

Performance Management, Caseloads and the Frontline Provision of Social Services

Rik van Berkel* and Eva Knies

School of Governance, Utrecht University, Utrecht, Netherlands

Abstract

Caseloads and performance management are important working conditions of workers delivering public social services in street-level organizations. The literature on these working conditions argues that high caseloads and performance management have considerable consequences for workers' performance in terms of the quality of services they provide and the results they realize. This article empirically investigates and compares these consequences, drawing on the results of a quantitative study of frontline workers in 14 local welfare agencies in the Netherlands. These workers are responsible for the delivery of welfare-to-work policies to social assistance recipients. The findings show that high caseloads do, indeed, have a detrimental effect on workers' performance, whereas the impact of performance management is more modest, though confirming some of the findings reported in other studies on performance management. The results also show that by focusing service provision on a proportion of their caseload, workers are able to reduce the negative impact of high caseloads somewhat. Overall, the article finds that the negative impact of high caseloads is more pervasive than that of performance management. The article concludes that the recent focus in the literature on performance management and its consequences for public services should not turn scholars' attention away from the 'traditional' public administration problem of high caseload sizes.

Keywords

Caseload; Performance management; Frontline delivery; Welfare-to-work; Social services; Street-level bureaucracy

Introduction

Civil servants' working conditions exert an important impact on how these workers actually do their work. In coping with working conditions, civil servants develop strategies and take decisions that have consequences for the services they provide and the outcomes they realize. These are important conclusions of a variety of studies that looked at how social policies and social services are implemented at the frontlines of organizations responsible for policy implementation: street-level bureaucracies or street-level organizations

Author Emails: r.vanberkel@uu.nl; e.knies@uu.nl

(Brodkin 2013; Lipsky 2010). Two working conditions have received considerable attention in the scientific literature: the size of caseloads and performance management (e.g. Jewell and Glaser 2006). Caseloads are a 'traditional' issue in street-level organizations where workers are structurally confronted with a shortage of resources (such as time, money, services) given the numbers of clients they are supposed to support. Performance management is a more recent issue that became increasingly popular in the public sector against the background of the rise of New Public Management (NPM). Since the introduction of NPM, the scientific literature has devoted considerable attention to the impact of performance management on civil servants (and especially on public professionals) (Talbot 2010). Nowadays, in comparison, attention on the 'traditional' street-level organization problem of high caseloads remains far behind. One of the reasons for this may be that performance management implies a more direct 'attack' on the discretion and autonomy of civil servants than do high caseloads. The introduction of performance management has often been explicitly framed as a political project legitimized in terms of 'curbing' the autonomy of professionals (Clarke and Newman 1997; Noordegraaf and Steijn 2013), whereas high caseloads are a less politicized issue, a 'fact of bureaucratic life'. Furthermore, critics of NPM often consider the introduction of performance management in the public sector as exemplary for NPM's damaging and, according to some, perverse impact on public services, public organizations and public workers (Diefenbach 2009).

This article takes up the 'old' public management issue of high caseloads and the 'new' public management issue of performance management by analyzing their consequences for the quality and outcomes of social services. Based on data from empirical research, we address the following research question: How are size of caseload and performance management related to the quality and outcomes of social services? The article seeks to contribute to the debate about the consequences of high caseloads and performance management for the frontline delivery of public policies and social services. In addition, it aims to contribute to the strand of research that considers conditions of work an important factor in determining how formal policies translate into policy practices (Jewell and Glaser 2006; Lipsky 2010). Changes in working conditions can have political implications (Brodkin 2013): they trigger changes in policy delivery processes that, compared to formal policy reforms, are far less transparent and subject to democratic decision-making processes but may still have a considerable impact on what policies look like in practice.

The type of social services that our empirical research focused on are services aimed at promoting unemployed people's employability and (labour-market) participation: services provided in the context of welfare-to-work or, in European parlance, activation policies. In the delivery of these policies, traditional street-level bureaucracies play an important role: benefit agencies, public employment services agencies, local welfare agencies. At the same time, the provision of welfare-to-work has been influenced by NPM in most countries (van Berkel *et al.* 2011), both in the form of corporate governance and market governance (Considine 2001). Therefore, we may expect high

caseloads as well as performance management to play a prominent role in the frontline delivery of welfare-to-work, making this a relevant context to study.

The next section explores the literature discussing caseloads and performance management and their impact on service quality and service outcomes. This exploration results in a number of hypotheses on the impact of caseloads and performance management. In the third section, our research project is described. Then, in the fourth section, we present the findings of our study. The final section discusses the findings and concludes.

Frontline Working Conditions and Performance

Caseload

The street-level bureaucracy literature has paid considerable attention to street-level bureaucrats' working conditions and the impact of these conditions on the delivery and implementation of public policies. In his famous study of street-level bureaucracies, Lipsky (2010) distinguished a variety of conditions of work structuring frontline policy implementation, such as inadequate resources, ambiguous or conflicting goal expectations, working with non-voluntary clients and, last but not least, large caseloads. Lipsky and other authors pointed out that what is problematic about high caseloads is not so much the numbers of clients that workers have *per se*, but the numbers of clients relative to workers' responsibilities and tasks: that is, workers with low caseloads may experience high workloads (Lee 2009). Because of this, various authors prefer the concept of workload to the concept of caseload. Jewell and Glaser, for example, argued that '[w]orkload encompasses not only the number of cases but also the kinds of tasks and decisions each case requires' (Jewell and Glaser 2006: 341). Some authors use an even broader definition of workload by including both client-related and non-client-related tasks (Stevens 2008). But although, theoretically, there is little reason to doubt that the concept of workload captures the work pressures frontline workers experience more adequately than the concept of caseload, measuring workload empirically is difficult, especially in quantitative research designs. Because of that, our study focuses on caseloads as a proxy of workloads; which is quite common practice in quantitative studies, as the studies discussed below illustrate.

