenten was in hoeverre zij vonden dat het advies dat zij aan de raad deden gevolgd werd door die raad. Er blijkt een 80% ingeschatte kans dat dit het geval is. Dit betekent dat zelfs wanneer de resultaten van dit onderzoek eerder het

advies van de maatschappelijk werker of een inschatting van de voorkeuren van de raad, dan de uiteindelijke beslissing bevatten, dit advies niet sterk zal afwijken van wat er weldegelijk zou gebeuren. Bijkomend vernamen we uit gesprekken met maatschappelijk werkers en raadsleden, dat één van de onderzochte beslissingen, namelijk 'het aanvragen van een sanctie of stopzetting van het leefloon', typisch iets is waarrond de maatschappelijk werker het initiatief neemt. De cliënt zal er immers niet om vragen en er bestaan weinig plaatsen waar maatschappelijk werkers systematisch moeten rapporteren over werkweigerings. Ze kiezen meestal zelf wanneer ze een werkweigerings met hun team en uiteindelijk de raad bespreken. Dit betekent dat een voorspelling van sanctie door de betrokken maatschappelijk werkers wellicht een relatief groot realiteitsgehalte heeft.

#### 4.2 Analyses

Om de onderzoeksvragen te beantwoorden hanteerde ik multi-level-modellen. Typisch aan een multi-level-model is dat de analyse rekening houdt met de clustering van de antwoorden. Ik ga ervan uit dat de antwoorden van de maatschappelijk werkers binnen één OCMW meer gelijkend zijn dan de antwoorden van maatschappelijk werkers in verschillende OCMW's. Verder zijn de voorspellingen over hypothetische cliënten beoordeeld door één maatschappelijk werker wellicht meer gelijkend dan de voorspellingen van verschillende maatschappelijk werkers. Dit betekent dat de data een drie-level-structuur hebben, met op het hoogste niveau de OCMW's (89), het tweede niveau de maatschappelijk werkers / respondenten (587) en het derde niveau de experimenteel gevarieerde hypothetische cliëntverhalen (9 per respondent).

In een eerste model onderzocht ik<sup>29</sup> of de verschillen in uitkeringsgerechtigdheid of sanctie vooral te voorspellen zijn aan de hand van verschillen tussen OCMW's, verschillen tussen maatschappelijk werkers in één OCMW of verschillen naar aanleiding van de cliëntkenmerken. Zo kunnen de verschillen tussen cliënten doorslaggevend zijn voor de begeleidingskeuzes. Mogelijk worden cliënten met kinderen systematisch anders begeleid dan cliënten zonder kinderen (De Wilde & Marchal, 2016). Daarnaast kan een maatschappelijk werker eigen begeleidingstradities of -voorkeuren of eigen inschattingen van de beslissingen door de raad hebben, wat resulteert in verschillende antwoorden op de survey naargelang de respondent (De Wilde, 2016). In dit hoofdstuk ben ik echter vooral geïnteresseerd in de variatie afhankelijk van het OCMW waar een cliënt zijn aanvraag doet. Als de variatie in antwoorden tussen OCMW's hoog is,

---

<sup>29</sup> De berekening gebeurde op de nul-modellen (= drie-level-model zonder toevoeging van verklarende variabelen) aan de hand van de onverklaarde variantie die in een multi-level-model typisch is opgedeeld in het aantal beschikbare levels.

betekent dit dat de gedecentraliseerde organisatie van de bijstand in België, zich ook toont in de praktijk.

In een tweede model trachtte ik de variatie in voorspellingen te verklaren aan de hand van kenmerken van het OCMW of de gemeente, door een aantal verklarende variabelen toe te voegen aan het bovenstaande model<sup>30</sup>. Ik testte vijf soorten variabelen (zie Table 6-1 voor een overzicht van de variabelen): ten eerste socio-demografische kenmerken van de bevolking van de gemeente, ten tweede de politieke context van het OCMW, ten derde de kenmerken van de deelnemende maatschappelijk werkers<sup>31</sup> (gemiddeld over alle medewerkers), ten vierde werkcontext gebonden kenmerken, en ten vijfde variabelen over de opinie over de welvaarsstaat, over de nood aan controle van uitkeringsgerechtigden en over de mate van misbruik van uitkeringen. Ik nam in de analyses zowel de gemiddelde opinie per gemeente (= gemiddelde van alle deelnemende maatschappelijk werkers), als de opinie van het teamhoofd sociale dienst op.

---

<sup>30</sup> Ik bespreek in dit hoofdstuk enkel de verklarende variabelen op OCMW- en gemeenteniveau. In het model zijn echter ook steeds de cliëntvariabelen en de significante respondentvariabelen opgenomen.

<sup>31</sup> Vermits een representatieve steekproef van maatschappelijk werkers werd genomen, komen de gemiddelde leeftijd, werkduur en gender van de deelnemers overeen met deze van de reële populatie maatschappelijk werkers.

Table 6-1. Overzicht kenmerken OCMW en gemeente

Kenmerken	Categorieën of minima en maxima	Gemiddelde <sup>32</sup>
Socio-demografische kenmerken		
Grootte gemeente	2110 – 502604	155899
Relatief aantal leefloongerechtigden	0.1 – 2.2%	1.1%
Relatief aantal leefloongerechtigden in activering	6.92 – 48.3%	23.4%
Relatief aantal inwoners van vreemde herkomst	3.6 – 47.3%	23.27%
Percentage werklozen in de gemeente	3.4 – 16.3%	10.1%
Welvaartsindex van de gemeente	88 – 139	103.82
Dekking kinderopvang voor 0-3-jarigen	19.2 – 68.9%	42.53%
Politieke kenmerken		
Politieke partij van de voorzitter van de OCMW-raad	1 = N-VA (27%), 2 = CD&V (34%), 3 = Open Vld (8%), 4 = SP.a (24%), 5 = Andere (8%)	
Kenmerken maatschappelijk werkers		
Gemiddelde leeftijd	25-59.29	38.21
Relatief aantal mannen bij maatschappelijk werkers	0-100%	19.8%
Relatief aantal maatschappelijk werkers met kinderen	0-100%	82.1%
Kenmerken werkcontext		
Anciënniteit	0.5 – 35.53	11.8
Mate van specialisatie in het OCMW	1 = voornamelijk 1 type functie (intake, begeleiding, activering of teamhoofd) (17%), 2 = voornamelijk 2 types functies (65%), 3 = voornamelijk 3 types functies (87%), 4 = gemengd (13%)	
Werkdruk in het OCMW (= aantal cliënten per maand / werkregime)	0.1 – 2.91	0.93
Gemiddelde inschatting van de invloed van maatschappelijk werker – team en teamverantwoordelijke op beslissingen over het toekennen van uitkeringen per OCMW (latente variabele)		
Gemiddelde inschatting van de invloed maatschappelijk werker – team en teamverantwoordelijke op beslissingen over de activering van cliënten per OCMW (latente variabele)		
Gemiddelde inschatting van de invloed van de raad op beslissingen over het toekennen van uitkeringen per OCMW		
Opinie-variabelen		
Gemiddelde waardering van het effect van de welvaartstaat op burgers per OCMW (latente variabele)		
Gemiddelde oordeel over de mate aan controle en straffen van uitkeringsgerechtigden per OCMW (latente variabele)		
Gemiddelde oordeel over misbruik van uitkeringen per OCMW (latente variabele)		
Waardering van het effect van de welvaartstaat op burgers door teamhoofd sociale dienst (latente variabele)		
Oordeel over de mate aan controle en straffen van uitkeringsgerechtigden door teamhoofd sociale dienst (latente variabele)		
Oordeel over gepercipieerd misbruik van uitkeringen door teamhoofd sociale dienst (latente variabele)		

Noot: NVA: nationalistische rechts geöriënteerde partij; CD&V: Christendemocratische centumpartij; Open VLD: Liberale centrum-rechtse partij; Sp.a: Socialistische centrum-linkse partij

## 5. Resultaten

### 5.1 Verscheidenheid tussen gemeenten

Een eerste vraag is of er zoveel variatie is in de begeleidingen omtrent het leefloon als men zou verwachten op basis van het type bijstandssysteem in België (= aanzienlijke

<sup>32</sup> De gemiddelden zijn gewogen op basis van de vignetten in de sample, dus niet op basis van de effectieve verdeling van leefloongerechtigden in Vlaanderen of op basis van de gemeenten (één meting per gemeente).

lokale autonomie). Table 6-2 geeft een eerste indruk. Op de eerste rij zijn de gemiddelde voorspellingen van de kans op een leefloon en op sanctionering opgenomen, en dit van een gemiddelde cliënt, begeleid door een gemiddelde maatschappelijk werker in een gemiddeld OCMW. Deze schattingen geven niet zoveel informatie vermits deze gemiddelde cliënten niet overeenkomen met de reële gemiddelde cliënt in de Vlaamse OCMW's (zie Sectie 4. Methodologie).

Interessanter is het om te kijken naar de variatie rond deze gemiddelden (Rij 2-5). Daaruit blijkt dat de variatie in voorspellingen het grootst is met betrekking tot de sanctie-variabelen (zie totale variatie op Rij 2). Dit betekent dat een grote gelijkenis is in de kans op het ontvangen van een leefloon, over alle cliënten, respondenten en OCMW's heen.

Opvallend is dat het stuk van de variatie op het niveau van de gemeente / het OCMW bijzonder klein is (zie Rij 3 en 6). Slechts zes tot negen procent van de variatie is te verklaren op dit niveau (zie Rij 6). De variatie op niveau van de begeleidende maatschappelijk werker en de cliënt is opmerkelijk veel hoger. Concreet betekent dit dat de voorspelde begeleidingsvariatie tussen gemeenten kleiner is dan men zou verwachten wetende dat de wetgeving veel ruimte tot eigen beleidskeuzes laat. Deze ruimte lijkt vooral te worden genomen door de maatschappelijk werker (36-60% van de variatie op dit niveau, berekend op basis van Rij 4).

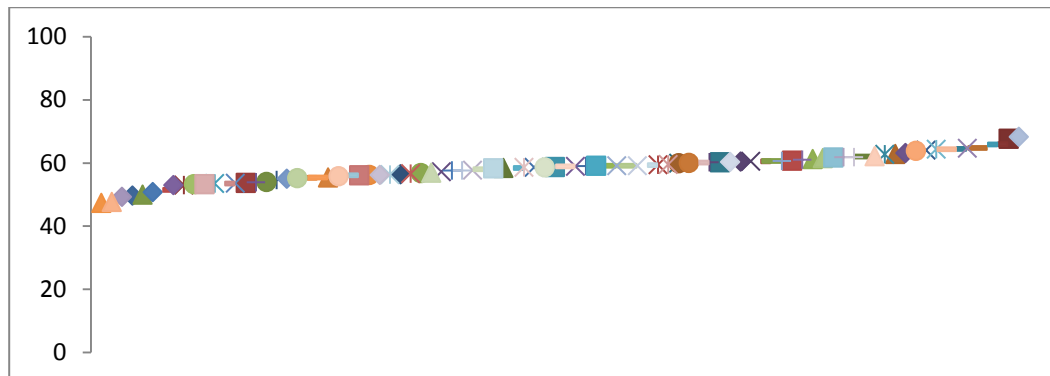
Table 6-2. Variatie in begeleiding naar gelang de actor (cliënt, maatschappelijk werker, OCMW)

	Kans op ontvangen leefloon	Kans op verlies leefloon		
		Ochtend-shift	Dag-contract	Diploma
Gemiddelde kans (onafhankelijk van cliënt, maatschappelijk werker of OCMW)	78.6%	46.2%	60.5%	63.6%
Totale variatie	354.51***	904.85***	844.85***	713.19***
Variatie op niveau gemeente / OCMW	25.31***	54.47***	75.18***	42.72***
Variatie op niveau maatschappelijk werker	154.86***	442.62***	489.12***	373.03***
Variatie op niveau cliënt	174.44***	407.76***	280.55***	297.44***
Relatieve variatie op niveau gemeente / OCMW (in verhouding tot totale variatie)	7.14%	6.02%	8.9%	6.0%

Figure 6-1 geeft het voorgaande resultaat visueel weer. De gemiddelden van de 90 onderzochte gemeenten met betrekking tot de voorspelde kans op sanctie na een weigering van een aanbod dat één dag duurde, liggen relatief dicht bij elkaar. Dit neemt echter niet weg dat sommige gemeenten weldegelijk sterk van elkaar verschillen. Zo

verschillen de uiterste gemeenten (punten helemaal links en rechts op de figuur) tot 20% in de kans dat zij een cliënt na een eerste werkweigering zouden sanctioneren<sup>33</sup>.

Figure 6-1. De gemiddelde kans op het verliezen van het leefloon in 89 Vlaamse OCMW's



*Noot : Case = Cliënt zonder kinderen, zonder moeilijke levensgebeurtenissen, met stabiele huisvesting, die gemotiveerd is om te werken en die een eerste keer een activeringsaanbod weigert dat slechts één dag zou duren*

## 5.2 Verklaringen voor de variatie

Hoewel de belangrijkste vaststelling voorlopig is dat de voorspelde variatie tussen OCMW's relatief beperkt is, tracht ik in deze sectie de toch bestaande variatie te verklaren. Zoals aangegeven in de methodologie sectie, testte ik vijf type variabelen, socio-demografische kenmerken van de gemeente, de politieke context van het OCMW, kenmerken van de maatschappelijk werkers, werkcontext gebonden kenmerken en welvaartsstaat gerelateerde attitudes (zie Table 6-1 voor een overzicht van de variabelen). Table 6-3 geeft de variabelen weer die een (bijna)<sup>34</sup> significant effect hadden.

De socio-demografische kenmerken van de gemeente blijken voorspellende factoren voor alle afhankelijke variabelen. Zo hebben cliënten in grotere gemeenten zowel minder kans op een uitkering als op sancties (Table 6-3, Rij 5). De welvaart van een gemeente lijkt ook een invloed te hebben. Hoe welvarender de gemeente, hoe minder kans op sanctie (Rij 10), dit onafhankelijk van de grootte van de gemeente of het aantal leefloongerechtigden. Een sterk verminderend effect op het sanctiebeleid is de

<sup>33</sup> De resultaten voor de andere afhankelijke variabelen zijn gelijkaardig. De figuren hiervan zijn op te vragen bij de auteur.

<sup>34</sup> Alle variabelen die bij uitsluiting voor een modelverslechtering zorgden werden behouden in het model. Twee controlevariabelen werden sowieso behouden: grootte van de gemeente en relatief aantal leefloongerechtigden.

aanwezigheid van een VDAB in de eigen gemeente (Rij 12). Het is vooralsnog moeilijk om te verklaren waarmee dit effect te maken heeft. Verder zijn er een aantal effecten die maar bij één van de afhankelijke variabelen voorkomen: meer kans op sanctie in gemeenten met een hoger percentage werklozen (Rij 8), met meer geactiveerde leefloongerechtigden (Rij 7), met meer inwoners met een vreemde herkomst (Rij 9) of met een meer denkende kinderopvang (Rij 11). Ten slotte, is er meer kans op een uitkering in gemeenten met veel leefloongerechtigden (Rij 6). Hier speelt er wellicht een omgekeerd causaal effect.

Ook de politieke partij van de voorzitter van de OCMW-raad lijkt een sterk bepalende factor. Maatschappelijk werkers in gemeenten met N-VA-voorzitters voorspellen gelijkaardige cliënten minder vaak een uitkering toe te kennen en vaker te sanctioneren dan hun collega's in gemeenten met een SP.a-bestuur (Rij 16). Met betrekking tot twee van de afhankelijke variabelen is er ook een verschil tussen gemeenten met N-VA en Open VLD-voorzitter (Rij 15).

Van de kenmerken van de (deelnemende) maatschappelijk werkers lijkt vooral de gemiddelde leeftijd in het OCMW doorslaggevend. Hoe ouder het deelnemende team maatschappelijk werkers of hoe meer werkervaring zij hebben, hoe vaker zij voorspellen een uitkering te weigeren en te sanctioneren (Rij 19 en 23). Andere socio-demografische kenmerken van de maatschappelijk werkers hebben slechts op één van de variabelen een effect. Zo lijkt het dat teams met meer mannen meer sancties voorspellen (Rij 20) en teams met meer maatschappelijk werkers met kinderen minder (Rij 21).

Er is een beperkte invloed van de – door de respondenten ingeschatte – beslissingsstructuur binnen het OCMW op de voorspelde begeleidingen. Het lijkt erop dat in OCMW's waar de raad sterker betrokken is bij beslissingen over uitkeringen en/of over activeringstrajecten, dat de maatschappelijk werkers daar vaker inschatten dat het niet aannemen van een job tot een sanctie leidt (Rij 31 en 32). De invloed van de betrokkenheid van de maatschappelijk werkers, de hoofdmaatschappelijk werkers en het team sterk bij deze beslissingen is minder duidelijk en slechts voor enkele van de afhankelijke variabelen doorslaggevend.

De organisatie van het OCMW lijkt dan wel weer erg belangrijk. De maatschappelijk werkers in sterk niet-gespecialiseerde teams (= de meeste maatschappelijk werkers oefenen drie of meer verschillende taken uit) voorspellen minder te sanctioneren (Rij 26). Verder lijkt een hoge werkdruk (verantwoordelijkheid voor een hoog aantal cliënten) gelinkt aan meer toekenningen en minder sancties (Rij 24). Dit is niet te verklaren door een hoog aantal gerechtigden door veel toekenningen en sancties, want daarvoor wordt gecontroleerd. Mogelijk betekent een hoge werkdruk dat er niet

voldoende tijd is voor de opvolging die nodig is om sanctiedossiers te initiëren en te documenteren.

De attitudes van de maatschappelijk werkers en van de hoofdmaatschappelijk werker hebben een duidelijke invloed. Elk van de attitude-variabelen heeft een invloed op één van de afhankelijke variabelen. In OCMW's waar de maatschappelijk werkers gemiddeld negatiever zijn over het functioneren van de welvaartsstaat (Rij 34), meer controle en sanctionering van uitkeringsgerechtigden willen (Rij 35) of meer misbruik van uitkeringen vermoeden (Rij 36), worden minder toekenningen of meer sancties voorspelt. In een OCMW waar een hoofdmaatschappelijk werker veel waarde hecht aan controle en sancties, voorspelt men meer sancties (Rij 38). Een moeilijk te verklaren effect is het negatieve effect van het vermoeden van de hoofd maatschappelijk werker dat uitkeringen (ziekte-uitkering, werkloosheidsuitkering en leefloon) misbruikt worden op de kans op sanctie (Rij 39).