One of the major merits of street-level bureaucracy studies is that they not only investigate how conditions of work are related to workers' performance, but also open the 'black box' of workers' agency in coping with working conditions (Brodkin and Marston 2013). Again, Lipsky's study has been path-breaking in identifying several coping strategies, such as creaming, favouritism (i.e. selecting preferred clients), parking, rubber stamping and client referrals (Lipsky 2010). Focusing on caseload size, theoretically, workers can cope with high caseloads in two ways.

First, they can decide to distribute time and resources more or less equally among their clients. Jewell pointed at this type of strategy when he argued that high caseloads threaten to reduce professionals to technicians who rely on informal shortcuts in doing their work and are unable to provide responsive

services (Jewell 2007). Thus, in this case high caseloads influence the quality of services clients receive. The strategies of rubber stamping and client referrals are examples of these shortcuts. Rubber stamping means that workers rely on judgements (e.g. client categorizations or needs assessments) of colleagues or workers in other agencies without checking their adequacy; client referrals mean that workers send clients to other service providers instead of supporting clients themselves. Several studies corroborate that high caseloads influence workers' performance. A US study of the delivery of welfare-to-work found that although the introduction of welfare-to-work policies increased workers' discretion, high caseloads restrained them from providing adequate services (Austin *et al.* 2009). Another study found a negative correlation between the size of workers' caseload and self-reported performance in various of their core activities (King *et al.* 2000). High caseloads may also increase the probability of workers making errors, such as unjust sanctions, because workers spend less time with clients and have less time for decision making (Godfrey and Yoshikawa 2012; Ryu *et al.* 2012). Lastly, workload was found to have an impact on the weight assigned to client preferences (Fraser and Estabrooks 2008).

Nevertheless, despite evidence of a detrimental impact of caseload size on workers' performance, some empirical studies revealed that the relation between caseload size and performance may be neither linear nor direct. A limited number of studies investigated the impact of caseload size on outcomes of welfare-to-work programmes (Bloom *et al.* 2003; Riccio *et al.* 1994). The study of Riccio *et al.* found no impact of caseload size on programme effectiveness, but the study by Bloom *et al.* reached the conclusion that large caseloads can undercut programme effectiveness. These conflicting results have been interpreted in two ways. First, lower caseloads do not guarantee that extra time for clients is used in effective ways (Riccio *et al.* 1994). In other words, lower caseloads may be a condition for improved performance but do not guarantee it. Second, detrimental effects of caseload size may only occur when a certain caseload size threshold is exceeded (Godfrey and Yoshikawa 2012). These interpretations echo Lipsky's caution regarding high expectations of relieving caseload pressures (Lipsky 2010).

Second, instead of distributing scarcity, workers may cope with high caseloads by deciding to focus their attention on a proportion of their clients. This strategy effectively reduces workers' 'active' caseload which may improve performance in terms of increasing the quality of services they provide and the outcomes they realize – although, as we see above, this does not necessarily have to be the case. Several studies looked at the issue of how workers select clients who are serviced and clients who are parked. Selection takes place on the basis of more or less routine client categorization processes, which distinguish between preferred and non-preferred clients (cf. favouritism), or between clients with whom targets and objectives are easy to realize and the cases considered 'difficult' (cf. creaming). The categorizations of clients do not merely reflect frontline workers' attitudes and preferences, even though these do play a role, but are *organizational* practices (Møller and Stone 2013; Rosenthal and Peccei 2006): they reflect organizational and policy priorities, the availability of services needed by various groups of clients, the available

time to provide groups of clients with adequate support, etc. What makes the selection issue politically salient is that even when this frontline strategy is effective in helping workers to improve their performance, selection, creaming and parking processes are often far from random and hit particular target groups harder than others (McDonald and Marston 2006).

Performance management

Performance management is a rather ‘fuzzy’ concept that refers to a variety of NPM and organizational practices. Diefenbach described performance management as the ‘systematic, regular and comprehensive capturing, measurement, monitoring and assessment of crucial aspects of organizational and individual performance through explicit targets, standards, performance indicators, measurement and control systems’ (Diefenbach, 2009: 894). Thus, performance management can be targeted at organizations as well as at individual workers. Organizations are subjected to performance management when they operate in competitive quasi-markets (market governance in Considine’s [2001] terminology) or when the outcomes they are expected to realize are defined, measured and evaluated with financial rewards for out-performers and financial penalties for underperformers (corporate governance, according to Considine). Performance management as an organizational human resource management strategy aims at managing the performance of workers, for example through performance appraisal or pay-for-performance systems (Boselie 2014; Weibel *et al.* 2010). The literature not always clearly distinguishes between performance management as part of governance models and as part of systems for managing workers’ performance. However, we may assume that managers whose organizations are subjected to forms of performance management will in one way or another send signals to workers about the performance expected from them, even though these signals do not necessarily include ‘tougher’ types of performance management such as pay-for-performance.

Several studies concluded that performance management is effective in steering workers’ actions in the desired direction (Riccucci *et al.* 2004; Ridzi 2004). By linking performance goals to desirable policy and organizational objectives, managers help to align frontline workers’ priorities and inform them about aspects of their work valued by the organization (cf. goal-setting theory; Locke and Latham 2002), which contributes to implementation integrity. Thus, the ‘measure what you treasure’ principle of performance management provides signals as to what workers are expected to focus on in their work. However, many authors come to more critical evaluations of performance management and point at its potential perverse consequences. There is considerable concern about the impact of performance management on workers’ discretion and their use of discretion. A frequently discussed issue is that performance management not only incentivizes workers to prioritize the goals that are measured but also discourages them to focus on other, non-measured goals (Brodkin 2008; Heinrich and Marschke 2010): under performance management workers’ discretion may be broad, but is not deep (Schram 2012). This becomes especially problematic when non-measured

goals are nonetheless valued goals: not all that is treasured is actually being measured, among others because measurement may be difficult. Johnson Dias and Maynard-Moody (2006) studied a private provider agency working under a contract that demanded the provider to offer clients case management but only rewarded job placements. This resulted in a conflict between management and workers: managers emphasized the objectives that were rewarded (i.e. job placements) whereas frontline workers adhered to the case management approach that the contract required but did not reward. Apart from the shift of focus in workers' priorities from valued to measured performance – and as we see, the two do not always coincide – studies pointed at another shift: instead of focusing on client needs, workers focus on meeting performance targets and organizational needs (Heinrich and Marschke 2010). Eventually, this may result in what has been called a performance paradox (Johnson Dias and Maynard-Moody 2006): workers' actions aimed at realizing performance targets may have a negative impact on policy practices and outcomes.

In the above, we see that high caseloads may encourage frontline workers to focus attention on part of their clients and in doing so, they develop certain selection practices. Selection processes may also take place as a consequence of performance management, although they may take different forms. For example, performance management may influence which clients are considered favourite and preferred (Rosenthal and Peccei 2006): it may, compared to high caseloads, provide stronger incentives to select clients with whom performance targets are easy to realize (Soss *et al.* 2011). So although we may expect similar types of selection processes (such as creaming, parking, favouritism) under conditions of high caseloads and under conditions of performance management, the outcomes of these processes may differ as a consequence of workers' stricter focus under performance management on measured results.

In short, performance management is likely to have consequences for the quality of services and outcomes. It will incentivize workers to provide services that are expected to realize measured outcomes while services considered unrelated to these outcomes are less likely to be provided. In addition, performance management may stimulate workers to realize quick wins as it encourages 'making the numbers' (Brodkin 2008) and 'speed over need' (Brodkin 2011) strategies. In the context of welfare-to-work, this may imply a focus on quick rather than sustainable job placements or a focus on increased sanctioning (Soss *et al.* 2011). Furthermore, workers may experience strong accountability pressures that increase the amount of time spent on administrative work, and workers may develop risk-minimizing rather than risk-taking strategies, which potentially has negative effects on service quality (Considine *et al.* 2011; McDonald and Marston 2008; van Berkel 2014). This makes clear why several authors point at the political rather than technical nature of performance management reforms (Brodkin 2006; Larsen 2013): they not merely encourage workers to do similar things in more efficient ways, they encourage workers to do different things.

Although the overall picture arising from the literature on performance management is rather gloomy, several authors emphasize that deterministic

accounts of performance management should be avoided. Several studies showed that diversity across countries regarding public management reforms and performance management systems is significant, which is why the concept of performance management regimes was introduced (Pollitt and Bouckaert 2000; Rhodes *et al.* 2012). In addition, workers respond in different ways to performance management: whereas some experience it as guiding rather than shaping their work, others experience strong pressures to meet performance objectives (Fletcher 2011; Morgen 2001). Furthermore, performance management systems are designed in various ways and system design is likely to influence the impact of performance management on frontline work (Foster and Hoggett 1999).

Conclusions and hypotheses

Based on this discussion, several conclusions can be drawn from which hypotheses about the impact of caseload size and performance management can be derived. First of all, the literature on caseloads suggests, despite several caveats concerning the linearity and directness of the relationship, that high caseloads are negatively related to workers' performance, both in terms of the quality of services they provide and in terms of the results they achieve. Therefore, our first hypotheses read as follows:

- H1a Total caseload size is negatively related to workers' results.
- H1b Total caseload size is negatively related to the quality of services workers provide.

However, one of the strategies workers can use to cope with high caseloads is to focus their attention on part of their clients: rather than servicing their total caseload, they provide services for a *reduced* caseload. This strategy intends to improve service quality and results for the clients who receive workers' attention. Therefore, the following hypotheses are formulated:

- H2a Reduced caseload size is not negatively related to workers' results.
- H2b Reduced caseload size is not negatively related to the quality of services workers provide.