Table 6-3. Effecten van verschillende gemeente en OCMW-kenmerken

		Kans op ontvangen leefloon	Kans op verlies leefloon		
			Ochtendshift	Dagcontract	Diploma
<b>OCMW of gemeentekennmerken</b>					
<i>Baseline</i>		76.92***	79.45***	32.17*	71.75***
<b>Socio-demografische kenmerken</b>					
Aantal inwoners		-	-0.00008***	-0.00002*	-0.00001
% leefloongerechtigden		0.00001**			
% leefloongerechtigden in activering		2.99***	2.37	2.85	1.0
% uitkeringsgerechtigd werklozen			36.48*		
% inwoners met vreemde herkomst			2.44*		
Welvaart gemeente				0.26*	
Kinderopvang			-0.38*		-0.27*
VDAB in de gemeente			0.34*		
<b>Politieke kenmerken</b>			-10.47***	-10.54***	-9.66***
Partij voorzitter OCMW-raad					
(ref = N-VA)					
	CD&V	3.92	-5.43	-6.67	-2.47
	Open VLD	1.83	-2.33	-16.66***	-10.33*
	SP.a	7.45**	-19.05***	-13.94***	-7.6
	Andere	2.35	-16.37***	-7.11	-5.07
<b>Kenmerken maatschappelijk werkers</b>					
Gemiddelde leeftijd respondenten		-0.36*		0.75**	
Percentage mannen in het team			0.15		
Percentage MA's met kinderen			-0.11	-18.97**	
<b>Kenmerken werkcontext</b>					
Gemiddelde werkervaring respondenten per OCMW					0.55*
Gemiddelde werkdruk per OCMW		5.33***	-5.81*		
Mate van specialisatie (ref = vooral 1 taak)					
	2 taken		-2.44	-4.2	-0.18
	3 of meer taken		-19.16***	-11.23*	-10.31*
	Combinatie		-7.3	-6.12	0.38
Invloed van MA, hoofdMA en team op beslissingen inzake uitkeringen		-2.62*	5.36		
Invloed van MA, hoofdMA en team op beslissingen inzake activering			-11.33		
Invloed van OCMW-raad op beslissingen uitkeringen			4.78*		
Invloed van OCMW-raad op beslissingen activering				3.38**	3.62***
Interne richtlijn over taalbeleid			6.64*		
<b>Opinie variabelen</b>					
Algemene attitude MA's over functioneren welvaartsstaat		-6.14*			
Algemene attitude MA's over nood aan controle			19.44**	19.34***	
Algemene attitude MA's over misbruik uitkeringen					14.4*
Attitude HMA over functioneren welvaartsstaat			6.01		
Attitude HMA over nood aan controle				6.67	10.32*
Attitude HMA over misbruik uitkeringen					-13.84*

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

## 6. Naar een meer administratieve toekenning van het leefloon?

In 1974 werden de gemeentelijk georganiseerde COO's verantwoordelijk voor de verstrekking van het bestaansminimum (nu leefloon). De keuze voor deze bestaande en gedecentraliseerde structuren was niet zo vreemd. Het betrof enerzijds het verderzetten, maar tevens formeler verankeren van een traditie, anderzijds betrof het aantal gerechtigden zo een kleine groep (ongeveer 10.000) dat het absurd leek om hiervoor een aparte nationaal georganiseerde dienst op te richten. Tevens is het niet

vreemd dat voor deze uitkering, die de laatst mogelijk uitkering is waar mensen beroep op kunnen doen (als ze al hun andere rechten uitgeput hebben), gekozen werd voor diensten die veel ruimte kregen om de toekenning aan te passen aan de specifieke noden van de cliënt en die bij die toekenning rekening konden houden met de lokale context (aanwezige diensten, bestaande problemen).

De situatie veranderde sindsdien sterk. Het aantal gerechtigden steeg tot meer dan 120.000 unieke leefloongerechtigden per maand (Van Mechelen e.a., 2016). Het lijkt steeds minder om een restgroep te gaan waarvan de kenmerken zeer divers zijn. In de huidige leefloonpopulatie zijn een aantal grote groepen terug te vinden, die hoewel ze een sterk omlijnd profiel hebben, moeilijk op te vangen zijn door de sociale zekerheid zoals ze nu georganiseerd is. Een eerste groep zijn de (laaggeschoolde) vrouwen (en soms mannen) die nooit of onvoldoende werkten om recht te hebben op sociale zekerheid en na de steeds frequentere echtscheidingen of relatiebreuken afhankelijk worden van het leefloon. Een tweede groep zijn de studenten die gehoor geven aan het groeiende belang van (extra) diploma's, maar die geen beroep kunnen doen op hun ouders om hun studie te ondersteunen en die omwille van hun jonge leeftijd en onbestaande of beperkte werkgeschiedenis nog niet voldoende bijdroegen aan de sociale zekerheid. Een derde groep zijn laaggeschoolde mensen die moeilijk werk vinden dat aangepast is aan hun mogelijkheden. Een vierde groep zijn vluchtelingen die nog niet (voldoende) werkten in België. Een vijfde de pas afgestudeerde jongeren die nog geen recht hebben op een inschakelingsuitkering (Van Mechelen e.a., 2016). Dit zijn slechts een aantal voorbeelden van groeiende groepen mensen die geen beroep kunnen doen op de Belgische sociale zekerheid die gebaseerd is op eerdere werkgeschiedenis en lastenbijdragen. Verder is er binnen de implementatie van de sociale zekerheidsuitkeringen ook een tendens naar verstrenging van de toepassing van de voorwaarden (en dus ook sancties), waardoor meer afgewezen of gestrafte cliënten (tijdelijk) afhankelijk raken van het leefloon.

Dit betekent dat de financiële dienstverlening aan kwetsbare groepen in groeiende mate gedecentraliseerd raakt. Dit gebeurt echter niet vanuit de idealen die voorstanders van decentralisatie naar voor schuiven, namelijk de kans tot het aanpassen aan lokale mogelijkheden en noden en tot het medebewindschap van lokale overheden (zie inleiding). De tendens is daarentegen bepaald door socio-demografische veranderingen en door een veranderde kijk op wie recht heeft op een werkloosheids- of ziekte-uitkering. Dit niet intentioneel decentraliseren van de sociale zekerheid kadert binnen de problematiek van de passieve subsidiariteit, waarbij groepen cliënten onder de verantwoordelijkheid van lokale overheden geplaatst worden, zonder dat daar duidelijke budgetten of beleidsdoelen tegenover staan (Kazepov, 2005).

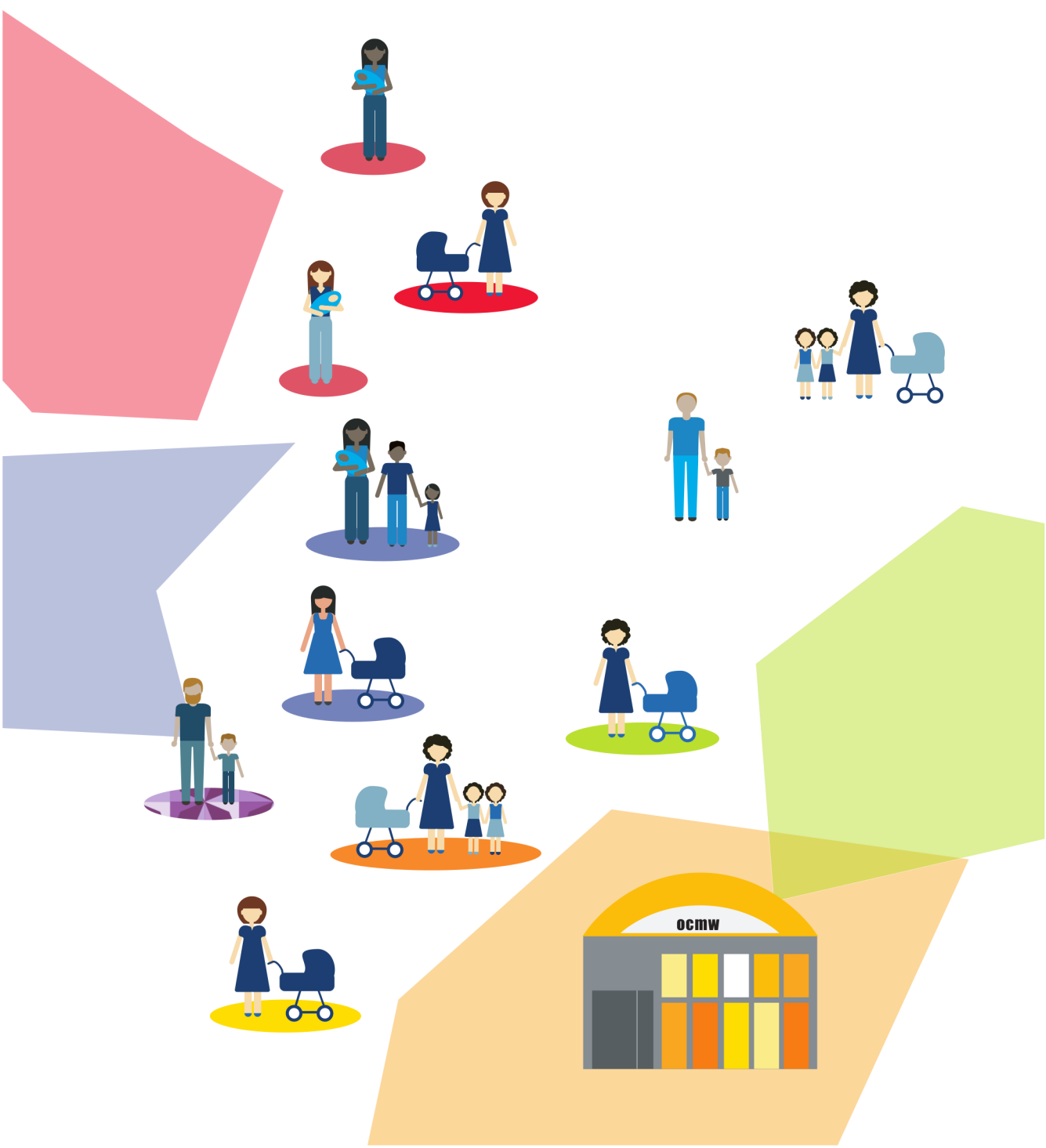
Het is in dit licht belangrijk na te gaan of er grote variatie is tussen gemeenten en indien zo waardoor deze variatie verklaard kan worden. De vraag is dan of de decentralisatie vooral zichtbaar wordt in variatie omwille van de socio-demografische context (eerste motivatie voor decentralisatie) of omwille van specifieke beleidsvoorkeuren van lokale overheden (tweede motivatie voor decentralisatie) of dat integendeel andere factoren die OCMW-gebonden en eerder toevallig zijn de variatie verklaren. Een eerste vaststelling uit het onderzoek voorgesteld in dit hoofdstuk, is dat de variatie tussen de gemeenten inzake de voorspellingen omtrent het uitkeren van een leefloon en het uitvaardigen van sancties, bijzonder klein was (zie Table 6-2). Dit neemt echter niet weg dat sommige gemeenten wel tot 20% van elkaar verschilden in de voorspelde kans op een uitkering of sanctie (zie Figure 6-1). Deze duidelijke verschillen tussen 10 tot 20 van de 90 onderzochte gemeenten, maakte het nagaan van de verklarende factoren relevant.

Zoals verwacht, speelt de context van de gemeente een rol (grootte, welvaart...). Het is echter moeilijk om de gevonden effecten als bewuste keuzes te interpreteren, namelijk keuzes met de bedoeling het OCMW-beleid aan te passen aan de lokale context, zodat de doelgroep of de ruimere gemeenschap er voordelen van ondervindt. Eén van de effecten, roept extra vragen op, met name de verlaging van de kans op sanctie in gemeenten met een VDAB-kantoor op het grondgebied. Mogelijk gaat het hier om een bewuste aanpassing aan de lokale situatie, maar het is vooralsnog moeilijk om de rationale erachter te interpreteren. Het belang van het zeggenschap van lokale overheden of bevolkingsgroepen is een tweede argument voor decentralisatie (Fording e.a., 2007; Soss e.a., 2011). Deze politieke invloed wordt duidelijk in het sterke effect van de politieke partij van de voorzitter van de raad. Daarnaast zijn er nog een aantal factoren die niet rechtstreeks met de lokale context en de beleidsvoorkeuren van de lokale overheid te maken hebben. Deels is de implementatie van het bijstandsbeleid afhankelijk van de samenstelling van het OCMW-team (bv. leeftijd en werkervaring). Belangrijker lijkt het effect van de algemene welvaartstaat attitudes van het team en van de het teamhoofd. Dit kan wijzen op de sanctiegevoeligheid van een gemeente (Holzner e.a., 2009). De kans bestaat dat de teamleden mede omwille van deze attitudes gekozen worden en dat dit dus eveneens een resultaat is van de beleidsvoorkeuren van de lokale overheid. Andere elementen waarrond een OCMW bewuste keuzes kan maken is de mate van specialisatie en de werkdruk van de maatschappelijk werkers. Deze organisatiekeuzes stellen OCMW's in staat om te experimenteren en toe te werken naar het implementatiebeleid dat ze voor ogen hebben (zie Table 6-3).

De bovenstaande resultaten stellen ons voor twee potentiële problemen. Ten eerste was de bijstand, met zijn zeer intensieve en betrokken begeleiding nooit bedoeld voor

grote groepen gerechtigden. Ten tweede is het de vraag of, hoewel de variatie tussen OCMW's relatief beperkt is, we de basisdienstverlening aan een groeiende en zeer precare groep afhankelijk willen maken van factoren als politieke voorkeur van de voorzitter, leeftijd van de maatschappelijk werkers, organisatie van het team of opinie van de maatschappelijk werkers. Vermits een aantal van de potentiële cliënten tot welomlijnde groepen behoren (zie opsomming in tweede paragraaf van deze sectie), kan er voor deze groepen misschien gekozen worden voor een – via eenvoudige criteria – meer administratieve toekenning van het recht op een leefloon. De veronderstelling is dat dit de werklast van maatschappelijk werkers en de versnelde afwikkeling van dossiers ten goede zal komen. Het eigenlijke maatschappelijke werk (trajectbegeleiding, activering...) zou dan kunnen starten na deze toekenning of na de weigering op basis van de administratieve data. Logischerwijs zal er immers een groep ontstaan die door de mazen van het net van zulke toekenning op basis van administratieve gegevens vallen. Voor deze groep zou het bijstandssysteem opnieuw kunnen werken zoals het werkte bij aanvang in 1974: sterk afhankelijk van de discretionaire ruimte van gemeenten en maatschappelijk werkers en als laatste vangnet van een groep die elders niet terecht kan.

Conclusions



## Chapter 7. Conclusion: Between legislation and realisation comes implementation

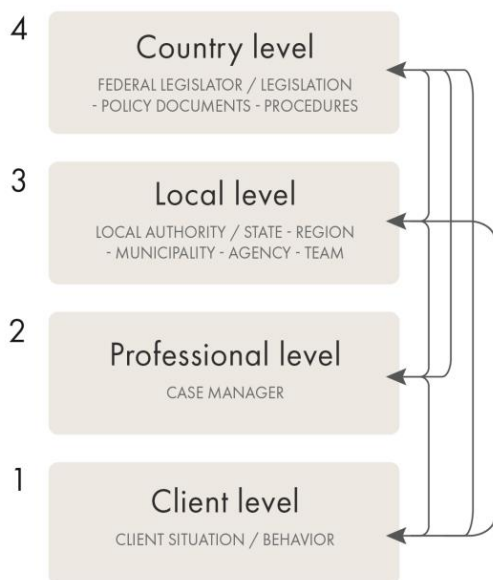
The consensus in social policy literature is that an understanding of how implementation works is key to understanding how policies become actual treatment for clients (Kazepov, 2010; Priem e.a., 2011; Rice, 2013). Surprisingly, the different levels that determine this implementation process (country, local setting, professional and client) are seldom studied in one research project or compared to the expected treatment outcomes for beneficiaries. The result is that the importance of professionals and of local policymaking levels may well be underestimated when discussing the successes or limitations of existing policies. Broadly speaking, policy proposals in social policy research are based on the study of legislation and outcomes, with no mention made of the importance of professionals or local processes in the actual realization of these proposals (Meyers e.a., 1998). At the same time, both the social work literature and public administration research broadly acknowledge the importance of professionals and local initiatives. However, the evaluation of this influence is seldom based on actual outcomes for clients. In this sense, discussing the need for more structuring or for more independence for professionals remains rather abstract.

In this dissertation I used an innovative and purpose-designed survey to determine the importance of the various levels. The experimental survey allowed me to study social assistance treatment for a large number of clients, social workers and social assistance agencies. Undertaking this research has raised a number of concerns concerning the premises of the different research traditions which have previously been taken for granted and on which I elaborate further in this conclusion.

As shown in Figure 7-1, my dissertation focused on four levels that influence conditionality in social policy implementation. Conditionality is defined in legislation (country level), elaborated in more detail in the local setting where the policy is implemented (local level), interpreted and shaped by the professionals who come into contact with clients (professional level) and determined by the situation and the behaviour of the client (client level) (Rice, 2013). In the context of this dissertation, it is impossible to make any statements about the fourth level, as I only studied Belgium, but, in the policy proposals at the end of this concluding chapter, I do refer to legislation. The extent to which the three other levels are important in determining the client's eventual treatment is influenced by several processes that are well described in literature, including federal policymaking, deservingness assessment, decentralization and discretion. Feedback loops also exist between the four levels (see arrows). To the

best of my knowledge, the three levels covered in this dissertation have never been studied in a single research project using one method and one dataset. Thus, to date, it has been impossible to determine which of the levels are more important and what the actual quantitative aspect might be.

Figure 7-1. Overview of the various levels that define conditionality or deservingness<sup>35</sup>



The general assumptions (without going into detail) concerning the influence of the several levels are:

- (1) Client characteristics determine the treatment that clients receive (= to a certain extent similar across professionals and agencies);
- (2) The professional is of decisive importance for treatment (= the use of discretion);
- (3) The level of decentralization of the policy determines treatment variation;

To test these hypotheses, I used Belgium as a case. Social assistance in Belgium is regulated by federal law (RMI law) and implemented by local welfare agencies. One of the conditions for eligibility is the client's willingness to work but the RMI law does not

---

<sup>35</sup> This graph is simply a schematic reproduction of the four levels I was able to account for in my dissertation. Deborah Rice (2013) drew a much more complete picture of implementation reality in her Micro-Institutionalist Theory of Policy Implementation.

detail how social workers should interpret willingness to work. It also refers to reasons for exemption from activation as 'health reasons' or 'fairness reasons', but no clearly described categories are included (Vande Lanotte, 2002a). Clearly, the RMI law offers important leeway to local regulations and preferences and the professional opinions and assessments of social workers. At the local level, each Belgian municipality has a social welfare agency headed by a politically composed non-professional board, which decides – within the framework of the law – on eligibility for financial assistance. Funding provision for the benefits is divided between the federal and local levels. The professionals who come into contact with the clients are always qualified social workers. In sum, the country's decentralization, rather vague descriptions of work willingness in legislation and the explicit recognition of the professional as a qualified professional with important discretion make Belgium an interesting case with which to look at the various determinants of eligibility (and their interaction).

Based on the interesting Belgian case, I raise the following concerns about the premises taken for granted:

- (1) *Client characteristics determine the treatment clients receive.* True, but – with regard to the dependent variables included in my research – to a lesser extent than might be expected from legislation. Client's work willingness affected eligibility only to a limited extent, but became more important later in the activation trajectory. The influence of specific difficult life circumstances on eligibility and sanction decisions was negligible.
- (2) *The professional is of decisive importance for treatment.* True, variation is considerable, and part of it is not random but can be explained by the general welfare state attitudes of professionals.
- (3) *The level of decentralization of the policy determines treatment variation.* This could not be confirmed. Overall, local agencies did not differ extensively from each other, but insofar as they did, organizational characteristics, political majorities and socio-demographic characteristics explained the variation.

In this concluding section, I first present the innovative method used in this dissertation, the factorial survey (Section 1). Next, I elaborate on each of the three research findings and discuss their major contributions to existing academic knowledge (Sections 2-4). Finally, I answer the policy questions arising from these results: (1) are human and civil rights guaranteed in the present implementation process and (2) do central governments need to take action to guarantee the implementation process as intended by policymakers? Policy proposals include a more administrative eligibility test and the reevaluation of discretion in the activation trajectory.



## 1. The factorial survey: an innovative research strategy

Before presenting the actual results of this PhD project, I briefly describe the method that was used and which proved to be a successful tool to combine the above-mentioned focuses into one quantitative study. Factorial survey experiments require each respondent to read and rate stories about individuals or situations (vignettes). Each vignette consists of several factors (e.g. gender) that correspond to certain levels (e.g. female and male). In a factorial experiment, the vignette population consists of all possible combinations of all levels across all factors. From these vignette populations, a sample can be taken and presented to the respondents. After viewing the vignettes, respondents are typically asked to answer a set of rating questions.