Whereas our hypotheses on total caseloads imply that we expect a negative impact of total caseload size on workers' performance, the literature makes clear that performance management works in a different way. It incentivizes workers to focus on the performance targets that are being measured, whereas it discourages workers to focus on results that are not being measured, irrespective of whether non-measured results are valued or not. Although performance management often focuses on the results that workers are expected to realize, we expect a similar impact on the services that are perceived as contributing to these results. In other words, when performance management is directed at realizing result A, it will also incentivize workers to provide clients with services that they believe will make realizing result A more likely. Against this background, we formulate the following hypotheses:

- H3a Performance management on a specific result is positively related to workers' performance on that result.
- H3b Performance management on a specific result is positively related to the quality of services that contribute to realizing that result.
- H4a Performance management on a specific result is negatively related to workers' performance on other, not-measured results.
- H4b Performance management on a specific result is negatively related to the quality of services that contribute to realizing other, not-measured results.

Our final hypotheses are based on our observation that performance management and its impact on service provision received considerably more attention than did caseloads in the academic literature during the last decades. In the introduction we mentioned that the politicized nature of the debates on NPM and performance management (compared to debates about caseload sizes) might explain this. An alternative explanation is that the impact of performance management on workers' performance is more reason for concern than the impact of high caseloads. One of the ways to investigate this is to compare the impact of performance management on non-measured results (and services contributing to these results) with the impact of caseloads on the same results (and services). If the impact of performance management is indeed more reason for concern than caseload size, then we should find support for the following hypotheses:

- H5a The negative relationship between workers' performance management on a specific result and their performance on other results is stronger than the negative relationship between caseload size and workers' performance on those results.
- H5b The negative relationship between workers' performance management on a specific result and workers' quality of services directed at other results is stronger than the negative relationship between caseload size and workers' quality of service directed at those results.

Research Context and Methods

The frontline workers in our study are employed at Dutch local welfare agencies that operate under the responsibility of municipalities. These agencies are responsible for the administration of social assistance benefits as well as for providing welfare-to-work services for social assistance recipients. Our research focused on frontline workers in these agencies responsible for the delivery of welfare-to-work. Currently, about 400,000 people are receiving social assistance in the Netherlands. Over the years, ever larger proportions of the social assistance population became target groups of welfare-to-work; by now, hardly anyone dependent on social assistance who is of working age is exempted from welfare-to-work related obligations. At the same time, these obligations themselves have been strengthened, and so have sanctions (and their practical enforcement) in cases of non-compliance. However,

municipalities are not obliged to offer social assistance recipients welfare-to-work services, the only exception being young people under 27 years old.

A new social assistance Act introduced in 2004 brought radical governance reforms. Decentralization and deregulation of welfare-to-work intended to increase the room of local welfare agencies to provide tailor-made services, and delegated many decisions concerning the nature of these services to the local level. Deregulation also affected service provision models: although the agencies are allowed to outsource the provision of welfare-to-work services to external public or private providers, the last decade saw an increase of the role of in-house service production. At the same time, the 2004 Act subjected municipalities and their local welfare agencies to forms of performance management. Whereas in the past, all local social assistance expenses were paid by the Ministry responsible for social affairs, nowadays each municipality receives, annually, a fixed budget for financing social assistance payments. This provides a strong incentive for municipalities to reduce the numbers of social assistance recipients. The performance management regime allows municipalities to spend a budget surplus in whatever way they like; but when they are confronted with budget deficits they will need to use local funds to pay for the extra spending. Practically, this performance management regime incentivizes municipalities to focus on welfare-to-work services that promote labour-market integration (and independence from social assistance) rather than on forms of social participation that may support social inclusion but do not end dependency on social assistance.

Decisions concerning the organization of frontline work in local welfare agencies are to a large extent a local issue. This includes several job design and working conditions issues relevant for this article. First, it is up to the municipalities to decide whether income support and the provision of welfare-to-work services are organized in separate, specialized or in integrated tasks. Second, the municipalities are free to decide whether or not frontline workers are made responsible for the provision of welfare-to-work to specific target groups. In the Netherlands, it is common to distinguish three target groups of social assistance recipients, dependent on their estimated labour-market distance: social recipients ready to return to work; recipients with a large labour-market distance who need welfare-to-work services to make them ready for work; and recipients with a very large labour-market distance who are considered unable to return to the labour market within the next one or two years. Although the latter are obliged to participate in welfare-to-work, labour-market participation is considered a long-term rather than an immediate goal. Third, municipalities are free to decide on workers' caseloads. As we see below, workers with clients most distant from the labour market have relatively high caseloads, which reflects that municipalities prioritize servicing clients closer to the labour market even though they are not explicitly obliged to do so. As a matter of fact, this reveals a clear tension between the performance management system to which local welfare agencies are subjected, which completely focuses on labour-market entry and social assistance exit, and Dutch social assistance policies that emphasize that all social assistance recipients irrespective of their labour-market distance should be 'activated'. Put differently: welfare-to-work services that do not immediately result in

labour-market participation and social assistance exit but contribute to reducing labour-market distance and social exclusion are valued but not promoted through the performance management system. Lastly, whether or not individual performance agreements are made with workers is a local decision as well, as is the nature of the performance agreements. Not all local welfare agencies define performance targets for individual workers, but when they do, labour-market entry targets are most common.

Methods and measures

In our study, frontline workers involved in providing welfare-to-work in 14 local welfare agencies were investigated. The study focused on workers working for clients with a large or very large distance from the labour market, who are the most important target groups of welfare to work. All workers in these 14 agencies working with clients with a large or very large labour-market distance were asked to participate in the study by completing a web-based survey. The response was 52 per cent ($n = 163$). Of the respondents, 30 per cent were male, 70 per cent female. The respondents' average age was 44 years.

The concepts mentioned in the hypotheses were operationalized as follows.

Total caseload. This refers to the total number of clients for whom workers are supposed to provide welfare-to-work services. In our study, we used self-reported caseloads (a continuous variable), which is not uncommon in studies of the impact of caseload size on services quality and outcomes. Using multiple data sources could have improved data validity, but we were not able to get access to administrative data in all agencies involved in our research. In order to adjust caseload sizes for workers' weekly working hours, they have been recalculated using a 40-hour working week as standard.

Reduced caseload. This is the number of clients who are actually being served by workers, calculated by multiplying workers' total caseload with the self-reported fraction of their clients for whom they actually provided welfare-to-work services. No data were available on the characteristics of clients who are being served or parked, so that we were not able to investigate what type of selection workers make when reducing their caseload.

Performance management. This refers to results workers have to realize according to performance agreements between workers and their supervisors. Two types of performance targets are distinguished: targets concerning the number of clients who need to find a job, or targets concerning the number of clients who need to start voluntary work. In both cases, dummy variables were created (targets versus no targets).

Workers' performance. This was operationalized in two ways: the self-reported results workers realized, and the perceived quality of welfare-to-work services they provided. Three types of results were distinguished: the proportion of workers' total caseload who found a job during the year in which our survey

took place,¹ the proportion of their total caseload who started voluntary work, and the proportion of their total caseload who were sanctioned because of non-compliance with welfare-to-work related obligations. Quality of services was operationalized in terms of the self-reported diversity of support that workers offer their clients. Support could be aimed at finding a paid job and at finding voluntary work. For each of these, respondents were presented with seven types of support covering support in searching vacancies and support in mediation. The following types of support were distinguished:

- looking for job or voluntary work vacancies;
- actively contacting employers or voluntary work organizations to find out whether job or voluntary work vacancies are available or can be created for clients;
- passing information about job or voluntary work vacancies on to clients;
- supporting clients in preparing for job interviews or introductory interviews at organizations providing voluntary work;
- accompanying clients during job interviews or introductory interviews;
- other support in searching vacancies;
- other support in mediation.

Control variables. We included the client group workers work with as well as number of years of work experience in providing reintegration services as control variables in our regression analyses.

Results

Descriptives

On average, respondents in our study have 11 years of work experience in providing reintegration services. Workers' average caseload is 105 clients. However, workers' caseloads differ considerably. Workers working for clients with a large labour-market distance (25 per cent of respondents) have 77 clients on average, those working for clients with a very large labour-market distance (26 per cent of respondents) have 170 clients. Workers with a mixed client group (49 per cent) have 84 clients on average. When we look at reduced caseloads, differences in caseload size are considerably smaller: reduced caseloads are 60, 75 and 63 clients respectively, with an overall average of 65 clients. Although focusing service provision on a proportion of clients occurs irrespective of the client group workers work for, this strategy is specifically helpful in reducing caseloads of workers with the largest caseloads: in this case, workers working for clients most remote from the labour market.

Of all respondents, 76 per cent had performance agreements with their supervisors. Of these workers, 70 per cent had performance agreements concerning the number of clients who should find a job; 21 per cent had performance agreements concerning the number of clients who should start voluntary work.² The first type of performance agreements is dominant among workers working with people with a large labour-market distance or with mixed client groups, whereas the second type of performance agreements

is most common among workers working with clients with a very large labour-market distance. Performance agreements are most common among workers working with people with a large labour market distance (72 per cent), followed by workers working with a mixed client group (60 per cent). Of workers working with clients with a very large labour-market distance, 51 per cent does not have performance agreements. Combining descriptive data on caseloads and performance management we see that workers having performance agreements on the number of clients who should start paid or voluntary work have a much lower average caseload than workers without these types of performance agreements (86 versus 132).

In terms of results we found that on average, 26 per cent of workers' total caseload found a job during the year in which the survey took place ($SD = 23.5$), 17 per cent started voluntary work ($SD = 21.5$) and 9 per cent were sanctioned due to non-compliance with welfare-to-work related obligations ($SD = 10.5$). In terms of service quality, workers offer their clients on average 2.6 types of support in finding a paid job ($SD = 1.8$) and 2.0 types of support in finding voluntary work ($SD = 1.7$).

Correlation matrix

The correlation matrix of all variables in our study is displayed in Table 1.

Table 1 shows that the total caseload is negatively and significantly related to all results and service quality indicators. The total caseload is most strongly related to the quality of services aimed at finding a job ($r = -.299, p < .01$). The reduced caseload is significantly related to only two of the five performance variables. Three out of five performance variables are non-significantly related to reduced caseloads (i.e. proportion of clients who were sanctioned, and the quality of services aimed at both finding a job and finding voluntary

Table 1
Correlation matrix

Performance	Total caseload	Reduced caseload	Performance management: no. of paid jobs	Performance management: no. of clients starting voluntary work
Results: paid job	-.281**	-.177*	.263**	-.200*
Results: voluntary work	-.210*	-.173*	-.193*	.067
Results: sanctions	-.258**	-.063	.156	-.155
Quality of service: aimed at job	-.299**	-.155	.241**	-.053
Quality of service: aimed at voluntary work	-.182*	-.120	-.111	.210**

Notes: * $p < .05$; ** $p < .01$.

work). This provides some support for hypothesis 2a, which stated that reduced caseload size is not negatively related to workers' results; and support for hypothesis 2b, stating that reduced caseload size is not negatively related to service quality.