Figure 7-2. Overview of the study design



36

In order to study the effect of multiple policy levels on social assistance treatment in Flanders (the northern region of Belgium), I designed, organized and implemented a

---

<sup>36</sup> Note with regard to the population sizes: 3: Flanders has 306 municipalities and thus 306 social assistance agencies; 2: from each of the 90 municipalities that contributed to the survey, I received lists of the professionals involved in social assistance eligibility decisions, a total of 1221 professionals. 3: the vignette population consisted of more than 3 million experimental combinations of the 14 client characteristics.

factorial survey, validated in a pilot study and carried out in a stratified sample of 90 Flemish municipalities among 606 stratified samples of employees involved in eligibility decision-making. I asked the employees to rate nine vignettes concerning hypothetical clients with regard to the likelihood that these clients would receive social assistance benefits, and the likelihood that they would lose their benefits if they refused a job offer (with responses on a seven-point scale). The characteristics of the hypothetical clients included in the vignettes are of particular interest in predicting treatment. I selected the vignette attributes on the basis of existing literature and interviews with professionals, team leaders and academics, keeping the deservingness criteria in mind (Nybom, 2013; Taylor, 2012; van Oorschot e.a., 2017; Wallander, 2009). All of the respondents rated nine vignettes, which varied across 14 client characteristics (see Table 4-1).

Chapter 1 reports on a study in which Peter Goos and I showed that the factorial survey is an excellent tool to combine a concern for the multi-layered reality of policy implementation with concern for the actual treatment of clients. With regard to the first concern, it is possible to distribute vignettes to a large number of respondents in multiple agencies, municipalities and various countries, using modern, inexpensive communication channels. This research setup, provided the resulting data are analysed with multi-level techniques, accounts for the multi-layered structure of the implementation reality. Furthermore, the method allows for the addition of a questionnaire about the respondent and the employer organization. As will be shown in Sections 3 and 4 of this concluding chapter, entering these factors into the model as explanatory variables greatly enriches our understanding of implementation processes.

Regarding the study of real treatment, the advantage of factorial surveys is that the dependent variables consider actual treatment and not the processes as such, as is usually the case in qualitative implementation research. One disadvantage and significant limitation is that factorial surveys use hypothetical rather than real treatment cases, which raises problems concerning external validity. While internal validity, meaning the ability to draw causal conclusions, is high, as the method operates on an experimental basis (Taylor, 2006; Wallander, 2012b) and ensures that no systematic errors occur, external validity, in all its forms, is more problematic. This is discussed further in Chapter 1 (Auspurg & Hinz, 2014). In defence of the factorial survey approach, several scholars believe that responses to experimentally created stories, such as in a factorial survey, are closer to reality than those given in interviews of respondents discussing real cases (Armacost e.a., 1991; Kirwan & Currey, 1984; Peabody, Luck, Glassman, Dresselhaus, & Lee, 2000b).

## **2. Implementing the work willingness condition**

One of the main questions in this dissertation is how behavioural conditions such as work willingness are implemented. Existing studies have argued that with the growing importance of activation, behavioural conditions are included in all European social assistance laws. Yet, relatively little is known about how these conditions are implemented, as most related studies are based on legislation itself or on datasets with client outcomes, which lack information about the work willingness of these clients. In this dissertation project, it was possible to combine awareness of existing legislation with information on the work willingness of hypothetical clients and with predicted outcomes. The main conclusion is that work conditionality is not decisive at the start of the process when eligibility issues are first considered, but becomes very important in the activation trajectory. Reasons to exempt clients from activation are not often used.

### **2.1 The work willingness condition**

The likelihood that the hypothetical clients would receive benefits ranged (without extremes) from 33% to 100% (see Chapter 4 and Chapter 3). Literature states that professionals differentiate between high-compassion and low-compassion clients (Scott, 1997), or between more deserving and less deserving clients (van Oorschot e.a., 2017), or between clients with different character types, such as ‘troublemaker’ and ‘nice lady’ (Maynard-Moody & Musheno, 2003), or between clients in specific situations, for example, in residential or non-residential care (Murray, 2006). To test this, the data were regressed on the client characteristics that could be considered to lead professionals in their assessments (multi-level model with all client characteristics included (no interaction terms) – see Chapter 4). Surprisingly, the difference in eligibility between two hypothetical clients who could be considered to be at either ends of the compassion and deservingness scales was rather small. A client who was not motivated to work was, on average, 79% likely to receive a benefit. For a client with a sick child this likelihood was 88%. Thus, the client characteristics included in the vignettes (Table 4-1) did not influence the eligibility likelihood to a large extent. This led to the conjecture that clients are considered eligible if they are in financial need and if they meet the objective citizen criteria (age, nationality, etc.) (see Chapter 4). Of course, I was unable to prove this with the available data, as these objective criteria (financial need, age, legally resident in Belgium) were not included as experimental factors. Furthermore, the effect of the lack of motivation to work is possibly stronger in real life, when a professional is confronted with a non motivated client. Here, I clash against the limits of standardized cases.

While the lack of motivation to work had almost no effect on eligibility, refusing job offers was clearly considered to be a violation of the work willingness condition. Clients with 'standard' characteristics (motivated to work, no children, no mental health problems, no addictions, etc.) who refused a job offer for the first time had a 52% to 67% risk of losing their benefit. This risk increased to 70-82% if they refused three different offers (Chapter 3).

These two results, namely the neglect of the work willingness condition to assess eligibility and the high sanction likelihood if the client refused a job offer, seem to contradict each other. Two explanations are possible. The first is that professionals see differences between the beginning of the trajectory and the trajectory itself. They might feel it is unacceptable to refuse social assistance based on their first contact with the client. This could be considered an intended or unintended refusal by professionals to implement legislation as it is designed. This is in line with the view of Tabin and Perriard (2016) that policies only have real effects if the professionals implementing them accept the ideology that lies behind them (Moller & Stone, 2013). It would be possible to conclude that professionals in Belgium do not appear to accept willingness to work as an exclusion criterion at the beginning of a social assistance trajectory. The difference between the beginning of the trajectory and the trajectory itself, however, may be practical rather than ideological. Perhaps professionals find it difficult to assess the work willingness condition reliably in the short period between the client claiming the benefit and the eligibility decision. Thus, it is perhaps for practical reasons that work willingness is not used as an exclusion criterion. This observation ties in with the finding of Meyers, Glaser and Mac Donald (1998) that US welfare workers rarely talk about the importance of work with their clients during the eligibility test, which appears to be both practically and ideologically motivated. Communication itself and communication about other elements than orientation towards work were considered more important (Meyers e.a., 1998).

## 2.2 Exemptions from work willingness

Another striking element was the limited use of fairness reasons to exempt clients from activation requirements. This possibility is included in legislation, but with almost no further specification. The question in Chapter 3 (sanctions) and Chapter 4 (eligibility) was which client characteristics professionals and agencies consider to be valuable fairness reasons. Parenthood, mental health problems and difficult life experiences only slightly mitigated the negative effect of not being motivated to work on eligibility. However, as the effect of not being motivated was itself rather small, this is not particularly remarkable (Chapter 4).

With regard to the likelihood of clients being sanctioned if they refuse an activation offer, only one client characteristic had a noticeable effect, namely parenthood, and only with regard to the refusal of a certain type of job offer, namely a job that started early in the morning. Other client circumstances such as mental health problems (depression or intellectual disability), addiction problems (beginning or severe), housing situation (unstable or homeless) or adverse life experiences (loss of a child or abusive upbringing) were significant predictors, but decreased the sanction likelihood by a maximum of only 4% (Chapter 3).

It is rather surprising that, while the legislation clearly leaves room to invoke several life circumstances as exemption criteria, these criteria do not seem to be used often, and certainly not clearly and uniformly across professionals and agencies. This suggests a rather administrative application of eligibility and sanction legislation, as was observed by Lens (2008) and Meyers, Laser and Mac Donalds (1998) when studying eligibility and sanction decisions in the US. Lens concludes that sanction treatment is not individualized, although legislation does provide for this possibility, and speaks of an *eligibility-compliance culture* to frame “depersonalised, mechanic and routine” decision-making (Lens, 2006, p. 576).

To conclude: my study – like several other previous studies – rejected the idea, previously taken for granted, that policies are decided upon at national level and implemented as intended by the legislators at lower levels (Tabin & Perriard, 2016). I showed quantitatively, for the first time, that in reality, several client characteristics are assessed in a different fasion than would be expected based on a straightforward application of the legislation. The importance of people and institutions at the lower levels in this implementation process is documented in the two next sections about the professional and the social assistance agency.

### **3. The importance of the professional level**

The most surprising finding in this dissertation project was the variation in eligibility and sanction treatment among professionals working for the same agency. This indicates a high use of discretion. There is growing consensus in social work literature that discretion is an inevitable aspect of the social work profession. Focusing on social assistance, one of the main explanations is that qualitative programmes should take into account the individual needs, strengths and weaknesses of clients. Targeting and adjusting treatment to the client’s needs increases the meaningfulness of the policy behind the treatment (Tummers & Bekkers, 2014). Furthermore, the situations in which clients ask for assistance are often so heterogeneous that it is impossible to provide specific legislation for all cases (Nothdurfter, 2016; Thoren, 2008b). In addition, social

assistance delivery entails more than just handing out money based on guidelines prescribed in legislation. Financial support is accompanied by administrative and psychosocial support, as well as support for reintegration, which all involve decisions that must be made by professionals. As the potential impact of such decisions is high, professionals must be professionals with the appropriate educational background and work ethic. The literature on professionalism suggests that employees are only 'professional' if they are assigned at least some decision-making power (Evetts, 2002; Handler, 2003; Hupe & Hill, 2007; Lipsky, 1980). If they are denied this decision-making freedom, they are merely executors of prescribed rules, much like clerks (Tummers & Bekkers, 2014).

However, my study is the first to show the magnitude of the use of discretion quantitatively. A three-level null model (client, professional, municipality) was used to calculate the unexplained variance at each of the levels studied. This variance was as substantial at the professional level as the variation found at the client level (see Chapter 5). In Chapter 5, it was argued that the likelihood of receiving social assistance when a client is willing to work and has no 'special' characteristics (such as parenthood or health problems) ranges (excluding outliers) from 49% to 100%, depending on the professional completing the survey. For the same client, the likelihood of keeping social assistance after refusing one, two or three activation offers was between 4% and 80%, with the only difference being the professional predicting the sanction likelihood. Below, I elaborate on these findings. First, I test the hypothesis that behavioural conditions add to the use of discretion. Second, I answer the question as to whether this use of discretion can be considered a normal human process, and third, I look for means of managing the use of discretion.

### 3.1 Legislation vagueness increases discretion

The growing importance of activation and reintegration in the labour market has been accompanied by the introduction of behavioural conditions for eligibility (Clasen & Clegg, 2007; Standing, 2005; Van Kersbergen & Hemerijck, 2012). Such conditions are both difficult to prescribe in legislation and to assess in practice (Meyers e.a., 1998). Therefore, implementation variation is likely to remain or to increase, even with regard to legislation that is designed to be more prescriptive, as a result of the introduction of behavioural conditions instead of categorical conditions (e.g. lone parent) or means tests alone (Fording e.a., 2007; Moller & Stone, 2013). It goes without saying that it is impossible to prove a direct causal effect, as no counterexample of a country that does not have behavioural conditions in its legislation was analysed. However, two observations are relevant. First, predicted treatment variation was higher with regard to

client characteristics that deviate from expectations in legislation (see Chapter 5). Second, some client characteristics that should not have any effect on eligibility proved significant, meaning that discretion transcends other characteristics.

First, the 'willingness to work' eligibility condition in Belgian legislation and the fairness and health reasons to exempt clients from activation requirements are typical examples of vague requirements. It is left to the agencies and the professionals to assess these conditions. The effect of such an assessment procedure is shown in Chapter 5. The differences between professionals grew bigger with regard to characteristics which tend to decrease eligibility. This means that this decrease is not common to all professionals. Professionals agree with each other more on client situations that are typically seen as high eligibility cases and disagree more if eligibility is in doubt. The assessment is even more subtle with regard to sanction treatment. The consensus in sanction likelihood was lower with regard to client characteristics that could be considered as things that the client had personal control over (addiction, housing situation, depression without consultation of the doctor). The variety in interpretations of what 'in control of the personal situation' means, adds to our understanding of the complexity of deservingness assessment.

Second, the PhD project found that some client characteristics had a significant influence on eligibility, though there is no mention of these characteristics in legislation, such as gender, the combination of second-generation migrant status and language proficiency, having a Bachelor's degree, housing situation and the desire to study (Chapter 4). A possible explanation for these rather unexpected results could be that because professionals are constantly assessing difficult situations, they use their discretion more broadly than they would do if they only needed to tick boxes.

### 3.2 Variation as an inevitable result of the human process?

This observation adds to the conviction that the use of discretion or treatment variation is an inevitable human process. For as long as offering or refusing benefits remains unautomated, based on objective data, variation will occur. However, two results stand out to contradict the notion that it is a 'normal human process'. First, there is the magnitude. As mentioned above, the variance found at the social worker's level was as substantial as the variation found at the client level, and the differences between (non-outlier) respondent-specific means were large. It is difficult to consider such huge differences as a random effect, caused by coincidence, and acceptable within the boundaries of what is human.

Second, the effects were not in fact random. Several professional characteristics determine the predicted likelihood that clients will receive or lose their benefits if they

appear to be unwilling to work. The seniority of the professional influences eligibility positively and sanction treatment negatively (De Wilde, 2016; De Wilde & Marchal, 2018). Other remarkable predictors include the general attitudes of the professionals towards the welfare state. Professionals who had lower trust in the usefulness of the system, and who thought that control and punishment of beneficiaries should be higher, less often considered clients eligible and more often predicted sanctions (De Wilde, 2016; De Wilde & Marchal, 2018). These results mean that the personality of the professional is of decisive importance for clients (Chapter 3). However, I found no support for the idea that certain professional characteristics, such as seniority, affected the use of discretion itself. No groups of professionals (e.g. in a certain age group or at a certain level of seniority) displayed more within-group variation than other groups (Chapter 5). However, possibly the respondent sample was too small to successfully perform the analysis I used in this chapter (modelling heteroscedasticity).

### 3.3 Can discretion be managed?

A logical question, based on the above finding that a professional's personality influences the treatment a client receives, is whether discretion can be managed or limited. Two possible means of managing discretion are, first, to introduce clearer and more detailed regulations and, second, to introduce more control mechanisms, modules and procedures that increase the standardization of implementation. The issue is seldom tested, and if so, results are not straightforward. An attempt to standardize eligibility and sanction treatment in Nordic countries did not seem to be successful as municipalities or professionals kept their discretion or kept on using it contrary to the central guidelines (Minas e.a., 2014). Researchers often refer to the decrease in discretion that is assumed to have occurred due to the increasing focus on managerial control, a result of evolutions in information and communication technology (ICT). These evolutions make it possible to monitor the extent to which legislation is put into practice more closely (Bovens & Zouridis, 2002; Buffat, 2015; Dunleavy e.a., 2006). However, other authors suggest that discretion is inevitable and highly necessary to translate policies into effective and client-tailored practices (Evans & Harris, 2004; Evetts, 2002; Handler, 2003; Hupe & Hill, 2007; Nothdurfter & Hermans, 2018; Thomann e.a., 2018; Thoren, 2008b; Tummers & Bekkers, 2014) and thus remains stable, even if additional regulations attempt to constrain discretion (Evans, 2011). Others state that the use of discretion is even increasing due to changes in policy focuses (Heidenreich & Aurich-Berheide, 2014; Jessen & Tufte, 2014; Künzel, 2012; van Berkel, 2010a, 2010b; Watkins-Hayes, 2009). To a large extent, these statements are based on in-depth case studies of organizations and on interviews with professionals and managers. However, I was able to test the links between inter-professional variation (discretion), predicted



outcomes and organizational features. In what follows, I describe the results with regard to the possibility of managing variation between professionals. It is important to note that these are initial, exploratory results – first, because Belgium has no strong tradition in managerial techniques focused on the control of professionals, and second, because the items used in my study are only proxies for the actual organizational settings in agencies.

Chapter 5 showed that none of the possible organizational characteristics that could have an impact on inter-professional variation, such as internal guidelines concerning some of the eligibility conditions, the importance of team involvement in decision-making or the involvement of the board in eligibility and activation considerations, proved significant. This means that, based on my – with regard to this item – limited data, organizational structure neither eliminates or encourages the use of discretion. However, professionals that experienced themselves as important in decision-making displayed more inter-professional variation. Thus, it is possible that a certain element external to the professional might lead to uniformity.

Furthermore, with regard to the effect of detailed regulations on the use of discretion, it is interesting to study the two cases that were detailed in law or in ministerial letters accompanying the law. The first case concerns clients undertaking full-time education. This is the only concrete reason for exemption from the work willingness condition that is included in Belgian legislation (Vande Lanotte, 2002b). The second case concerns a client with a severe drug addiction. This situation is included in a circular instruction of a ministerial letter explaining the new social assistance law of 2002 as an example of health reasons that exempt clients from the activation condition (Vande Lanotte, 2002a).

With regard to clients involved in full-time education, the limits of such a specified condition become apparent in the circular instructions of a ministerial letter that was sent to all social assistance agencies to explain this new element of the legislation (Dupont, 2004). First, the ministerial letter is extensive, which shows that adding a concrete prescription does not necessarily simplify things. Second, the document includes quite vague descriptions, such as: ‘the study needs to increase the labour market opportunities of the client’. It is clear that this condition is again subject to the assessment and thus the discretion of the professional. This is also shown in my data. Against all expectations, a client who wanted to study was less likely to receive benefits than a client who wanted to work. This likelihood was not effected by the level of education that the client already had (primary education, secondary education or a

Bachelor's degree). Furthermore, inter-professional variation was as high for this client characteristic as for the characteristic 'wants to work' (Chapter 4).

The same was true for the assessment of addiction problems. A client with a severe addiction was 13% less likely to be directed to the labour market, compared to a person with no addiction problems. Although the activation likelihood remained 73%, this decrease is substantial, which might mean that agencies and professionals follow the recommendation given by the government. However, the differences between professionals were large, with likelihoods from 41% to 100% likelihoods (extremes excluded).

Lastly, as shown in the previous section, personal attitudes are of decisive importance in the implementation process. The study undertaken with Bart Meuleman and Koen Abts (reported in Chapter 2) demonstrated that the welfare state attitudes of social workers do not reflect the welfare state attitudes of the general public. This was partly explained by socio-demographic and attitude characteristics, but there were other differences that could not be explained in these terms. We hypothesized that experience of specific problems associated with poverty and/or self-interest, such as feeling responsible for the system in which one is working, produces a more positive view of the welfare state system and of welfare state beneficiaries. It is doubtful that more strict regulations and more management control would significantly alter the influence of particular welfare state attitudes on actual treatment.

This dissertation project is too limited to conclude that discretion cannot be managed at all, but insofar as I was able to investigate the issue, I found no indication that more detailed legislation or more elements typical of standardization would reduce the use of discretion (Minas e.a., 2014).