Performance targets concerning the number of clients who have to find a job are significantly related to three performance variables. As expected, these targets are positively related to the proportion of clients who found a job ($r = .263, p < .01$) and to the quality of services aimed at supporting clients to find a job ($r = .241, p < .01$). These targets are negatively related to the proportion of clients who found voluntary work ($r = -.193, p < .05$). Performance targets concerning the number of clients who have to find a job are non-significantly related to the proportion of clients who have been sanctioned and the quality of services aimed at supporting clients to find voluntary work.

Performance targets concerning the number of clients who have to find voluntary work are significantly related to two performance variables: the proportion of clients who found a job ($r = -.200, p < .05$) and the quality of services aimed at supporting clients to find voluntary work ($r = .210, p < .01$). The other outcome variables are non-significantly related to this type of performance targets.

Regression analyses: workers' results

Table 2 displays the results of the regression analyses including both total caseload and performance management as independent variables and the three types of results as dependent variables.

The table shows that, when we control for client group and work experience, neither total caseload size nor performance management is related to the proportion of clients who found a job; 28.2 per cent of the variance is explained in this analysis.

Table 2

Regression analyses performance outcomes as dependent variables

	Results: paid job	Results: voluntary work	Results: sanctions
Total caseload	-.148	-.251**	-.188
Performance management: no. of jobs	.105	-.231**	.131
Performance management: no. of clients starting voluntary work	-.095	.076	-.083
Client group (1 = large distance)	-.123	-.004	-.144
Client group (1 = very large distance)	-.389**	-.024	-.250*
Work experience	.134	-.054	-.127
R²	.282	.102	.170

Notes: standardized regression coefficients are reported. * $p < .05$; ** $p < .01$.

In case of the proportion of clients who found voluntary work, both caseload size and performance management concerning the number of clients who need to find a job are negatively and significantly related ($\beta = -.251$ and $\beta = -.231$, $p < .01$ respectively); 10.2 per cent of the variance is explained in this analysis.

Lastly, in case of the proportion of clients who were sanctioned, only one of the control variables was significant. This means that, when controlled for client group and work experience, neither caseload nor performance management has a significant relationship with this result variable.

These results provide moderate support for hypotheses 1a and 4a. In one of the three analyses, caseload had a negative effect on results (H1a). In one analysis, we found support for the hypothesis that performance management on a specific result is negatively related to workers' performance on other, not-measured results (H4a): performance management on the number of clients who need to find a job is negatively related to the number of clients who start voluntary work. At the same time, we have to reject hypothesis 3a: we found no evidence that performance management on a specific result is positively related to workers' performance on that result. We also found no evidence supporting hypothesis 5a. Performance management concerning the number of clients who need to find a job has a weaker negative relationship with performance on one of the other, not-measured results (proportion of clients starting voluntary work) than total caseload size. In the case of performance management concerning the number of clients who need to start voluntary work, neither performance management nor total caseload size were significantly related to one of the other, not-measured results (proportion of clients entering paid jobs).

Regression analyses: workers' service quality

In table 3 the results of the regression analyses including the two service quality indicators as dependent variables are displayed.

The table shows that, when we control for client group and work experience, total caseload as well as performance targets concerning the number of clients who need to find a job are significantly related to the quality of services aimed at supporting clients to find a job ($\beta = -.187$, $p < .01$; $\beta = .153$, $p < .05$ respectively). 22.6 per cent of the variance is explained in this analysis. A quite similar picture emerges when we look at the quality of services aimed at supporting clients to find voluntary work. Both caseload size and performance targets concerning the number of clients who have to find voluntary work are significantly related ($\beta = -.225$, $p < .01$; $\beta = .207$, $p < .05$ respectively); 11.9 per cent of the variance is explained in this analysis.

The results provide full support for hypothesis 1b: total caseload is negatively related to service quality. We also found full support for hypothesis 3b, which states that performance management on a specific target is positively related to workers' quality of service directed at that target. We have to reject hypothesis 4b as we found no evidence that performance management on a specific target is negatively related to workers' quality of service directed at other targets. We also found no support for hypothesis 5b: performance

Table 3

Regression analyses service quality as dependent variable

	Quality of services aimed at job finding	Quality of services aimed at starting voluntary work
Total caseload	-.187*	-.225**
Performance management: nr of jobs	.153*	-.131
Performance management: nr of clients starting voluntary work	.030	.207*
Client group (1 = large distance)	-.164*	-.024
Client group (1 = very large distance)	-.338**	.086
Work experience	.026	.064
R²	.226	.119

Notes: standardized regression coefficients are reported. * $p < .05$; ** $p < .01$.

management on a specific result is not stronger negatively related than total caseload to the quality of services directed at other, not-measured results.

Discussion and Conclusion

This study set out to answer the question of how the size of caseloads and performance management relate to the quality and outcomes of social services. The findings of our analyses concerning total caseload size show that it is negatively related to three out of five performance variables: the number of clients who start voluntary work and the two service quality variables. If we assume that these relationships can be interpreted as causal relationships and that performance can be conceived as an effect rather than a cause of caseload size (which seems plausible from a logical point of view), our findings show that high caseloads are reason for concern where performance is concerned. Our study also revealed that focusing service provision on a proportion of workers' total caseload is somewhat effective. Interestingly, though, this effectiveness is most clear when we look at service quality: when looking at reduced instead of total caseload, we no longer find a negative and significant relation between caseload size and service quality. In terms of results (job entries and starting voluntary work), the negative relation between caseload size and performance becomes smaller but remains significant. Thus, focusing attention on part of their caseload helps workers to improve the quality of services they provide but the impact of this strategy on the results they realize seems to be less strong.

As far as our findings concerning performance management are concerned, we found that performance management is not related to the results workers need to realize. In other words, performance management does not improve workers' results on their targets. These findings could bring us to the

conclusion that performance management has little added value as a goal-setting device, for example because workers know what is expected from them anyhow or because performance management regimes in Dutch local welfare agencies are relatively mild in terms of the quantitative targets that are set or in terms of the consequences of not meeting targets. However, the picture changes when we look at the impact of performance management on the quality of services: performance management on a target is positively related to the quality of services directed at realizing that target. Thus, performance management seems to affect workers' efforts to realize their targets even if it does little to improve their actual results. Performance management does, however, affect results in another way: performance management on job placements directed attention away from voluntary work placements. The same does not happen when workers have performance targets on voluntary work placements. A possible explanation may be that workers know that job placements are more valued than voluntary work placements by policymakers and by their agencies, irrespective of the nature of their performance targets.

Lastly, in the context of our research, total caseload size seems to be a more important determinant in affecting performance negatively than performance management. In other words, our study provides no indications that in terms of the potential impact on performance, performance management deserves more attention than caseload size. Therefore, on the basis of our study, we recommend researchers interested in the quality and results of social services to focus attention to the impact of both caseloads and performance management.

Our study has some weaknesses that should make us reserved in drawing firm conclusions, and that result in suggestions for future research. First, for reasons elaborated above we looked at caseload which is at best an acceptable proxy of workers' actual workload. As our research shows, distinguishing between total and reduced caseload is useful to come to more nuanced conclusions regarding the impact of caseload size. However, further refinements to bridge the gap between measuring caseloads and workloads, especially in quantitative studies, are highly desirable. Second, our operationalization of service quality has its limitations, as we only looked at service diversity and at two types of services: services aimed at finding a paid job and services aimed at finding voluntary work. Especially in the context of services for the most difficult to employ groups of unemployed, including a broader set of social services in studies of the impact of caseload size and performance management is recommendable. Third, our operationalization of performance management was dichotomous (workers made or did not make performance agreements with their supervisors) without including more qualitative aspects such as the size and feasibility of performance targets, the enforcement of performance management agreements, consequences of realizing or failing to realize agreed targets, etc. Given the contested nature of performance management, future research should preferably avoid statements about performance management without including these qualitative aspects. Fourth, our data does not provide insight into the consequences of reduced caseloads in terms of what groups of unemployed people are being

served and what groups excluded from services. Even though our study showed that caseload reductions may improve service quality, the consequences for people excluded from services should be included in the evaluation of effects of caseload reductions as well. Lastly, the sample used in this study is rather small ($n = 163$). This could have resulted in Type II errors (i.e. false negatives). We suggest that larger samples are collected in follow-up studies to overcome this drawback.

In the introduction, we raised the issue of the political implications of decisions concerning working conditions in agencies delivering social policies. Some indications of these implications were found in our study. In Dutch local welfare agencies, working conditions seem to affect clients with a very large labour-market distance specifically. Workers who work for this group have relatively high total caseloads and, as we see, this has a negative impact on services and voluntary work placements. We may assume that this target group needs more complex and multifarious types of support, but as our findings show, workers' services aimed at finding paid work are more diversified than those aimed at finding voluntary work. Workers respond to high caseloads by focusing attention on part of their caseload, but this excludes considerable proportions of their total caseload from any welfare-to-work support whatsoever. So even though Dutch formal social assistance policies do not prioritize some client groups over others, this is exactly what happens in practice. The funding system for social assistance expenses (see above) stimulates municipal and organizational decisions that direct resources to clients considered close to the labour market and lead to high caseloads for workers working for the most vulnerable client groups, resulting in selection processes, reductions of service quality and deteriorating results.

In addition, we saw that performance management on the number of clients who need to find a job is negatively related to the number of clients who start voluntary work. This could be an indication that workers work more effectively, since job placements are preferred to voluntary work placements. But it could also indicate that workers – for example workers working with mixed target groups – focus attention on those clients who are closest to the labour market, whereas clients who could benefit from participating in voluntary work (as a form of social participation or a stepping-stone towards paid work) are parked. Thus, performance management may have similar negative effects as caseload size on services provided for people with a very large labour-market distance. And although performance agreements regarding the number of people who should start voluntary work are positively related to the quality of services directed at this target, this type of agreement is not common among workers working with mixed target groups or those working for clients closer to the labour market.

The political implications of working conditions and how they affect frontline work often remain unnoticed in public and political debates. At the same time, our study contributes to the increasing number of studies pointing at these political implications. Therefore, social scientists' attempts to investigate these working conditions and their impact on the frontline provision of social services deserve continuation.

Acknowledgements

The study reported in this article was funded by the Dutch foundation, Stichting Kennis Ontwikkeling HBO, reference number PRO-2-012. The authors would like to thank three anonymous reviewers for their constructive and useful comments.

Notes

* Corresponding author.

1. The survey took place at the end of 2012.
2. A group of workers made performance agreements concerning the number of clients who should be in welfare-to-work activities. As this is an output rather than an outcome target, we did not include this performance target in our analyses.