### 3.4 Awareness of the use of discretion

Awareness of the high use of discretion appears to be limited. First, professionals see themselves as less important than the agency board in deciding what happens with eligibility issues (60% agreement with their own importance in eligibility issues compared to 89% agreement with the importance of the board). Second, in 2018, I distributed a post-questionnaire on the existence of internal guidelines in 2015 among the 90 municipalities included in the research project. Only 13% of the agencies reported having issued internal guidelines to regulate the legal concept 'work willingness'. In a follow-up question, the agencies were asked to indicate whether the professionals knew how to interpret the criterion. Almost all agency respondents answered 'yes' and referred to extensive case-by-case intervention in the team. Only 10% of the agencies stated that the professionals were quite free to assess client

characteristic and behaviour. A total of 62% of the agencies referred to the team as the central actor in interpreting work willingness issues. In contrast, the 600 respondents in the main study had a low appreciation of the role of the team in eligibility and activation decisions. The respondents assessed the importance of the team to be only 42% (0% = not important at all – 100% = very important). While it is difficult to evaluate the precise level of unawareness of the importance of discretion based on this study, it is fair to say that the awareness was limited. To be able to use discretion as a positive tool, it seems crucial to raise awareness of treatment choices.

#### **4. The relative importance of local policymaking**

##### **4.1 Low implementation variation across municipalities**

The decentralized assessment of eligibility criteria and activation requirements appears to be a general characteristic of contemporary social assistance schemes. While decentralized assessment is self-evidently relevant in other policy domains (Brodkin, 2011), Sabatinelli (2010) has argued that it is especially relevant at the intersection of activation and social assistance. The expectation is that – as far as national or federal legislation allows – decentralized organization will result in local entity-specific policy accents, and thus in treatment variation. However, the main finding in this dissertation with regard to variation across municipalities is that this variation was rather limited. Only 6% to 9% of all variation in eligibility or sanction predictions was situated at the local level (Chapter 3 and Chapter 5).

The magnitude of the variation across different local entities has been analysed before (Forssen, 1998; Keiser, 1999). A number of studies report finding large variation (Carpentier, 2016; Carpentier e.a., 2014; Fording e.a., 2007; Keiser e.a., 2004; Monnat, 2010; Soss e.a., 2011; van Oorschot, 2000). There are five possible explanations for this difference in results. First, the entities investigated in these studies were mostly states, not municipalities, with some exceptions (Carpentier, 2016; Minas e.a., 2014). This might influence the magnitude of the variation. Second, the multi-level analysis I used might distort the actual results. In multi-level analyses, distribution around the level mean is considered random. This means that extreme items that have few measurements at the lower levels are shrunk to the mean. One third of all municipalities in the survey had only two collaborating respondents. The deviation from the mean for these municipalities may – due to the homoscedasticity assumption – have been calculated too conservatively. It is not possible to establish the extent to which this miscalculation could have influenced the results. Third, the dependent variables investigated in this dissertation are, of course, limited. The conclusions of this study may not hold for issues such as variation in types of activation programmes, additional

financial support, duration of benefit dependence or other client situations not included in the vignettes.

The last two explanations, however, show the uniqueness of the method used in this dissertation. The fourth possible reason for my results, concerns the way in which variation is traditionally measured. Most studies compare the two most extreme local entities and conclude – without using a clear benchmark – that variation is high (Carpentier e.a., 2014; Fording e.a., 2007; Keiser e.a., 2004). If I were to base my conclusion on such a parameter, the difference in mean eligibility likelihoods between the most extreme municipalities in the survey would be 15% to 19%, which can be considered high<sup>37</sup>. Focusing on the extremes, however, provides no insight into the general variation across agencies. Other studies argue that local variation is high because there are parameters at that level that have a significant effect on the dependent variable (Carpentier e.a., 2014). These studies do not compare the z-scores of the variables across levels (local level and client level) and, thus, do not account for the relative importance of these parameters. In my project, multi-level null models were used to measure the unexplained variance at each level. This, and the calculation of the variance partition coefficient, provides a clear idea of the importance of the variation at the local level, while avoiding the need to rely on artificial benchmarks.

Fifth, these studies were not able to account for the variation in the use of discretion across professionals from one agency (see previous section). If other studies used multi-level models, these were mostly two-level models (agency/municipality and client) (Carpentier e.a., 2014; Minas e.a., 2014). However, if one crucial hierarchal level (the professional) is neglected in the analysis, the variation found at that level is spread over the two other levels (agency or client) and variation estimates are biased. If I repeated the multi-level analyses in my study without including the individual professional level, the variation situated at the municipality level would be estimated at 20%, which is considerably higher and similar to the intra-municipality variation results of Minas and colleagues (2014) in Denmark, Finland, Norway and Sweden. Thus, part of the variation that is situated at the municipality level in previous studies is actually variation between professionals in the same agency.

---

<sup>37</sup> The difference between the two different municipalities would be even higher if I did not control for client characteristics, as I did in this analysis, and if I did not use a multi-level model, whose main feature is that it considers distribution around the mean at each level to be random. As a result, the results for level-specific elements (e.g. municipalities) with low frequencies at a lower level (e.g. professionals) would be shrunk to the mean.

#### 4.2 Does decentralization live up to its promises?

In several countries, policymakers are calling for more local decision-making power with regard to social policy. Belgium, with its long tradition of semi-decentralized social assistance delivery and policymaking, can be used as a case study to determine whether the decentralized organization of social assistance is successful in the sense that the explicit goals of decentralization are reflected in outcome results. The results are shown below, though it should be acknowledged that treatment variation at the local level is actually very low (see previous section). Advocates of the decentralized organization of social policy services emphasize three important goals: (1) the organization of poverty relief is more efficient if it occurs at the local level (Fording e.a., 2007; Kazepov, 2010; Sheely, 2012; Soss e.a., 2001; Vandenbroucke e.a., 2016); (2) it provides room for local experimentation to resolve poverty (Kazepov, 2010); and (3) it provides local authorities with more political power concerning highly contested poverty relief policies (Fording et al., 2007; Kazepov, 2010; Soss, Fording, & Schram, 2011).

First, if the adjustment of poverty strategies is a reason to decentralize, we would expect local factors, such as size, labour-market situation, poverty indicators or socio-demographic composition of the population (age, gender, educational level, ethnicity) to affect treatment (Carpentier e.a., 2014). The study described in Chapter 3 and Chapter 5 found some effects of municipal composition, but it was not easy to understand these as conscious responses to a unique local situation (Künzel, 2012). The higher sanction rates in municipalities with higher unemployment rates could be understood as such, in the sense that agencies in areas with fewer job opportunities expect their clients to take any opportunity available. Overall, however, the often claimed need to give local institutions the freedom to make policy choices because they understand their situation best, was not shown in our data. Of course, it might be possible that we were unable to control for the precise local factors that do guide specific choices and policies, or that the local circumstances are more important in relation to other aspects, such as the types of activation programmes that are offered – treatment options that were not studied here.

The second reason given for decentralization is to increase the legitimacy of the policies to local government and citizens (Fording e.a., 2007; Kazepov, 2010; Soss e.a., 2011). This led us to expect that political preferences influence treatment decisions. This hypothesis was confirmed, as the predictions of the professionals in social assistance agencies governed by the Flemish right-wing party (N-VA) differed considerably from predictions in socialist (and central-liberal) governed agencies. This is quite in line with their party programmes, which, in the case of the N-VA, emphasize the importance of self-reliance, activation and control of beneficiaries.

The third reason, namely the wish to give local entities the opportunity to experiment, is more difficult to grasp. I had no information about recent decisions made by agencies to adjust policies or organization structure with the purpose of ameliorating poverty reduction. What was found is that the specialization choices of an agency appear to matter: the more specialized the teams, the more the professionals tended to predict sanctioning (Blom, 2004). These organizational differences can be seen as ways to experiment. It might be concluded that the more professionals go along with a client in all aspects of treatment, the less they see sanctioning as the appropriate treatment for their clients. Of course, the research undertaken here does not give us any insight into the actual effect of such treatment preferences on the poverty outcomes of clients.

In conclusion, I would state that the results in my dissertation suggest that decentralization and the possibility of designing locally adapted policy lines do not live up to their promises. This conclusion is based on both the limited variation across municipalities and on the limited evidence that the specific goals of decentralization are being reached. This limited variation may be explained by the moderating effect of the professionals preparing the files for the agency board and following up on the clients' activation trajectories. As shown in the previous section, variation among these professionals is high.

## **5. Policy questions**

The above results (low use of conditionality in the eligibility procedure, high use of conditionality in the activation trajectory, high use of discretion by professionals, low use of discretion by agencies) lead to two policy concerns: (1) are human and civil rights guaranteed in the present implementation process and (2) do central governments need to take action to guarantee the implementation process as intended by the policymakers?

### **5.1 Guaranteeing of human and civil rights**

In general, on the basis of my data, we can say that the civil right to financial support of a person in financial need is guaranteed in Flanders. The likelihood that hypothetical clients who had no money or resources would receive benefits was high. However, this result does not consider possible non-take-up, caused by several factors, such as information deficits, personal barriers, administrative barriers or problems in the policy design (Van Mechelen & Janssens, 2017). Furthermore, the discretion used, might be regarded as problematic, because of three elements: the magnitude (see Section 3); the non-randomness of the variation (see Section 3.2); and the fact that agencies do not appear to be completely aware of the high levels of variation among their professionals

(see Section 3.4). If the variation found in this research project was completely random and small, it could be seen as an unavoidable aspect of professional discretion. However, it is rather difficult to justify (to clients) the large differences in treatment for the same client, reported Section 3, and the importance of professionals' preferences in explaining these differences. The use of discretion does not seem to be an act of creativity, or tailoring treatment. Rather, professionals seem to believe that they are doing what is expected of them by legislation and by the agency board<sup>38</sup>, while, to some extent, they are simply acting in accordance with their own preferences.

The question then is: is there a way out? Is it possible or desirable to strive for more equality? In addressing this option, a goal could be to make legislation stricter and more concrete, so that professionals can do nothing but follow the rules. Several arguments can be made against this strategy, partly based on the results presented in Section 3. First, such detailed legislation is in fact very arbitrary. Evans (2011) described how attempts to make rules clearer and more specific often lead to more rules and regulations. The more detailed or specialized legislation is, the more choice becomes available to professionals, which encourages the use of discretion. This is because specialized legislation offers professionals more options with regard to what 'to follow and [what] to ignore' (Evans, 2011, p. 381). In Section 3.3, this was shown with regard to the possible fairness reasons 'full-time education' and 'drug addiction'. Second, the post-questionnaire I sent to all participating agencies showed that the municipalities that already have detailed regulations did not exhibit less inter-professional variation (Chapter 5).

Third, there is a moral consideration. Social assistance schemes are systems of last resort, which means that there is no protection left if a person is denied social assistance. The question then is whether it is acceptable to include rules that force professionals and municipalities to exclude people based on detailed prescriptions, without allowing for or encouraging the possibility of assessing individual cases, needs and possibilities. In this way, the system risks excluding needy individuals. To test these arguments, follow-up research should be carried out in countries where legislators have chosen to prescribe the work willingness requirement in more detail.

It seems crucial to strike a good balance between cultivating awareness, acceptance and appreciation of the discretion required for the precarious work that social assistance

---

<sup>38</sup> This statement is also based on conversations with professionals following the presentation of the results of the dissertation that I gave in around 20 agencies.

workers undertake every day, attempting sincerely to ensure that policies are implemented as intended and guaranteeing the rights of clients.

## 5.2 Guaranteeing the intended implementation of legislation

The second policy question concerns the difficult tension between giving responsibility to local actors and the wish to see legislation implemented as intended. A first observation with regard to this issue is that – in line with research by Minas and colleagues (2014) in the Scandinavian countries – inter-municipality variation is small when it comes to the parameters investigated here. One conclusion could be: why should we bother?

On the other hand, however, 15% of the organizations do have eligibility and sanction treatments that are significantly more or less generous than those of others. Belonging to this group correlates with the number of social assistance beneficiaries in the municipality, with the political majority and the sanction preferences of the head of the social assistance team. In this context, Kazepov (2008) mentions the implicit rescaling of social policies. This is a process in which, without explicit intent, local levels become more important and higher level/national levels lose some policymaking power. One of the mechanisms behind this is the increase in the client population of decentralized benefit schemes. Due to several factors, one of them being the increasing conditionality of nationally organized social security institutions, social assistance populations are growing throughout Europe. Bonoli and Trein (2016) speak of ‘cost shifting’. In systems with multi-layered responsibilities, the different policy levels attempt to shift clients and their ‘costs’ to other systems at other levels. This may occur indirectly if federal levels are made less generous or accessible (Bonoli & Trein, 2016). The growing conditionality of social security, both in legislation and in implementation, might, in this sense, be a form of cost shifting, caused by growing numbers of clients and shrinking budgets. While small differences between municipalities are negligible if the numbers of beneficiaries are low, these differences become more pronounced if the dependent population increases. In this sense, a ‘residence lottery’ is increasingly coming into play, with potentially large consequences for some vulnerable people (Chapter 3 and Chapter 5).

Third, another reason to be sceptical about the decentralization trend in several countries is the fact that the high expectations of the advocates of decentralization (more tailored and locally supported treatment of people in poverty) do not materialize in the data. Thus, if these higher goals are not being reached, why should we then accept the small but potentially problematic variation that was found? One specific goal of decentralization, namely the influence of the board’s political party, was found to be important in predicting treatment. Decentralizing policymaking power is a means of



responsibilizing local governments to the difficult problem of the fight against poverty. Here, the advantage of decentralization might be that local communities are more inclined to invest in the problem if they bear more responsibility for solving it.

Again, an important trade-off needs to be made here. Variation is limited and local communities may feel involved in the issue because of their policymaking mandate. As a result, the system appears acceptable as it is. Yet, if the social assistance population continues to grow, the question is whether it is still acceptable to make claimants dependent on local preferences.

## **6. Policy proposals**

Policy proposals based on these results need to take into account: (1) the growing importance of conditionality in public opinion and in political discourse (Chapter 2), (2) discretion dynamics that are inevitable and important, but possibly disturbing (Chapter 3 and Chapter 5), and (3) the growing number of clients becoming dependent on the last resort system, which actually enhances decentralization tendencies (Chapter 6).

### **6.1 Administratization of the eligibility test**

Considering all of these elements together, I propose to construct a simple administrative eligibility test for social assistance benefits in Belgium, for five reasons. First, if discretion or local variation is ever problematic, then this is because of the right to a last resort benefit for eligible citizens. Being considered eligible or ineligible is the difference between having money and not having money. While variation is more acceptable and desirable with regard to activation trajectories, it should not be allowed with regard to eligibility tests. Second, other research has shown that non-take-up of social assistance benefits is unacceptably high. It has been estimated that around 50% of all eligible citizens are not receiving benefits. At present, it remains impossible to determine the main driving factor for this large number. However, it has been shown that complexity in the administration of eligibility procedures increases the process costs for clients, which makes both not starting and not finishing the process more likely (Van Mechelen & Janssens, 2017). Well-designed administrative eligibility tests might decrease such process costs. Third, one of the criteria which requires assessment by a professional, in the sense that it is difficult to prescribe intra-municipality regulations<sup>39</sup>, namely the client's willingness to work, concerns behaviour and is thus not an objective factor. However, it was not found to be an important predictor of eligibility. Therefore, it

---

<sup>39</sup> Some discretion is allowed in assessing the means test (income from selling a house and income from an inheritance).

seems reasonable to drop this criterion from the eligibility procedure, while continuing to emphasize its importance in the remainder of the trajectory. Fourth, the results of this dissertation suggest that leaving leeway with regard to some conditions creates effects for certain criteria that should not be under assessment, such as gender and nationality. Fifth, as numbers of beneficiaries are increasing in Belgium and other European countries, the question is whether tailored treatment with regard to all aspects of a trajectory remains still possible and desirable. It might be a gain in both efficiency and workload for the social workers to make applications more administrative. However, this step should only be taken provided that certain conditions are met.

The application criteria should be as simple as possible (Watts, 2014), because the less clear the conditions, the more assessment by professionals will be required and the more variation and risk of losing rights might occur (Evans, 2011). One possible technique is the use of data available in electronic databases. Such databases include information about family composition, activity status and income, though often provide insufficient detail to compose the means test that leads to social assistance. First, the official domicile information does not always match reality. Second, most income data are at least a few months old, so a form could be made available to allow clients to verify and update the information. This form could be made available online, giving clients a first impression of their rights. France is an example of a country that has such online forms for social assistance (<https://www.service-public.fr/particuliers/vosdroits/R558>). In Belgium, online intake exists for social assistance for the elderly ([http://www.onprvp.fgov.be/demo/pensioenaanvraag\\_NL.swf](http://www.onprvp.fgov.be/demo/pensioenaanvraag_NL.swf)). In the Netherlands, some municipalities have opted for administrative intake in which most clients only need to have one telephone conversation before receiving the benefit<sup>40</sup>. Another possibility is to administratize processes involving clients who have typical profiles, such as those who move from one municipality to another but do not change life circumstances in any other way, or clients who have been temporarily withdrawn from other types of benefits. Lastly, an option could be to use the administrative data to determine eligibility and to guarantee the lowest social assistance level (e.g. the category 'cohabiting') and to start the social investigation if the client claims to belong to another category.

---

<sup>40</sup> Information obtained in conversations with Dutch social assistance professionals.

Furthermore, there will still need to be a service which is easy to reach and access, where people can receive help with applying for benefits. Moreover, if a claimant is not considered eligible there must be the possibility of reapplying through a tailored procedure, which is again easy to reach and to access. Otherwise, the administratization of the eligibility test will introduce an *eligibility-compliance culture* in which administration is more important than human need (Lens, 2006; Meyers e.a., 1998). This would defeat the purpose of the policy proposal. Lastly, the activation and integration tasks of social assistance agencies should remain as important as before.

## 6.2 Re-evaluation of discretion with regard to other forms of treatment

A second policy proposal would be to re-evaluate discretion with regard to forms of treatment other than eligibility treatment. In this respect, this research agrees with traditional social work literature, which states that – certainly if treatments become more complex and oriented to behavioural change or assessment – discretion is key in social work practice, both because it is inevitable and because it makes the treatment more meaningful. Concretely, this means that professionals should be educated and allowed to look for solutions for every situation they might encounter, without having internal guidelines that formulate answers to all of these situations. This means giving professionals the autonomy they need to look for tailored and valuable solutions that are in line with the needs and opportunities available to the client, taking into account their human rights, social assistance legislation, local expectations and the preferences of and possibilities available to the professional. The goal would be to raise awareness, first, that discretion is already being used and, second, that it can be used, for example with regard to the interpretation of exemptions from work willingness in sanction treatment (Lens, 2008).

Two insights seem valuable in this respect. First, team intervention is an important tool in developing client and community tailored treatment proposals. The analyses in this research did not provide clear proof that more team involvement leads to less variation among professionals, however. Thus, team intervention is not necessarily a suitable tool for fostering equal treatment, although it may be an opportunity for professionals to become more aware of differences between colleagues and of the welfare state attitudes that guide their treatment and those of colleagues. In this regard, the team is not the place for professionals to exchange ideas and distinguish the right treatment proposal together, but instead a place for professionals to become aware of several possible proposals and to choose between them individually based on a conscious consideration of treatment possibilities. Second, changes in the organization of the agency might help to stimulate autonomy. Choosing to rely on less specialized teams is

one of the possible changes. It is well known that more specialization leads to more routinization; however, if discretion is sincerely valued, routinization and specialization will not be appropriate solutions.

#### **7. Methodological implication: combining the study of outcomes and implementation processes**

Finally, it would not have been possible to arrive at the conclusions above if the study of predicted outcomes had not been combined with information about the different implementation levels. If I had focused on legislation or the interpretation of legislation by stakeholders (see the follow-up survey in 2018), I would not have concluded that work willingness is not important in the eligibility process and that fairness reasons are seldom used. If, instead, I had used outcome data, such as registration databases, I would not have been able to include information on work conditionality in the analyses, as it is not usually included in registration files. I would also have been blind to the importance of the professional level. I would probably have formulated policy proposals for legislation, without taking the implementation process into account (Meyers e.a., 1998). Furthermore, if I had studied variation across agencies quantitatively, again without investigating the professional level, I would have concluded mistakenly that variation between agencies is relatively high. Lastly, if the context in which discretion takes shape had been my main research interest and I had interviewed clients, team coordinators or professionals about their insights and experiences, I would not have gained any insight into the magnitude of variation among professionals.

I believe this dissertation shows that, both in the social policy branch and in the social work and social administration tradition, research misses a vast amount of information and insight if it focuses too narrowly on one of the policymaking levels. Drafting good policies is only useful if their implementation also receives attention and if the role of professionals and local policymaking levels is considered. Similarly, discussions of the positive and negative consequences of discretion and decentralization are only useful if the outcomes for actual clients are used as a touchstone.

## References

---

- Abts, K., Swyngedouw, M., Meuleman, B., Baute, S., Galle, J., & Gaasendam, C. (2015). Belgian National Election Study 2014. Codebook: Questions and Frequency tables. Leuven: ISPO-KU Leuven & CLEO - Université de Liege.
- Aguinis, H., & Bradley, K. J. (2014). Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies. *Organizational Research Methods, 17*(4), 351–371.
- Aiman-Smith, L., Scullen, S. E., & Barr, S. H. (2002). Conducting Studies of Decision Making in Organizational Contexts: A Tutorial for Policy-Capturing and Other Regression-Based Techniques. *Organizational Research Methods, 5*(4), 388–414.
- Armocost, R. L., Hosseini, J. C., Morris, S. A., & Rehbein, K. A. (1991). An Empirical Comparison of Direct Questioning, Scenario, and Randomized Response Methods for Obtaining Sensitive Business Information. *Decision Sciences, 22*(5), 1073–1090.
- Asparouhov, T., & Muthén, B. (2014). Multiple-Group Factor Analysis Alignment. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(4), 495–508.
- Atzmüller, C., & Steiner, P. M. (2010). Experimental Vignette Studies in Survey Research. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 6*(3), 128–138.
- Auspurg, K., & Hinz, T. (2014). *Factorial Survey Experiments*. SAGE Publications.
- Auspurg, K., Hinz, T., Liebig, S., & Sauer, C. (2015). The Factorial Survey as a Method for Measuring Sensitive Issues. In *Improving Survey Methods. Lessons from Recent Research*.
- Auspurg, K., Hinz, T., & Sauer, C. (2017). Why Should Women Get Less? Evidence on the Gender Pay Gap from Multifactorial Survey Experiments. *American Sociological Review, 82*(1), 179–210.
- Auspurg, K., & Jäckle, A. (2015). First Equals Most Important? Order Effects in Vignette-Based Measurement. *Sociological Methods & Research*.
- Bargain, O., Immervoll, H., & Viitamäki, H. (2012). No claim, no pain. Measuring the non-take-up of social assistance using register data. *Journal of Economic Inequality, 10*(3), 375–395.
- Barretti, M. (2004). What do we know about the professional socialization of our students? *Journal of Social Work Education, 40*(2), 255–283.
- Baughman, K., Ludwick, R., Merolla, D., Palmisano, B., Hazelett, S., Winchell, J., & Hewit, M. (2012). Professional Judgments About Advance Care Planning With Community-Dwelling Consumers. *Journal of Pain and Symptom Management, 43*(1), 10–19.
- Bean, C., & Papadakis, E. (1998). A Comparison of Mass Attitudes towards the Welfare State in Different Institutional Regimes, 1985–1990. *International Journal of Public Opinion Research, 10*(3), 211–236.

- Belfius. (2007). Lokale financiën. Sociaaleconomische typologie van de gemeenten. Belfius.
- Bell, H. (2005). Caseworkers' Assessment of Welfare Reform. *Journal of Human Behavior in the Social Environment*, 12(2–3), 243–259.
- Benz, A. (2000). Two types of multi-level governance: Intergovernmental relations in German and EU regional policy. *Regional & Federal Studies*, 10(3), 21–44.
- Bergmark, Å., Bäckman, O., & Minas, R. (2017). Organizing local social service measures to counteract long-term social assistance receipt. What works? Experiences from Sweden. *European Journal of Social Work*, 20(4), 548–559. <https://doi.org/10.1080/13691457.2016.1206851>
- Blekesaune, M., & Quadagno, J. (2003). Public Attitudes toward Welfare State Policies: A Comparative Analysis of 24 Nations. *European Sociological Review*, 19(5), 415–427.
- Blom, B. (2004). Specialization in Social Work Practice: Effects on Interventions in the Personal Social Services. *Journal of Social Work*, 4(1), 25–46.
- Blomberg, H., Kallio, J., Kroll, C., & Niemela, M. (2015). What explains frontline workers' views on poverty? A comparison of three types of welfare sector institutions. *International Journal of Social Welfare*, 24(4), 324–334.
- Blomberg, H., Kroll, C., Kallio, J., & Erola, J. (2013). Social workers' perceptions of the causes of poverty in the Nordic countries. *Journal of European Social Policy*, 23(1), 68–82.
- Blommesteijn, M., van Geuns, R. C., Groenewoud, M., & Slotboom, S. T. (2012). *Vakkundig aan de slag. Een onderzoek naar vakmanschap in de gemeentelijke re-integratiesector* (p. 74). Amsterdam: Regioplan.
- Bonoli, G., & Trein, P. (2016). Cost-Shifting in Multitiered Welfare States: Responding to Rising Welfare Caseloads in Germany and Switzerland. *Publius: The Journal of Federalism*, 46(4), 596–622.
- Bovens, M., & Zouridis, S. (2002). From street-level to system-level bureaucracies: How information and communication technology is transforming administrative discretion and constitutional control. *Public Administration Review*, 62(2), 174–184.
- Brodkin, E. Z. (2011). Policy Work: Street-Level Organizations Under New Managerialism. *Journal of Public Administration Research and Theory*, 21(suppl 2), i253–i277.
- Bryman, A. (2012). *Social research methods*. Oxford; New York: Oxford University Press.
- Buffat, A. (2015). Street-Level Bureaucracy and E-Government. *Public Management Review*, 17(1), 149–161.
- Bullock, H. E. (2004). From the Front Lines of Welfare Reform: An Analysis of Social Worker and Welfare Recipient Attitudes. *The Journal of Social Psychology*, 144(6), 571–590.
- Byers, B., & Zeller, R. A. (1998). Measuring Subgroup Variation in Social Judgment Research: A Factorial Survey Approach. *Social Science Research*, 27(1), 73–84.

- Carpentier, S. (2016). *Lost in Transition? Essays on Socio-Economic Trajectories of Social Assistance Beneficiaries*. Universiteit Antwerpen, Antwerpen.
- Carpentier, S., Neels, K., & Van den Bosch, K. (2014). How Do Exit Rates from Social Assistance Benefit in Belgium Vary with Individual and Local Agency Characteristics? In S. Carcillo, H. Immervoll, S. Jenkins P., S. Königs, & K. Tatsiramos (Red.), *Safety Nets and Benefit Dependence* (pp. 151–187). Emerald Group Publishing Limited.
- Castillo, J. T., & Becerra, D. (2012). The Perception of Poverty and Social Welfare Policies among Undergraduate and Graduate Social Work Students in the United States. *Journal of Human Behavior in the Social Environment*, 22(4), 375–391.
- Caussade, S., Ortúzar, J. de D., Rizzi, L. I., & Hensher, D. A. (2005). Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B: Methodological*, 39(7), 621–640.
- Champion, C., & Bonoli, G. (2011). Institutional fragmentation and coordination initiatives in western European welfare states. *Journal of European Social Policy*, 21(4), 323–334.
- Cherlin, A. J., Bogen, K., Quane, J. M., & Burton, L. (2002). Operating within the Rules: Welfare Recipients' Experiences with Sanctions and Case Closings. *Social Service Review*, 76(3), 387–405.
- Clasen, J., & Clegg, D. (2007). Levels and levers of conditionality: measuring change within welfare states. In J. Clasen & Nico Siegel A., *Investigating Welfare State Change: The "Dependent Variable Problem" in Comparative Analysis* (pp. 166–197). Cheltenham: Edward Elgar Publishing.
- Corazzini, K. (2003). How State-Funded Home Care Programs Respond to Changes in Medicare Home Health Care: Resource Allocation Decisions on the Front Line. *Health Services Research*, 38(5), 1263–1282.
- Costa, L. P. da, & Dias, J. G. (2015). What do Europeans Believe to be the Causes of Poverty? A Multilevel Analysis of Heterogeneity Within and Between Countries. *Social Indicators Research*, 122(1), 1–20.
- Cozzarelli, C., Wilkinson, A. V., & Tagler, M. J. (2001). Attitudes Toward the Poor and Attributions for Poverty. *Journal of Social Issues*, 57(2), 207–227.
- Csikai, E. L., & Rozensky, C. (1997). "Social Work Idealism" and Students' Perceived Reasons for Entering Social Work. *Journal of Social Work Education*, 33(3), 529–538.
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement Equivalence in Cross-National Research. *Annual Review of Sociology*, 40(1), 55–75.
- De Bie, M., & Vandenbussche, J. (2016). 40 jaar OCMW-praktijk op de spanning tussen bijstand als sociaal beleidsinstrument en bijstand als sociaal grondrecht. In M. De Wilde, B. Cantillon, F. Vandenbroucke, & M. De Bie (Red.), *40 jaar OCMW en bijstand*. Leuven: Acco.
- De Deken, J., & Kittel, B. (2007). Social expenditure under scrutiny: the problems of using aggregate spending data for assessing welfare state dynamics. In J. Clasen & N. A.

Siegel, *Investigating Welfare State Change: The “Dependent Variable Problem” in Comparative Analysis*. Cheltenham, Northampton: Edward Elgar.

De Wilde, M. (2016). Het gebruik van de discretionaire ruimte door Vlaamse OCMW-medewerkers. In M. De Wilde, B. Cantillon, F. Vandenbroucke, & M. De Bie (Red.), *40 jaar OCMW en bijstand* (pp. 104–120). Leuven: Acco.

De Wilde, M. (2017). Deservingness in social assistance administrative practice: a factorial survey approach. In W. van Oorschot, B. Meuleman, F. Roosma, & T. Reeskens (Red.), *The Social Legitimacy of Targeted Welfare: Attitudes to Welfare Deservingness*. Cheltenham: Edward Elgar Publishing.

De Wilde, M., & Goos, P. (2017). Studying the multidimensionality of social policy implementation: the factorial survey approach. *CSP Working Papers*, 17(6).

De Wilde, M., & Marchal, S. (2016). De voorwaardelijkheid van de bijstand: de werkbereidheidseis. In M. De Wilde, B. Cantillon, F. Vandenbroucke, & M. De Bie (Red.), *40 jaar OCMW en bijstand* (pp. 84–103). Leuven: Acco.

De Wilde, M., & Marchal, S. (2018). Weighing up work willingness in social assistance: a balancing act on multiple levels. *CSP Working Papers*, 18(08).

Degenholtz, H. B., Kane, R. A., Kane, R. L., & Finch, M. D. (1999). Long-Term Care Case Managers’ Out-of-Home Placement Decisions: An Application of Hierarchical Logistic Regression. *Research on Aging*, 21(2), 240–274.

DeShazo, J. R., & Fermo, G. (2002). Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. *Journal of Environmental Economics and Management*, 44(1), 123–143.

Ditch, J. (1999). The structure and dynamics of social assistance in the European Union. In M. Heikkilä (Red.), *Linking welfare and work* (pp. 59–70). Luxembourg: Office for Official Publications of the European Communities.

Dubois, V. (2010). *The Bureaucrat and the Poor: Encounters in French Welfare Offices*. Aldershot: Ashgate Publishing.

Dülmer, H. (2007). Experimental Plans in Factorial Surveys Random or Quota Design? *Sociological Methods & Research*, 35(3), 382–409.

Dülmer, H. (2016). The Factorial Survey: Design Selection and its Impact on Reliability and Internal Validity. *Sociological Methods & Research*, 45(2), 304–347.

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). New Public Management Is Dead—Long Live Digital-Era Governance. *Journal of Public Administration Research and Theory*, 16(3), 467–494.

Dupont, C. (2004, augustus 3). Omzendbrief wet van 26 mei 2002 betreffende het recht op maatschappelijke integratie - studenten en het recht op een leefloon. Bestuur van de Maatschappelijke Integratie.

Eardley, T., Bradshaw, J., Ditch, J., Gough, I., & Whiteford. (1996). *Social Assistance in OECD Countries: Synthesis Report, Department of Social Security Research Report* (No. 46). London: HMSO.



- Edlund, J. (1999). Trust in government and welfare regimes: attitudes to redistribution and financial cheating in the USA and Norway. *European Journal of Political Research*, 35(3), 341–370.
- Eichhorst, W., Kaufmann, O., Konle-Seidl, R., & Reinhard, H.-J. (2008). Bringing the Jobless into Work? An Introduction to Activation Policies. In D. W. Eichhorst, D. O. Kaufmann, & R. Konle-Seidl (Red.), *Bringing the Jobless into Work?* (pp. 1–16). Springer Berlin Heidelberg.
- Eifler, S. (2010). Validity of a Factorial Survey Approach to the Analysis of Criminal Behavior. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 6(3), 139–146.
- Eleveld, A. (2016). *Work-Related Sanctions in European Welfare States: An Incentive to Work or a Violation of Minimum Subsistence Rights?* (SSRN Scholarly Paper No. ID 2802656). Rochester, NY: Social Science Research Network.
- Elkins, D. J., & Simeon, R. E. B. (1979). A Cause in Search of Its Effect, or What Does Political Culture Explain? *Comparative Politics*, 11(2), 127–145.
- Ervasti, H. (2012). Who hates the welfare state? Criticism of the welfare state in Europe. In *The future of the welfare state: Social policy attitudes and social capital in Europe* (p. 231).
- Evans. (2011). Professionals, Managers and Discretion: Critiquing Street-Level Bureaucracy. *British Journal of Social Work*, 41(2), 368–386.
- Evans. (2012). *Professional Discretion in Welfare Services: Beyond Street-Level Bureaucracy*. Ashgate Publishing, Ltd.
- Evans, & Harris, J. (2004). Street-Level Bureaucracy, Social Work and the (Exaggerated) Death of Discretion. *British Journal of Social Work*, 34(6), 871–895.
- Evans, P. M. (2007). (Not) taking account of precarious employment: Workfare policies and lone mothers in Ontario and the UK. *Soc Policy Admin*, 41(1), 29–49.
- Evetts, J. (2002). New directions in state and international professional occupations: discretionary decision-making and acquired regulation. *Work Employment and Society*, 16(2), 341–353.
- Feldman, S., & Steenbergen, M. R. (2001). The Humanitarian Foundation of Public Support for Social Welfare. *American Journal of Political Science*, 45(3), 658–677.
- Fording, R. C., Soss, J., & Schram, S. F. (2007). Devolution, Discretion, and the Effect of Local Political Values on TANF Sanctioning. *Social Service Review*, 81(2), 285–316.
- Fording, R. C., Soss, J., & Schram, S. F. (2011). Race and the Local Politics of Punishment in the New World of Welfare. *American Journal of Sociology*, 116(5), 1610–1657.
- Forssen, K. (1998). Decentralization of decision making: the case of payment policies for children's daycare. *Scandinavian Journal of Social Welfare*, 7(4), 277–287.
- Godfrey, E. B., & Yoshikawa, H. (2012). Caseworker-Recipient Interaction: Welfare Office Differences, Economic Trajectories, and Child Outcomes: Offices, Economic Trajectories, and Child Outcomes. *Child Development*, 83(1), 382–398.

- Goos, P., & Jones, B. (2011). *Optimal Design of Experiments: A Case Study Approach* (1ste dr.). West-Sussex: Wiley.
- Groves, R. M., Jr, F. J. F., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology*. John Wiley & Sons.
- Guy, A. (2011). Vocational Choice and Attitudes toward Welfare Policy. *The British Journal of Social Work, 41*(7), 1321–1339.
- Hackett, S., Kuronen, M., Matthies, A.-L., & Kresal, B. (2003). The motivation, professional development and identity of social work students in four European countries. *European Journal of Social Work, 6*(2), 163–178.
- Handler, J. F. (2003). Social Citizenship and Workfare in the US and Western Europe: From Status to Contract. *Journal of European Social Policy, 13*(3), 229–243.
- Harrison, G., & Healy, K. (2016). Forging an Identity as a Newly Qualified Worker in the Non-government Community Services Sector. *Australian Social Work, 69*(1), 80–91.
- Hasenfeld, Y. (2010). Organizational Responses to Social Policy: The Case of Welfare Reform. *Administration in Social Work, 34*(2), 148–167.
- Hasenfeld, Y., Ghose, T., & Larson, K. (2004). The Logic of Sanctioning Welfare Recipients: An Empirical Assessment. *Social Service Review, 78*(2), 304–319.
- Hasenfeld, Y., & Rafferty, J. A. (1989). The Determinants of Public Attitudes Toward the Welfare State. *Social Forces, 67*(4), 1027–1048.
- Heeringa, S., & Berglund, P. A. (2010). *Applied survey data analysis*. Boca Raton, FL: Chapman & Hall/CRC.
- Heidenreich, M., & Aurich-Beerheide, P. (2014). European worlds of inclusive activation: The organisational challenges of coordinated service provision. *International Journal of Social Welfare, 23*, S6–S22.
- Hermans, K. (2005). *De actieve welvaartsstaat in werking. Een sociologische studie naar de implementatie van het activeringsbeleid op de werkvloer van de Vlaamse OCMW's*. KULeuven, Leuven.
- Hölsch, K., & Kraus, M. (2006). European schemes of social assistance: an empirical analysis of set-ups and distributive impacts. *International Journal of Social Welfare, 15*(1), 50–62.
- Holzner, C., Munz, S., & Büttner, T. (2009). Evaluating the Organisational Performance of Local Job Centers.
- Hooghe, L. (1996). *Cohesion Policy and European Integration*. Oxford: Oxford University Press.
- Hooghe, L., & Marks, G. (2003). Unraveling the central state, but how? Types of multi-level governance. *American Political Science Review, 97*(2), 233–243.
- Hox, J. J., Kreft, I. G. G., & Hermkens, P. L. J. (1991). The Analysis of Factorial Surveys. *Sociological Methods & Research, 19*(4), 493–510.

- Hupe, P., & Hill, M. (2007). Street-level bureaucracy and public accountability. *Public Administration*, 85(2), 279–299.
- Immervoll, H. (2009). Minimum-Income Benefits in OECD Countries: Policy Design. *Effectiveness and Challenges (IZA Discussion Paper)*, 4627.
- Jæger, M. M. (2006). What Makes People Support Public Responsibility for Welfare Provision: Self-Interest or Political Ideology? A Longitudinal Approach. *Acta Sociologica*, 49(3), 321–338.
- Jasso, G. (2006). Factorial Survey Methods for Studying Beliefs and Judgments. *Sociological Methods & Research*, 34(3), 334–423.
- Jasso, G., & Milgrom, E. M. M. (2008). Distributive Justice and CEO Compensation. *Acta Sociologica*, 51(2), 123–143.
- Jessen, J. T., & Tufte, P. A. (2014). Discretionary Decision-Making in a Changing Context of Activation Policies and Welfare Reforms. *Journal of Social Policy*, 43(2), 269–288.
- Johnson, F. R. (2006). Comment on “Revealing Differences in Willingness to Pay Due to the Dimensionality of Stated Choice Designs: An Initial Assessment”. *Environmental and Resource Economics*, 34(1), 45–50.
- Kallio, J., & Kouvo, A. (2015). Street-level Bureaucrats’ and the General Public’s Deservingness Perceptions of Social Assistance Recipients in Finland. *Social Policy & Administration*, 49(3), 316–334.
- Karren, R. J., & Barringer, M. W. (2002). A Review and Analysis of the Policy-Capturing Methodology in Organizational Research: Guidelines for Research and Practice. *Organizational Research Methods*, 5(4), 337–361.
- Kazepov, Y. (1999). At the Edge of Longitudinal Analysis. Welfare Institutions and Social Assistance Dynamics, 33(3), 305–322.
- Kazepov, Y. (2005). Cities of Europe: Changing Contexts, Local Arrangements, and the Challenge to Social Cohesion. In Y. Kazepov (Red.), *Cities of Europe* (pp. 3–42). Blackwell Publishing Ltd.
- Kazepov, Y. (2008). The Subsidiarization of Social Policies: Actors, Processes and Impacts. *European Societies*, 10(2), 247–273.
- Kazepov, Y. (2010). *Rescaling Social Policies: Towards Multilevel Governance in Europe*. Ashgate Publishing, Ltd.
- Kazepov, Y., & Barberis, E. (2013). Social Assistance Governance in Europe: Towards a Multilevel Perspective. In I. Marx & K. Nelson (Red.), *Minimum Income Protection in Flux* (pp. 217–248). Palgrave Macmillan UK.
- Keiser, L. R. (1999). State Bureaucratic Discretion and the Administration of Social Welfare Programs: The Case of Social Security Disability. *Journal of Public Administration Research and Theory: J-PART*, 9(1), 87–106.
- Keiser, L. R. (2010). Understanding Street-Level Bureaucrats’ Decision Making: Determining Eligibility in the Social Security Disability Program. *Public Administration Review*, 70(2), 247–257.

- Keiser, L. R., Mueser, P. R., & Choi, S.-W. (2004). Race, Bureaucratic Discretion, and the Implementation of Welfare Reform. *American Journal of Political Science*, 48(2), 314–327.
- Killick, C., & Taylor, B. J. (2012). Judgements of Social Care Professionals on Elder Abuse Referrals: A Factorial Survey. *British Journal of Social Work*, 42(5), 814–832.
- Kirwan, J. R., Chaput de Saintonge, D. M., Joyce, C. R., & Currey, H. L. (1983). Clinical judgment in rheumatoid arthritis. I. Rheumatologists' opinions and the development of "paper patients". *Annals of the Rheumatic Diseases*, 42(6), 644–647.
- Kirwan, J. R., & Currey, H. L. (1984). Clinical judgment in rheumatoid arthritis. IV. Rheumatologists' assessments of disease remain stable over long periods. *Annals of the Rheumatic Diseases*, 43(5), 695–697.
- Kittel, B., & Obinger, H. (2003). Political parties, institutions, and the dynamics of social expenditure in times of austerity. *Journal of European Public Policy*, 10(1), 20–45.
- Kraus, S. J. (1995). Attitudes and the Prediction of Behavior: A Meta-Analysis of the Empirical Literature. *Personality and Social Psychology Bulletin*, 21(1), 58–75.
- Kroeger, N. (1975). Bureaucracy, Social-Exchange, and Benefits Received in a Public-Assistance Agency. *Social Problems*, 23(2), 182–196.
- Krumer-Nevo, M., & Lev-Wiesel, R. (2005). Attitudes of social work students toward clients with basic needs. *Journal of Social Work Education*, 41(3), 545–556.
- Kulin, J., & Svallfors, S. (2013). Class, Values, and Attitudes Towards Redistribution: A European Comparison. *European Sociological Review*, 29(2), 155–167.
- Künzel, S. (2012). The local dimension of active inclusion policy. *Journal of European Social Policy*, 22(1), 3–16.
- Kuppens, T., & Yzerbyt, V. Y. (2014). Predicting variability: Using multilevel modelling to assess differences in variance. *European Journal of Social Psychology*, 44(7), 691–700.
- Langley, G. R., Tritchler, D. L., Llewellyn-Thomas, H. A., & Till, J. E. (1991). Use of written cases to study factors associated with regional variations in referral rates. *Journal of Clinical Epidemiology*, 44(4–5), 391–402.
- Larsen, C. A. (2008). The Institutional Logic of Welfare Attitudes: How Welfare Regimes Influence Public Support. *Comparative Political Studies*, 41(2), 145–168.
- Lens, V. (2006). Examining the Administration of Work Sanctions on the Frontlines of the Welfare System. *Social Science Quarterly*, 87(3), 573–590.
- Lens, V. (2008). Welfare and Work Sanctions: Examining Discretion on the Front Lines. *Social Service Review*, 82(2), 197–222.
- Lepianka, D., Gelissen, J., & van Oorschot, W. (2010). Popular Explanations of Poverty in Europe: Effects of Contextual and Individual Characteristics across 28 European Countries. *Acta Sociologica*, 53(1), 53–72.
- Limb, G. E., & Organista, K. C. (2006). Change Between Entry and Graduation in MSW Student Views on Social Work's Traditional Mission, Career, Motivations, and Practice

Preferences: Caucasian, Student of Color and American Indian Group Comparisons. *Journal of Social Work Education*, 42(2), 269–290.

Lipsky, M. (1980). *Street-Level Bureaucracy. Dilemmas of the Individual in Public Service*. New York: Russell Sage Foundation.

Lødemel, I., & Trickey, H. (2001). *“An Offer You Can’t Refuse”: Workfare in International Perspective*. Bristol: The Policy Press.

Lorentzen, T., Dahl, E., & Harsløf, I. (2012). Welfare risks in early adulthood: A longitudinal analysis of social assistance transitions in Norway. *International Journal of Social Welfare*, 21(4), 408–421.

Maloy, K., Pavetti, L., Shin, P., Darnell, J., & Scarpulla-Nolan, L. (1998). *Description and Assessment of State Approaches to Diversion Programs and Activities Under Welfare Reform*. Washington: The George Washington University Center for Health Policy Research.

Marchal, S., Marx, I., & Van Mechelen, N. (2014). The Great Wake-Up Call? Social Citizenship and Minimum Income Provisions in Europe in Times of Crisis. *Journal of Social Policy*, 43(02), 247–267.

Marchal, S., & van Mechelen, N. (2017). A New Kid in Town? Active Inclusion Elements in European Minimum Income Schemes. *Social Policy & Administration*, 51(1), 171–194.

Marsh, H. W., Guo, J., Parker, P. D., Nagengast, B., Asparouhov, T., Muthén, B., & Dicke, T. (2017). What to do When Scalar Invariance Fails: The Extended Alignment Method for Multi-Group Factor Analysis Comparison of Latent Means Across Many Groups. *Psychological Methods*.

May, P. J., & Winter, S. C. (2009). Politicians, Managers, and Street-Level Bureaucrats: Influences on Policy Implementation. *Journal of Public Administration Research and Theory*, 19(3), 453–476.

Maynard-Moody, S., & Musheno, M. (2003). *Cops, Teachers, Counselors: Stories from the Front Lines of Public Service*. University of Michigan Press.

Mcdonald, C., & Marston, G. (2005). Workfare as welfare: governing unemployment in the advanced liberal state. *Critical Social Policy*, 25(3), 374–401.

Meyers, M. K., Glaser, B., & Donald, K. M. (1998). On the front lines of welfare delivery: Are workers implementing policy reforms? *Journal of Policy Analysis and Management*, 17(1), 1–22.

Minas, R., Bäckman, O., Jakobsen, V., Korpi, T., Lorentzen, T., & Kauppinen, T. (2014). Rescaling inequality? Welfare reform and local variation in social assistance payments. In *DIVA* (pp. 239–258). Policy Press.

Minas, R., Whrighth, S., & Van Berkel, R. (2012). Decentralization and centralization: Governing the activation of social assistance recipients in Europe. *International Journal of Sociology and Social Policy*, 32(5/6), 286–298.

Mion, L. C., Sandhu, S. K., Khan, R. H., Ludwick, R., Claridge, J. A., Pile, J., ... Winchell, J. (2010). Effect of Situational and Clinical Variables on the Likelihood of Physicians

Ordering Physical Restraints. *Journal of the American Geriatrics Society*, 58(7), 1279–1288.

MISSOC. (2015). MISSOC Comparative Tables Database. Geraadpleegd van <http://www.missoc.org/MISSOC/comparativeTables>

Mohan, D., Fischhoff, B., Farris, C., Switzer, G. E., Rosengart, M. R., Yealy, D. M., ... Barnato, A. E. (2014). Validating a Vignette-Based Instrument to Study Physician Decision Making in Trauma Triage. *Medical Decision Making*, 34(2), 242–252.

Moller, M. O., & Stone, D. (2013). Disciplining Disability under Danish Active Labour Market Policy. *Social Policy & Administration*, 47(5), 586–604.

Monnat, S. M. (2010). The Color of Welfare Sanctioning: Exploring the Individual and Contextual Roles of Race on TANF Case Closures and Benefit Reductions. *Sociological Quarterly*, 51(4), 678–707.

Morley, C. P. (2010). The effects of patient characteristics on ADHD diagnosis and treatment: a factorial study of family physicians. *BMC Family Practice*, 11(1), 11.

Morrison, E. W. (2006). Doing the Job Well: An Investigation of Pro-Social Rule Breaking. *Journal of Management*, 32(1), 5–28.

Müller-Engelmann, M., Krones, T., Keller, H., & Donner-Banzhoff, N. (2008). Decision making preferences in the medical encounter – a factorial survey design. *BMC Health Services Research*, 8(1), 260.

Murray, C. (2006). State intervention and vulnerable children: Implementation revisited. *Journal of Social Policy*, 35, 211–227.

Nothdurfter, U. (2016). The street-level delivery of activation policies: constraints and possibilities for a practice of citizenship. *European Journal of Social Work*, 19(3–4), 420–440.

Nothdurfter, U., & Hermans, K. (2018). Meeting (or not) at the street level? A literature review on street-level research in public management, social policy and social work. *International Journal of Social Welfare*, 0(0).

Nybohm, J. (2013). Activation and ‘coercion’ among Swedish social assistance claimants with different work barriers and socio-demographic characteristics: What is the logic? *International Journal of Social Welfare*, 22(1), 45–57.

OECD. (2013). An explanation of social assistance, pension schemes, insurance schemes and similar concepts. *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth*.

Ortega, E. G., Baz, B. O., & Sánchez, F. L. (2012). Professionals’ Criteria for Detecting and Reporting Child Sexual Abuse. *The Spanish Journal of Psychology*, 15(03), 1325–1338.

Pavetti, L., Derr, M. K., & Hesketh, H. (2003). *Review of Sanction Policies and Research Studies*. Washington: Department of Health and Human Services.

Peabody, J. W., Luck, J., Glassman, P., Dresselhaus, T. R., & Lee, M. (2000a). Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study

of 3 methods for measuring quality. *JAMA: the journal of the American Medical Association*, 283(13), 1715–1722.

Peabody, J. W., Luck, J., Glassman, P., Dresselhaus, T. R., & Lee, M. (2000b). Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *Journal of the American Medical Association*, 283(13), 1715–1722.

Piattoni, S. (2009). Multi-level Governance: a Historical and Conceptual Analysis. *Journal of European Integration*, 31(2), 163–180.

Priem, R. L., Walters, B. A., & Li, S. (2011). Decisions, Decisions! How Judgment Policy Studies Can Integrate Macro and Micro Domains in Management Research. *Journal of Management*, 37(2), 553–580.

Raeymaeckers, P., & Dierckx, D. (2013). To Work or Not to Work? The Role of the Organisational Context for Social Workers' Perceptions on Activation. *British Journal of Social Work*, 43(6), 1170–1189.

Rehwald, K., Rosholm, M., & Rouland, B. (2016). Does Activating Sick-Listed Workers Work? Evidence from a Randomized Experiment. Geraadpleegd van <https://papers-ssrn-com.dianus.libr.tue.nl/abstract=2742559>

Reingold, D. A., & Liu, H. K. (2009). Do Poverty Attitudes of Social Service Agency Directors Influence Organizational Behavior? *Nonprofit and Voluntary Sector Quarterly*, 38(2), 307–332.

Rice, D. (2013). Street-Level Bureaucrats and the Welfare State: Toward a Micro-Institutionalist Theory of Policy Implementation. *Administration & Society*, 45(9), 1038–1062.

Roosma, F., Gelissen, J., & Oorschot, W. van. (2013). The Multidimensionality of Welfare State Attitudes: A European Cross-National Study. *Social Indicators Research*, 113(1), 235–255.

Roosma, F., van Oorschot, W., & Gelissen, J. (2014). The weakest link in welfare state legitimacy. European perceptions of moral and administrative failure in the targeting of social benefits. *Journal of Comparative Sociology*, 55(6), 489–508.

Sabatinelli, S. (2010). Activation and rescaling: Interrelated questions in social policy. In Y. Kazepov (Red.), *Rescaling Social Policies towards Multilevel Governance in Europe*. (pp. 75–102). Avebury: Ashgate.

Sandfort, J. R. (2000). Moving Beyond Discretion and Outcomes: Examining Public Management from the Front Lines of the Welfare System. *Journal of Public Administration Research and Theory*, 10(4), 729–756.

Saraceno, C. (2002). *Social Assistance Dynamics in Europe: National and Local Poverty Regimes*. Bristol: The Policy Press.

Scott, P. G. (1997). Assessing Determinants of Bureaucratic Discretion: An Experiment in Street-Level Decision Making. *Journal of Public Administration Research and Theory*, 7(1), 35–58.

- Sewpaul, V., & Jones, D. (2005). Global standards for the education and training of the social work profession\*. *International Journal of Social Welfare*, *14*(3), 218–230.
- Sheely, A. (2012). Devolution and Welfare Reform: Re-evaluating “Success”. *Social Work*, *57*(4), 321–331.
- Shlay, A. B. (2010). African American, White and Hispanic child care preferences: A factorial survey analysis of welfare leavers by race and ethnicity. *Social Science Research*, *39*(1), 125–141.
- Skarlicki, D. P., & Turner, R. A. (2014). Unfairness begets unfairness: Victim derogation bias in employee ratings. *Organizational Behavior and Human Decision Processes*, *124*(1), 34–46.
- Soss, J., Fording, R., & Schram, S. F. (2011). The Organization of Discipline: From Performance Management to Perversity and Punishment. *Journal of Public Administration Research and Theory*, *21*(suppl\_2), i203–i232.
- Soss, J., Schram, S. F., Vartanian, T. P., & O’Brien, E. (2001). Setting the Terms of Relief: Explaining State Policy Choices in the Devolution Revolution. *American Journal of Political Science*, *45*(2), 378–395.
- Standing, G. (2005). *Promoting Income Security as a Right: Europe and North America*. London, New York: Anthem Press.
- Stephenson, P. (2013). Twenty years of multi-level governance: ‘Where Does It Come From? What Is It? Where Is It Going?’ *Journal of European Public Policy*, *20*(6), 817–837.
- Stokes, J., & Schmidt, G. (2012). Child Protection Decision Making: A Factorial Analysis Using Case Vignettes. *Social Work*, *57*(1), 83–90.
- Stranz, H., Karlsson, P., & Wiklund, S. (2016). The wide-meshed safety net. Decision-making on social assistance eligibility in Sweden. *European Journal of Social Work*, *0*(0), 1–13.
- Swaan, A. D., Manor, J., Øyen, E., & Reis, E. P. (2000). Elite Perceptions of the Poor: Reflections for a Comparative Research Project. *Current Sociology*, *48*(1), 43–54.
- Tabin, J.-P., & Perriard, A. (2016). Active social policies revisited by social workers. *European Journal of Social Work*, *19*(3–4), 441–454.
- Taylor. (2006). Factorial Surveys: Using Vignettes to Study Professional Judgement. *British Journal of Social Work*, *36*(7), 1187–1207.
- Taylor. (2012). Models for professional judgement in social work. *European Journal of Social Work*, *15*(4), 546–562.
- Taylor, B. J. (2006). Factorial Surveys: Using Vignettes to Study Professional Judgement. *British Journal of Social Work*, *36*(7), 1187–1207.
- Taylor, T. (2014). No Discretion Required? Caseworker Autonomy and the Rules of Welfare Reform. *Sociological Inquiry*, *84*(3), 412–434.
- Terum, L. I., & Heggen, K. (2016). Identification with the SocialWork Profession: The Impact of Education. *British Journal of Social Work*, *46*(4), 839–854.



- Thomann, E., van Engen, N., & Tummers, L. (2018). The Necessity of Discretion: A Behavioral Evaluation of Bottom-Up Implementation Theory. *Journal of Public Administration Research and Theory*.
- Thoren, K. H. (2008a). "Activation Policy in Action": A Street-level Study of Social Assistance in the Swedish Welfare State. Växjö University, Växjö.
- Thoren, K. H. (2008b). *Activation Policy in Action. A Street-Level Study of Social Assistance in the Swedish Welfare State*. Växjö University Press, Växjö.
- Tummers, L., & Bekkers, V. (2014). Policy Implementation, Street-level Bureaucracy, and the Importance of Discretion. *Public Management Review*, 16(4), 527–547.
- Tummers, L., Bekkers, V., Vink, E., & Musheno, M. (2015). Coping During Public Service Delivery: A Conceptualization and Systematic Review of the Literature. *Journal of Public Administration Research and Theory*, 25(4), 1099–1126.
- Tummers, L., Steijn, B., & Bekkers, V. (2012). Explaining the Willingness of Public Professionals to Implement Public Policies: Content, Context, and Personality Characteristics. *Public Administration*, 90(3), 716–736.
- van Berkel, R. (2006). The decentralisation of social assistance in The Netherlands. *International Journal of Sociology and Social Policy*, 26(1/2), 20–31.
- van Berkel, R. (2010a). Professionals without a profession? Redesigning case management in Dutch local welfare agencies. *European journal of social work*, 13(4), 447.
- van Berkel, R. (2010b). The Provision of Income Protection and Activation Services for the Unemployed in "Active" Welfare States. An International Comparison. *Journal of Social Policy*, 39(1), 17–34.
- Van Berkel, R., & Aa, P. V. D. (2012). Activation Work: Policy Programme Administration or Professional Service Provision? *Journal of Social Policy*, 41(3), 493–510.
- van Berkel, R., & Knies, E. (2016). Performance Management, Caseloads and the Frontline Provision of Social Services. *Social Policy & Administration*, 50(1), 59–78.
- Van Kersbergen, K., & Hemerijck, A. (2012). Two Decades of Change in Europe: The Emergence of the Social Investment State. *Journal of Social Policy*, 41, 475–492.
- Van Mechelen, N., & Janssens, J. (2017). Who is to blame? An overview of the factors contributing to the non-take-up of social rights. *CSP Working Papers*, 17(08).
- Van Mechelen, N., Marchal, S., Goedemé, T., Marx, I., & Cantillon, B. (2011). The CSB-Minimum Income Protection Indicators dataset (CSB-MIPI). *CSB-Workingpaper*.
- Van Mechelen, N., Zamora, D., & Cantillon, B. (2016). De groei en de diversificatie van de bijstandspopulatie. In M. De Wilde, B. Cantillon, F. Vandenbroucke, & M. De Bie (Red.), *40 jaar OCMW en bijstand* (pp. 13–32). Leuven: Acco.
- van Oorschot, W. (2000). Who should get what, and why? On deservingness criteria and the conditionality of solidarity among the public. *Policy & Politics*, 28(1), 33–48.

- van Oorschot, W. (2006). Making the difference in social Europe: deservingness perceptions among citizens of European welfare states. *Journal of European Social Policy*, 16(1), 23–42.
- van Oorschot, W., & Meuleman, B. (2012). Welfarism and the multidimensionality of welfare state legitimacy: Evidence from The Netherlands, 2006. *International Journal of Social Welfare*, 21(1), 79–93.
- van Oorschot, W., & Meuleman, B. (2014). Popular deservingness of the unemployed in the context of welfare state policies, economic conditions and cultural climate. In S. Kumlin & I. Stadelmannsteffen (Red.), *How Welfare States Shape The Democratic Public: Policy Feedback, Participation, Voting, and Attitudes* (pp. 244–268). Edward Elgar Publishing.
- van Oorschot, W., Meuleman, B., Roosma, F., & Reeskens, T. (Red.). (2017). *The Social Legitimacy of Targeted Welfare: Attitudes to Welfare Deservingness*. Edward Elgar Publishing.
- van Oorschot, W., Reeskens, T., & Meuleman, B. (2012). Popular perceptions of welfare state consequences: A multilevel, cross-national analysis of 25 European countries. *Journal of European Social Policy*, 22(2), 181–197.
- van Oorschot, W., Uunk, W. J. G., & Jeene, M. D. (2008). Who should get what and why, under which conditions: Descriptions and explanations of public deservingness opinions.
- Vande Lanotte, J. (2002a). Algemene omzendbrief betreffende de wet van 26 mei 2002 betreffende het recht op maatschappelijke integratie. Bestuur van de Maatschappelijke Integratie.
- Vande Lanotte, J. Wet van 26 mei 2002 betreffende het recht op maatschappelijke integratie (2002).
- Vandenbroucke, F., Luigjes, C., & Lievens, K. (2016). Activering en bijstand na de staatshervorming: buitenlandse lessen, uitdagingen en kansen. In M. De Wilde, B. Cantillon, F. Vandenbroucke, & M. De Bie (Red.), *40 jaar OCMW en bijstand* (pp. 184–206). Leuven: Acco.
- Versailles, P. (2008). *Het bestaansminimum en de maatschappelijke dienstverlening door de rechtspraak van het jaar 2006* (Verslag opgesteld op aanvraag van de Minister van sociale Integratie). Namur: Centre Droits fondamentaux & Lien social F.U.N.D.P.
- Wallander, L. (2009). 25 years of factorial surveys in sociology: A review. *Social Science Research*, 38(3), 505–520.
- Wallander, L. (2012a). Measuring social workers' judgements: Why and how to use the factorial survey approach in the study of professional judgements. *Journal of Social Work*, 12(4), 364–384.
- Wallander, L. (2012b). Measuring social workers' judgements: Why and how to use the factorial survey approach in the study of professional judgements. *Journal of Social Work*, 12(4), 364–384.

- Wallander, L., & Blomqvist, J. (2008). Modeling Ideal Treatment Recommendations: A Factorial Survey of Swedish Social Workers' Ideal Recommendations of Inpatient or Outpatient Treatment for Problem Substance Users. *Journal of Social Service Research, 35*(1), 47–64.
- Wastell, D., White, S., Broadhurst, K., Peckover, S., & Pithouse, A. (2010). Children's services in the iron cage of performance management: street-level bureaucracy and the spectre of Svejksism. *International Journal of Social Welfare, 19*(3), 310–320.
- Watkins-Hayes, C. (2009). *The New Welfare Bureaucrats*. Chicago: The University of Chicago Press.
- Watkins-Hayes, C. (2011). Race, Respect, and Red Tape: Inside the Black Box of Racially Representative Bureaucracies. *Journal of Public Administration Research and Theory, 21*(suppl\_2), i233–i251.
- Watts, B. (2014). Homelessness, Empowerment and Self-reliance in Scotland and Ireland: The Impact of Legal Rights to Housing for Homeless People. *Journal of Social Policy, 43*(4), 793–810.
- Webster, S. W., O'Toole, R., O'Toole, A. W., & Lucal, B. (2005). Overreporting and underreporting of child abuse: Teachers' use of professional discretion. *Child Abuse & Neglect, 29*(11), 1281–1296.
- Weiss, & Gal, J. (2007). Poverty in the Eyes of the Beholder: Social Workers Compared to Other Middle-Class Professionals. *The British Journal of Social Work, 37*(5), 893–908.
- Weiss, Gal, J., & Cnaan, R. A. (2004). Social work education as professional socialization: A study of the impact of social work education upon students' professional preferences. *Journal of Social Service Research, 31*(1), 13–31.
- Weiss, I. (2005). Is there a global common core to social work? A cross-national comparative study of BSW graduate students. *Social Work, 50*(2), 101–110.
- Weiss-Gal, I. (2008). The person-in-environment approach: Professional ideology and practice of social workers in Israel. *Social Work, 53*(1), 65–75.
- Weiss-Gal, I., Benyamini, Y., Ginzburg, K., Savaya, R., & Peled, E. (2009). Social workers' and service users causal attributions for poverty. *Social Work, 54*(2), 125–133.
- Weiss-Gal, I., & Gal, J. (2008). Social workers and policy-practice: The role of social and professional values. *Journal of Social Service Research, 34*(4), 15–27.



### Appendix 1. Vignette example

This appendix shows an example of a vignette, the questions (dependent variables) asked about that vignette and the different categories per attribute / characteristic in the vignette. The vignettes are experimentally varied, with one category from each attribute in each vignette, in order to produce unique vignettes.

Vignette (The words or words groups between brackets are the attributes that change from vignette to vignette):

X is a 22-year-old single [woman]. She has no income or savings and no debts.

X [and her parents were born in Belgium]. She [speaks Dutch well]. She has [secondary vocational education qualifications].

X [lives in an apartment in your municipality which she rents at a reduced price because of a government subsidy. She has enough space]. She has no contact with her parents and says there isn't really anyone she can count on. She has [no children].

X has no physical health problems [and no known mental problems]. She has [no known addiction problems]. [X's childhood was marked by violence, abuse and poverty.] [She has spent two months in prison for repeated shoplifting.]

X [wants to work]. She [does not yet have any work experience]. She [does not yet have any experience with activation projects]. She [has thus far been very conscientious about keeping appointments with your agency].

Question 1: What is the likelihood that X would receive social assistance in the office where you are working? (1-7 scale)

Question 2: What is the likelihood – in the office where you are working – that X would lose social assistance if refusing an activation or job offer that

- started at 5 o'clock in the morning? (1-7 scale)
- only last one day (1-7 scale)
- is not in line with the own qualifications (1-7 scale)

## Appendix 2. Categories of the 14 attributes included in the vignettes

Attribute	Category1	Category2	Category2
Gender	Female	Male	
Nationality background	Client and parents born in Belgium	Client born in Belgium – parents born in Morocco	Syrian refugee – one year in Belgium. Registered
Command of the national language	Good	Limited	Poor
Level of education	Primary education	Secondary education	Bachelor's degree
Parenthood	No children	One healthy child aged two years	One sick child aged two years (immune disorder)
Housing situation	Stable housing in a subsidized apartment	Lives with a friend – lots of quarrels	Homeless – staying with various friends
Mental health	Good	Depressed – no treatment	Lack of intellectual ability – not diagnosed by a doctor
Drug addiction	None	Slightly addicted – refuses treatment	Severely addicted – refuses treatment
Adverse life experiences	<i>No information included</i>	Youth with violence and abuse	Lost a baby one year ago
Alternative ways of earning money	<i>No information included</i>	Been in prison for shoplifting	Has been prostituting her/himself
Aspirations	Wants to work	Wants to study	Is not motivated to work or study
Employment experience	No experience	Positive experience – made redundant	Negative experience – quarrels with superior
Activation experience	No experience	One negative experience – not diligent	Several negative experiences – not diligent
Attitude	Diligent	Missed one appointment	Missed several appointments

Note: the sick child had an immune system disorder, the symptoms of the illness were unpredictable, or no sufficient medical assistance was available, which made working difficult (knowing that the client had no network to fall back on); the mental health problems had not been diagnosed by a doctor; the drug-addicted clients refused help; or the deceased child (adverse life experience) had been born with heart problems.

## Appendix 3. Respondent sample procedure

We divided the 306 Flemish municipalities into 38 strata based on number of inhabitants, relative number of benefit recipients and a socioeconomic index (Belfius, 2007). Drawing a stratified sample from these groups, we selected around 90 municipalities to ask to participate in the first wave. When a municipality declined to participate, we drew another municipality from the same stratum. In total, we invited 105 municipalities, 15 of which declined to participate mainly due to time constraints and one that did not provide enough responses to be included in the research. In the selected municipalities we surveyed professionals who were both in direct contact with clients and who were involved in the decision-making process concerning eligibility to social assistance<sup>41</sup>. In each stratum we chose one municipality at random in which all employees involved in the decision-making process were invited to participate. In the other municipalities, we invited 25% of the employees stratified by gender, age and job

<sup>41</sup> This means that professionals who were only responsible for activation trajectories or for debt counselling were not included in the survey.

title (management position or not). If a professional declined to participate, one of her colleagues with similar characteristics was asked to participate instead, as we aimed to include 600 professionals. In total, 839 social workers were invited, of which 582 provided a completed questionnaire.

#### Appendix 4. Overview and descriptive statistics of variables at the three levels

		Frequency	Mean	Min-Max
<b>Level 1 = Client level (n=4785)</b>		All client characteristics = 33%		
<b>Level 2 = Professional level (n=582)</b>				
Age			38	22-64
Gender (female)		Male	20%	
Ethnicity (both parents born in B)		Both born in Europe	3%	
		At least one born outside of Europe	3%	
Housing (tenant)		House owner	79%	
Qualification (Bachelor's degree)		Master's	11%	
Experience with long-term unemployment (no)		Yes	44%	
Seniority			12	0-40
Work regime in percentages			90	33-100
Specialisation (number of tasks) (1)		2	41%	
		3 or 4	28%	
Decision_eligibility			0	-2.9-2.2
Decision_activation			0	-1.2-0.9
Board follows advice of professional (1-7)			5.82	3-7
Involvement of board in decision making (1-7)			6.32	1-7
Primary cause of unemployment (other)		Laziness	7%	
Political orientation (1: left – 10 : right)			4.23	1-10
Responsibility of welfare of citizens (1: individual – 10 : state)			5.53	1-10
WS Criticism			0	-1.3-2.0
WS Strictness			0	-1.3-0.9
WS Overuse			0	-1.4-1.6
<b>Level 3 = municipality/social assistance agency (n=89)</b>				
Percentages of social assistance receivers			0.56	0.1-2.2
Size municipality			30442.7	2110-502604
Poverty				
Unemployment rate			7.1	3.4-16.3
Welfare index			108.1	88-139
Percentage inhabitants with foreign background				
Job centre in the same municipality (no)		Yes	38.9%	
Subsidised psycho-social organisation in the same municipality (no)		Yes	24.4%	
Political party of the chair of the board of the agency (Nationalist right wing party (N-VA))		Centrist Christian Democratic party (CD&V)	47.8%	
		Centre-right liberal party (Open VLD)	12.2%	
		Centre-left socialist party (Sp.a)	10%	
		Other	15.6%	
WS Criticism_munic			-0.0	-1.2-0.7
WS Strictness_munic			0.0	-0.6-0.6
WS Overuse_munic			0.0	-0.6-0.6
Eligibility_munic			0.2	-2.9-2.2

Activation_munic			0.1	-0.7-0.9
Specialisation (average number of tasks) (1)	2	39.3%		
	3	13.5%		
	No clear pattern	25.8%		
Average age social workers			38.4	25-59.3
Percentage male professionals			17.3	0-100
Percentage professionals with children			67.6	0-100
Average seniority			11.7	0.5-35.5

## Appendix 5. Construction of latent attitude variables

The construction of three latent variable attitudes, used in several chapters (Chapter 2, Chapter 3, Chapter 5 and Chapter 6), is as follows: ‘perceived negative consequences of the welfare state (*Welfare State Criticism*)’, ‘opinion on the extent to which uncooperative social assistance recipients should be monitored or punished (*Welfare State Strictness*)’ and ‘opinion on over use of benefits (*Welfare State Over-use*)’. Latent variables are based on confirmatory factor analyses on different Likert-scale (1 to 5) variables reflecting the respondent’s answers on the below questions.

The latent variable *Welfare State Criticism* including four items (1-5 Likert scale), namely:

- The welfare state costs too much money compared to what it yields (*costs*).
- The welfare state makes people irresponsible and lazy (*lazy*).
- The welfare state is too much of a hammock that people become dependent on (*dependent*).
- The welfare state causes people to no longer be able to take care of themselves (*no self*).

The second latent variable, *Welfare State Strictness*, was expected to influence the following items (1-5 Likert scale):

- The government is too strict on social benefit recipients (*to strict*).
- The government should check more closely whether the unemployed are applying for jobs sufficiently (*contr job*).
- Social benefit beneficiaries who do not do what is required of them should be punished more harshly (*more puni*).
- The government should check more closely whether the unemployed do additional illicit work (*control il*).

Finally, *Welfare State Over-use* is the latent construct capturing perceptions about how often (1-5 Likert scale) the following situations occur:

- People use their health insurance although they are not sick (*sick*).



- People receive unemployment benefits although they could get a job if they wanted (*unemploy*).
- People receive a living wage (minimum income) although they are not actually poor (*welfare*).

## Appendix 6. Correlations among attitude items Chapter 2

Correlations among attitude items for the general population (N=871)

	costs	lazy	dependen	noself	toostric	contrjob	morepuni	contrill	sick	unemploy	welfare
costs	1										
lazy	<b>0.40</b>	1									
dependen	<b>0.36</b>	<b>0.55</b>	1								
noself	<b>0.34</b>	<b>0.47</b>	<b>0.50</b>	1							
toostric	-0.07	-0.19	-0.14	-0.09	1						
contrjob	0.09	0.21	0.25	0.18	-0.23	1					
morepuni	0.16	0.25	0.24	0.21	-0.25	<b>0.52</b>	1				
contrill	0.08	0.16	0.16	0.18	-0.11	<b>0.34</b>	<b>0.37</b>	1			
sick	0.15	0.23	0.29	0.23	-0.10	0.23	0.27	0.12	1		
unemploy	0.16	0.21	0.22	0.17	-0.16	0.25	0.27	0.15	<b>0.57</b>	1	
welfare	0.15	0.22	0.24	0.26	-0.10	0.25	0.23	0.13	<b>0.50</b>	<b>0.49</b>	1

Correlations among attitude items for the social workers (N=603)

	costs	lazy	dependen	noself	toostric	contrjob	morepuni	contrill	sick	unemploy	welfare
costs	1										
lazy	<b>0.52</b>	1									
dependen	<b>0.50</b>	<b>0.56</b>	1								
noself	<b>0.41</b>	<b>0.57</b>	<b>0.58</b>	1							
toostric	-0.18	-0.27	-0.26	-0.20	1						
contrjob	0.28	<b>0.32</b>	<b>0.39</b>	0.27	-0.45	1					
morepuni	<b>0.34</b>	<b>0.34</b>	<b>0.40</b>	0.29	-0.43	<b>0.58</b>	1				
contrill	0.17	0.19	0.20	0.16	-0.25	<b>0.43</b>	<b>0.39</b>	1			
sick	0.19	0.22	0.25	0.26	-0.26	0.29	<b>0.31</b>	0.23	1		
unemploy	<b>0.30</b>	<b>0.32</b>	<b>0.44</b>	0.29	-0.36	<b>0.46</b>	<b>0.43</b>	0.26	<b>0.38</b>	1	
welfare	0.23	0.27	0.29	0.19	-0.22	0.20	<b>0.30</b>	0.19	<b>0.38</b>	0.27	1

## Appendix 7. Results SEM-models Chapter 2

All estimates of the measurement part of the model, full sample

Measurement M	Model 1				Model 2			
	Coefficients		Intercepts		Coefficients		Intercepts	
	SW	GP	SW	GP	SW	GP	SW	GP
<b>WS criticism -&gt;</b>								
Dependen	0.73 (0.0)***	0.7 (0.0)***	2.74 (0.0)***	2.56 (0.1)***	0.73 <i>Constr.</i>	0.7 <i>Constr.</i>	2.74 <i>Constr.</i>	2.56 <i>Constr.</i>
Lazy	0.61 (0.0)***	0.73 (0.0)***	2.22 (0.0)***	2.29 (0.1)***	0.62 (0.0)***	0.74 (0.0)***	2.22 (0.0)***	2.28 (0.0)***
Noself	0.55 (0.0)***	0.57 (0.0)***	2.31 (0.0)***	2.25 (0.1)***	0.56 (0.0)***	0.58 (0.0)***	2.31 (0.0)***	2.25 (0.0)***
Costs	0.59 (0.0)***	0.52 (0.0)***	2.42 (0.0)***	2.58 (0.1)***	0.58 (0.0)*	0.53 (0.0)***	2.42 (0.0)***	2.58 (0.0)***
<b>WS sanctioning -&gt;</b>								
Contrjob	0.74 (0.0)***	0.66 (0.1)***	3.54 (0.0)***	3.57 (0.0)***	0.74 <i>Constr.</i>	0.66 <i>Constr.</i>	3.54 <i>Constr.</i>	3.57 <i>Constr.</i>
Morepuni	0.72 (0.0)***	0.77 (0.0)***	3.16 (0.0)***	3.35 (0.1)***	0.72 (0.0)***	0.76 (0.1)***	3.17 (0.0)***	3.36 (0.1)***
Contrill	0.44 (0.0)***	0.53 (0.1)***	3.93 (0.0)***	3.61 (0.1)***	0.44 (0.0)***	0.53 (0.1)***	3.94 (0.0)***	3.61 (0.1)**
Toostric	-0.46 (0.0)***	-0.34 (0.1)***	2.54 (0.0)***	2.59 (0.1)***	-0.46 (0.0)***	-0.34 (0.1)***	2.53 (0.0)***	2.59 (0.0)***
<b>WS overuse -&gt;</b>								
Sick	0.41 (0.0)***	0.41 (0.0)***	3.03 (0.0)***	3.04 (0.0)***	0.41 <i>Constr.</i>	0.41 <i>Constr.</i>	3.03 <i>Constr.</i>	3.04 <i>Constr.</i>
Unemploy	0.51 (0.0)***	0.39 (0.0)***	3.44 (0.0)***	3.38 (0.0)***	0.53 (0.1)***	0.4 (0.0)***	3.46 (0.0)***	3.38 (0.0)***
Welfare	0.31 (0.0)***	0.37 (0.0)***	2.43 (0.0)***	2.76 (0.1)***	0.31 (0.0)***	0.37 (0.0)***	2.44 (0.0)***	2.76 (0.0)***
<b>Latent means</b>								
Mean (WS criticism)			0 <i>Constr.</i>	0.69 (0.2)***			0 <i>Constr.</i>	0.37 (0.1)**
Mean (WS sanctioning)			0 <i>Constr.</i>	0.65 (0.1)***			0 <i>Constr.</i>	0.5 (0.1)***
Mean (WS overuse)			0 <i>Constr.</i>	1.15 (0.1)***			0 <i>Constr.</i>	1.11 (0.2)***
N	603	871	603	871	603	871	603	871

Model 1 = Alignment model without explanatory variables

Model 2 = Alignment within CFA with explanatory variables

All estimates of the measurement part of the model, restricted sample

Measurement M	Model 3				Model 4			
	Coefficients		Intercepts		Coefficients		Intercepts	
	SW	GP	SW	GP	SW	GP	SW	GP
<b>WS criticism -&gt;</b>								
Dependen	0.73 (0.0)***	0.69 (0.0)***	2.74 (0.0)***	2.55 (0a)***	0.73 <i>Constr.</i>	0.69 <i>Constr.</i>	2.74 <i>Constr.</i>	2.55 <i>Constr.</i>
Lazy	0.61 (0.0)***	0.75 (0.1)***	2.22 (0.0)***	2.27 (0.1)***	0.62 (0.0)***	0.76 (0.1)***	2.22 (0.0)***	2.28 (0.0)***
Noself	0.55 (0.0)***	0.58 (0.0)***	2.31 (0.0)***	2.22 (0.1)***	0.56 (0.0)***	0.58 (0.1)***	2.26 (0.1)***	2.22 (0.1)***
Costs	0.59 (0.0)***	0.54 (0.1)***	2.42 (0.0)***	2.52 (0.1)***	0.61 (0.0)***	0.55 (0.1)***	2.42 (0.0)***	2.51 (0.1)***
<b>WS sanctioning -&gt;</b>								
Contrlol	0.74 (0.0)***	0.71 (0.1)***	3.54 (0.0)***	3.48 (0.1)***	0.74 <i>Constr.</i>	0.71 <i>Constr.</i>	3.54 <i>Constr.</i>	3.48 <i>Constr.</i>
Morepuni	0.72 (0.0)***	0.76 (0.1)***	3.16 (0.0)***	3.32 (0.1)***	0.72 (0.0)***	0.73 (0.1)***	3.17 (0.0)***	3.34 (0.1)***
Contrill	0.44 (0.0)***	0.48 (0.1)***	3.93 (0.0)***	3.61 (0.1)***	0.44 (0.0)***	0.47 (0.1)***	3.94 (0.0)***	3.62 (0.1)***
Toostric	-0.46 (0.0)***	-0.38 (0.1)***	2.54 (0.0)***	2.52 (0.0)***	-0.47 (0.0)***	-0.38 (0.1)***	2.54 (0.0)***	2.52 (0.1)***
<b>WS overuse -&gt;</b>								
Sick	0.41 (0.0)***	0.41 (0.0)***	3.03 (0.0)***	3.04 (0.0)***	0.41 <i>Constr.</i>	0.41 <i>Constr.</i>	3.03 <i>Constr.</i>	3.04 <i>Constr.</i>
Unemploy	0.51 (0.0)***	0.41 (0.1)***	3.44 (0.0)***	3.38 (0.1)***	0.53 (0.1)***	0.43 (0.0)***	3.43 (0.0)***	3.37 (0.1)***
Welfare	0.31 (0.0)***	0.37 (0.1)***	2.43 (0.0)***	2.70 (0.1)***	0.31 (0.0)***	0.37 (0.0)***	2.43 (0.0)***	2.70 (0.1)***
<b>Latent means</b>								
Mean (WS criticism)			0 <i>Constr.</i>	0.56 (0.1)***			0 <i>Constr.</i>	0.58 (0.5)
Mean (WS sanctioning)			0 <i>Constr.</i>	0.59 (0.2)***			0 <i>Constr.</i>	1.33 (0.4)***
Mean (WS overuse)			0 <i>Constr.</i>	0.80 (0.1)***			0 <i>Constr.</i>	2.3 (0.8)**
N	603	364	603	364	603	364	603	364

Model 3 = Alignment model without explanatory variables with only respondents with a Bachelor's or Master's degree

Model 3 = Alignment within CFA with explanatory variables with only respondents with a Bachelor's or Master's degree

## Appendix 8. Results for other dependent variables in Chapter 3

The effects on the likelihood of being sanctioned when refusing a job offer that lasts one day

Likelihood of being sanctioned when refusing a job offer that lasts one day	Model1		Model2		Model3	
	Coef	% of total	Coef	compared to M1	Coef	compared to M1
<b>Variances</b>						
Level 3 = municipality	75.2**	8.9%	69.3**	-7.6%	8.7	-88.4%
Level 2 = professional	489.1***	57.9%	409.3***	-16.3%	367.7***	-24.8%
Level 1 = client	280.6***	33.2%	253.2***	-9.6%	252.8***	-9.9%
Total	844.9	100%	721.7	-14.57%	628.2	-25.64%
	<b>Coef</b>	<b>z-score</b>	<b>Coef</b>	<b>z-score</b>	<b>Coef</b>	<b>z-score</b>
<b>Intercept</b>	60.5***	31.7	56.6***	22.6	66.0***	12.2
<b>Level 1 = client characteristics</b>						
<b>Aggravating characteristics</b>						

	Not motivated to work (ref = wants to work)			-0.24	-0.4	-0.26	-0.44
	1 time late for an appointment (ref = diligent)			-2.15*	-2.06	-2.22*	-2.13
	Several times late for an appointment (ref = diligent)			-0.16	-0.16	-0.16	-0.15
	Negative work experience (ref = no work experience)			0.0	0.0	0.01	0.02
	Positive work experience (ref = no work experience)			0.36	0.58	0.35	0.58
	1 negative activation exp (ref = no activ exp.)			0.02	0.03	-0.01	-0.01
	Several negative activation exp (ref = no activ exp.)			0.05	0.08	0.06	0.1
	2 <sup>nd</sup> work refusal (ref = 1 <sup>st</sup> )			16.58***	6.65	15.93***	6.68
	3 <sup>rd</sup> work refusal (ref = 1 <sup>st</sup> )			25.47***	10.25	26.27***	11.05
<b>Mitigating characteristics</b>							
	1 healthy child of 2 years (ref = no children)			-1.31	-1.25	-1.23	-1.07
	1 sick child of 2 years (ref = no children)			-4.76***	-4.01	-5.35***	-4.59
	Unstable housing (ref = stable housing)			-0.73	-1.22	-0.72	-1.2
	Homeless (ref = stable housing)			-1.86**	-3.08	-1.91***	-3.18
	Depressed (ref = no mental problems)			-0.19	-0.18	-0.23	-0.22
	Mental health problems (ref = no mental problems)			-1.88	-1.77	-1.84	-1.73
	Beginning addiction (ref = no addiction)			-0.36	-0.61	-0.34	-0.58
	Severe addiction (ref = no addiction)			-5.89***	-9.69	-5.91***	-9.74
	Lost a child (ref = no special life exp mentioned)			-4.76***	-4.56	-4.79***	-4.6
	Violent upbringing (ref = no special life exp)			-3.07**	-2.94	-3.18**	-3.05
<b>Interaction between aggravating and mitigating characteristics</b>							
	2 <sup>nd</sup> refusal * healthy child			-1.79	-1.22	-1.67	-1.15
	2 <sup>nd</sup> refusal * sick child			-4.71**	-3.12	-4.71**	-3.13
	3 <sup>rd</sup> refusal * healthy child			-0.81	-0.54	-0.66	-0.45
	3 <sup>rd</sup> work refusal * sick child			-3.6*	-2.5	-3.3*	-2.2
	2 <sup>nd</sup> work refusal * depression			-4.1**	-2.8	-3.9**	-2.7
	2 <sup>nd</sup> work refusal * mental health problems			-1.3	-0.9	-1.3	-0.9
	3 <sup>rd</sup> work refusal * depression			-1.1	-0.4	-1.1	-0.8
	3 <sup>rd</sup> work refusal * mental health problems			-1.5	-1.0	-1.6	-1.1
<b>Level 2 = Professional level</b>							
	Parenthood cm (ref = no)	Yes				0.3	0.1
	Parenthood client * parenthood cm	Client healthy child * cm child(ren)				0.4	0.3
		Client sick child * cm child(ren)				-2.8*	-2.2
	Involvement board with activation					1.8***	3.2
	Specialisation (ref = 1 task of intake, treatment, activation or management)	2 tasks				7.2***	3.4
		3 or 4 tasks				2.5	1.1
	WS Strictness					10.3***	4.2
	WS Over-use					6.2**	2.9

Level 3 = Municipalities							
	% inhabitants with foreign background					0.0*	2.2
	Availability of a job centre (VDAB) (ref. = no)					-10.8***	-4.0
	Party of the chair of the board (ref = nationalist right-wing party, N-VA)	Centrist Christian Democratic party (CD&V)				-5.6	-1.7
		Centre-right liberal party (Open VLD)				-14.7***	-3.2
		Centre-left socialist party (Sp.a)				-11.2**	-3.1
		Other				-5.0	-1.1
	Mean age professionals					0.5*	2.1
	% professionals with children					-0.2*	-2.4
	Specialisation in the organisation (ref = most cm's 1 task)	Most cm's two tasks				-6.3*	-2.3
		Most cm's three or four tasks				-12.4**	-2.6
		No clear specialisation pattern in the agency				-6.0	-1.7
N	Municipalities	89		89		89	
	Professionals	582		582		582	
	Vignettes	4785		4785		4785	

The effect on the likelihood of being sanctioned when refusing a job offer that is not in line with the client's qualifications

		Model1		Model2		Model3	
		Coef	% of total	Coef	compared to M1	Coef	compared to M1
<b>Variances</b>							
	Level 3 = municipality	42.7**	6.1%	45.8*	-7.12%	0	-100%
	Level 2 = professional	373.0***	53.1%	348.5***	-6.57%	316.35***	-15.19%
	Level 1 = client	297.4***	42.3%	269.4***	-9.43%	269.33***	-9.45%
	Total	703.2	100%	663.7	-5.62%	585.68	-16.71%
		<b>Coef</b>	<b>z-score</b>	<b>Coef</b>	<b>z-score</b>	<b>Coef</b>	<b>z-score</b>
<b>Intercept</b>		63.6***	31.7	66.9***	22.55	47.2***	6.3
<b>Level 1 = client characteristics</b>							
<b>Aggravating characteristics</b>							
	Not motivated to work (ref = wants to work)			0.4	0.58	0.3	0.6
	One time late for an appointment (ref = diligent)			-1.0	-1.68	-1.0	-1.7
	Several times late for an appointment (ref = diligent)			0.5	0.79	0.5	0.8
	Negative work experience (ref = no work experience)			0.3	0.4	0.3	0.4
	Positive work experience (ref = no work experience)			0.8	1.2	0.8	1.3
	1 negative activation exp (ref = no activ exp.)			-0.1	-0.2	-0.1	-0.2
	Several negative activation			-0.58	-1.0	-0.6	-1.0

	exp (ref = no activ exp.)						
	2 <sup>nd</sup> work refusal (ref = 1 <sup>st</sup> )			5.56*	2.3	5.4*	2.4
	3 <sup>rd</sup> work refusal (ref = 1 <sup>st</sup> )			14.98***	6.4	15.7***	6.9
	<b>Mitigating characteristics</b>						
	One healthy child of 2 years (ref = no children)			-2.53***	-4.1	-2.6***	-4.2
	One sick child of 2 years (ref = no children)			-6.79***	-10.9	-6.8***	-11.0
	Unstable housing (ref = stable housing)			-0.24	-0.4	-0.3	-0.5
	Homeless (ref = stable housing)			-2.0***	-3.2	-2.1***	-3.4
	Depressed (ref = no mental problems)			-4.0***	-3.9	-4.0***	-3.9
	Mental health problems (ref = no mental problems)			-5.29***	-5.0	-5.3***	-5.0
	Beginning addiction (ref = no addiction)			-0.04	-0.1	-0.0	-0.1
	Severe addiction (ref = no addiction)			-6.34***	-5.0	-6.4***	-10.2
	Lost a child (ref = no special life exp mentioned)			-2.28*	-2.2	-2.3*	-2.1
	Violent & abusive upbringing (ref = no special life exp)			-0.9	-0.9	-0.9	-0.9
	<b>Interaction between aggravating and mitigating characteristics</b>						
	2 <sup>nd</sup> work refusal * lost a child			3.06*	2.0	2.9*	1.9
	2 <sup>nd</sup> work refusal * violent upbringing			0.94	0.6	1.0	0.6
	3 <sup>rd</sup> work refusal * lost a child			-1.66	-1.1	-1.8	-1.2
	3 <sup>rd</sup> work refusal * violent upbringing			-2.48	-1.7	-2.6	-1.7
	2 <sup>nd</sup> work refusal * depression			3.58*	2.4	3.5*	2.3
	2 <sup>nd</sup> work refusal * mental health problems			1.93	1.3	1.9	1.3
	3 <sup>rd</sup> work refusal * depression			0.35	0.2	0.4	0.2
	3 <sup>rd</sup> work refusal * mental health problems			2.83	1.9	2.7	1.8
	<b>Level 2 = Professional level</b>						
	<b>Job characteristics</b>						
	Involvement board with activation					1.5***	3.4
	Specialisation (ref = 1 task of intake, treatment, activation or management)	2 tasks				4.2*	2.2
		3 or 4 tasks				-0.0	-0.0
	<b>Ideology</b>						
	WS Strictness					11.1***	5.0
	WS Over-use					5.9**	2.9
	Perceived strictness of the agency					4.6**	3.0
	<b>Level 3 = Municipalities</b>						
	<b>Socio-demographic characteristics</b>						
	Availability of a job centre					-8.0***	-4.0

	(VDAB) (ref. = no)						
	<b>Ideology</b>						
	Party of the chair of the board (ref = nationalist right-wing party, N-VA)	Centrist Christian Democratic party (CD&V)					-3.8
		Centre-right liberal party (Open VLD)					-7.6*
		Centre-left socialist party (Sp.a)					-10.1**
		Other					-6.4
	<b>Organisational characteristics</b>						
	Mean seniority professionals						0.5*
N	Municipalities	89		89		89	
	Professionals	582		582		582	
	Vignettes	4785		4785		4785	

### Appendix 9. Results of multi-level fixed effects models in Chapter 5

Model 1		Likelihood to receive a benefit	Likelihood to keep benefit after activation refusal(s)
<b>Variances</b>			
	Level 1 = client	162.27***	419.24***
	Level 2 = social worker	143.92***	218.55***
	Level 3 = municipality/organization	0	0
<b>Intercept</b>			
		90.1***	42.43***
<b>Level 1 = client characteristics</b>			
Frequency of activation refusal	Two (ref = one)		-10.52***
	Three (ref = one)		-18.46***
Type of activation offer	Day contract (ref = early morning)		-12.54
	Not in line with qualification (ref = early morning)		-15.8
Gender	Man (ref = woman)	-1.86***	-1.71***
Nationality background	Second generation migrant (ref = cl and parents born in Belgium)	-1.74*	0.62
	Refugee (ref = cl and parents born in Belgium)	0.39	-0.88
Language	Speaks Dutch poorly (ref = speaks good Dutch)	0.05	-0.25
	Speaks no Dutch (ref = speaks good Dutch)	-1.73*	1.44
Interaction nationality * language	Second generation migrant who speaks Dutch poorly	0.65	0.17
	Second generation migrant who speaks no Dutch	3.07**	-0.84
	Refugee who speaks Dutch poorly	-0.0	2.05
	Refugee who speaks no Dutch	1.21	1.98
Parenthood	One healthy child (ref = none)	1.39**	6.31***
	One sick child (ref = none)	1.77***	12.02***
Qualifications	Qualification secondary education (ref = primary education)	-0.32	-0.73
	Bachelor's degree (ref = qualification primary education)	-2.36***	-2.79***

Housing situation	Unstable housing (ref = stable housing)	-1.41**	
	Homeless (ref = stable housing)	-2.66***	
Mental health	Depressed (ref = no mental problems)	-0.62	2.54***
	Mental disability (ref = no mental problems)	-0.01	3.72***
Drug addiction	Slight addiction (ref = none)	-0.87	0.2
	Severe addiction (ref = none)	-1.52***	5.52***
Adverse life events	Violent & abusive upbringing (ref = no adverse life events)	0.83	1.45***
	Lost a child (ref = no adverse life events mentioned)	1.14	2.47***
Aspirations	Wants to study (ref = willing to work)	-3.32***	1.46***
	Not motivated to work (ref = willing to work)	-7.3***	2.47***
Diligence	Missed one appointment (ref = diligent)	-2.0***	0.89*
	Missed several appointments (ref = diligent)	-3.93***	-0.57
Work experience	Positive work experience (ref = none)	0.66	-0.73
	Negative work experience (ref = none)	-1.44**	-0.31
Activation experience	One negative activation experience (ref = none)	-0.35	0.24
	Several negative activation exp (ref = none)	-0.79	0.49
Interaction aspirations * mental health	Wants to study and depression	2.35*	
	Wants to study and mental disability	2.3*	
	Not motivated and depression	1.15	
	Not motivated and mental disability	1.46	
Interactions aspirations * adverse life events	Wants to study and violent & abusive upbringing	-0.23	
	Wants to study and lost a child	-0.39	
	Not motivated and violent & abusive upbringing	2.95**	
	Not motivated and lost a child	3.39**	
<b>Level 2 = professional characteristics</b>			
Seniority		0.2**	0.15
Parenthood	Yes (ref = no)		-0.08
Parenthood client * parenthood professional	Cl healthy child * Pr child(eren)		0.32
	Cl sick child * Pr child(eren)		2.82**
Housing situation	Owner (ref = tenant)	2.52	
Housing situation client * housing situation professional	Cl unstable housing * Pr owner	-0.45	
	Cl homeless * Pr owner	-2.12	
Specialisation (ref = 1 task of intake, treatment, activation or management)	2 tasks	-1.47	-5.09**
	3 or 4 tasks	2.84	-1.96



Advise_professional		3.62***	
Eligibility_board		1.09*	
Eligibility			-1.48*
Activation_board			-1.05**
Activation		2.97	
Appreciation strictness agency			-2.3
Welfare state criticism			-3.38*
Welfare state strictness		-4.82***	-8.26***
Welfare state overuse			-5.71***
<b>Level 3 = agency characteristics</b>			
Inhabitants		-0.0*	0.0***
Quadratic term of inhabitants			-0.0**
Relative number beneficiaries		5.84***	-8.44**
Relative number beneficiaries in activation trajectory			-0.19
Availability of a job centre (VDAB)	Yes (ref. = no)		8.27
Party of the chair of the board (ref = nationalist right-wing party, N-VA)	Centrist Christian Democratic party (CD&V)	3.51	4.72
	Centre-right liberal party (Open VLD)	1.13	13.13***
	Centre-left socialist party (Sp.a)	3.93	8.46*
	Other	2.73	10.11**
Specialisation in the organisation (ref = most cm's 1 task)	Most cm's two tasks	-2.63	4.28
	Most cm's three or four tasks	-2.01	17.2***
	No clear specialisation pattern in the agency	-4.41*	7.59**
Eligibility_board_municipality		-3.52*	
Activation_municipality		-9.49**	10.3*
Welfare state strictness team coordinator		4.52	-11.48**
Welfare state overuse			6.84
Mean age professionals		-0.43*	
Mean seniority professionals			-0.69**
% parenthood			0.09
% male		-0.08*	
<b>N</b>	Municipalities	76	
	Professionals	572	
	Vignettes	5139	







The consensus in social policy literature is that the study of how implementation works is key to understand how policies become actual treatment for clients. Surprisingly, the different levels that determine this implementation process (country, local setting, professional and client) are seldom studied in one research project and with respect to the expected treatment outcomes for beneficiaries. In this dissertation I used an innovative and purpose designed survey to determine the importance of the several levels.

**The experimental survey (factorial survey) allowed me to study social assistance treatment with regard to a large amount of clients (5247), social workers (583) and social assistance agencies (89).**

I used Belgium as a case. The country's decentralization, rather vague descriptions of work willingness in legislation and the explicit recognition of the case manager as a qualified professional with important interpretative discretion make Belgium an interesting case with which to look at the various determinants of eligibility. Undertaking this research has raised a number of reservations concerning the three following premises of the different research traditions which have previously been taken for granted:

**(1) Client characteristics determine the treatment clients receive.**

True, but – with regard to the dependent variables included in my research – to a lesser extent than might be expected from legislation. Client's work willingness affected eligibility only to a limited extent, but became more important later in the activation trajectory. The influence of specific difficult life circumstances on eligibility and sanction decisions was negligible.

**(2) The professional is of decisive importance for treatment.**

True, variation is considerable, and part of it is not random but can be explained by the general welfare state attitudes of social work professionals.

**(3) The level of decentralization of the policy determines treatment variation.** This could not be confirmed. Overall, local agencies did not differ extensively from each other, but insofar as they did, organizational characteristics, political majorities and socio-demographic characteristics explained the variation.

**Policy recommendations**, derived from this dissertation, focus on both the administratisation of the eligibility test and on the reevaluation of discretion in other parts of the social assistance trajectory.