References

- Austin, M., Johnson, M., Chow, J., De Marco, A. and Ketch, V. (2009), Delivering welfare-to-work services in county social service organizations: an exploratory study of staff perspectives, *Administration in Social Work*, 33, 1: 105–26.
- Bloom, H., Hill, C. and Riccio J. (2003), Linking program implementation and effectiveness: lessons from a pooled sample of welfare-to-work experiments, *Journal of Policy Analysis and Management*, 22, 4: 551–75.
- Boselie, P. (2014), *Strategic Human Resource Management: A Balanced Approach*, 2nd edn, London: McGraw-Hill Higher Education.
- Brodkin, E. (2006), Bureaucracy redux: management reformism and the welfare state, *Journal of Public Administration Research and Theory*, 17, 1: 1–17.
- Brodkin, E. (2008), Accountability in street-level organizations, *International Journal of Public Administration*, 31, 3: 317–36.
- Brodkin, E. (2011), Policy work: street-level organizations under new managerialism, *Journal of Public Administration Theory and Practice*, 21, suppl2: i253–77.
- Brodkin, E. (2013), Street-level organizations and the welfare state. In E. Brodtkin and G. Marston (eds), *Work and the Welfare State. Street-Level Organizations and Workfare Politics*, Washington, DC: Georgetown University Press, pp. 17–35.
- Brodtkin, E. and Marston, G. (eds) (2013), *Work and the Welfare State. Street-Level Organizations and Workfare Politics*, Washington, DC: Georgetown University Press.
- Clarke, J. and Newman, J. (1997), *The Managerial State*, London: Sage.
- Considine, M. (2001), *Enterprising States. The Public Management of Welfare-to-Work*, Cambridge: Cambridge University Press.
- Considine, M., Lewis, J. and O'Sullivan, S. (2011), Quasi-markets and service delivery flexibility following a decade of employment assistance reform in Australia, *Journal of Social Policy*, 40, 4: 811–33.
- Diefenbach, T. (2009), New Public Management in public sector organizations: the dark sides of managerialistic 'enlightenment', *Public Administration*, 87, 4: 892–909.
- Fletcher, D. (2011), Welfare reform, Jobcentre Plus and the street-level bureaucracy: towards inconsistent and discriminatory welfare for severely disadvantaged groups? *Social Policy and Society*, 10, 4: 445–58.
- Foster, D. and Hoggett, P. (1999), Change in the benefits agency: empowering the exhausted worker? *Work, Employment and Society*, 13, 1: 19–39.
- Fraser, K. and Estabrooks, C. (2008), What factors influence case managers' resource allocation decisions? A systematic review of the literature, *Medical Decision Making*, 28: 394–410.

- Godfrey, E. B. and Yoshikawa H. (2012), Caseworker-recipient interaction: welfare office differences, economic trajectories, and child outcomes, *Child Development*, 83, 1: 382–98.
- Heinrich, C. and Marschke, G. (2010), Incentives and their dynamics in public sector performance management systems, *Journal of Policy Analysis and Management*, 29, 1: 183–208.
- Jewell, C. (2007), *Agents of the Welfare State. How Caseworkers Respond to Need in the United States, Germany, and Sweden*, Basingstoke: Palgrave MacMillan.
- Jewell, C. and Glaser, B. (2006), Toward a general analytic framework: organizational settings, policy goals, and street-level behavior, *Administration & Society*, 38, 3: 335–64.
- Johnson Dias, J. and Maynard-Moody, S. (2006), For-profit welfare: contracts, conflicts, and the performance paradox, *Journal of Public Administration Research and Theory*, 17, 2: 189–211.
- King, R., Le Bas, J. and Spooner, D. (2000), The impact of caseload on the personal efficacy of mental health case managers, *Psychiatric Services*, 51, 3: 364–9.
- Larsen, F. (2013), Active labor-market reforms in Denmark. The role of governance in policy change. In E. Brodtkin and G. Marston (eds), *Work and the Welfare State. Street-Level Organizations and Workfare Politics*, Washington, DC: Georgetown University Press, pp. 103–24.
- Lee, J. (2009), Another dimension of welfare reform: the implementation of the Employment Insurance Programme in Korea, *International Journal of Social Welfare*, 18, 3: 281–90.
- Lipsky, M. (2010), *Street-Level Bureaucracy. Dilemmas of the Individual in Public Services*, New York, NY: Russell Sage Foundation.
- Locke, E. and Latham, G. (2002), Building a practically useful theory of goal setting and task motivation: a 35-year odyssey, *American Psychologist*, 57, 9: 705.
- McDonald, C. and Marston, G. (2006), Room to move? Professional discretion at the frontline of welfare-to-work, *Australian Journal of Social Issues*, 41, 2: 171–82.
- McDonald, C. and Marston, G. (2008) Motivating the unemployed? Attitudes at the front line, *Australian Social Work*, 61, 4: 315–26.
- Møller, M. and Stone, D. (2013), Disciplining disability under Danish active labour market policy, *Social Policy & Administration*, 47, 5: 586–604.
- Morgen, S. (2001), The agency of welfare workers: negotiating devolution, privatization, and the meaning of self-sufficiency, *American Anthropologist*, 103, 3: 747–61.
- Noordegraaf, M. and Steijn, B. (2013), *Professionals under Pressure. The Reconfiguration of Professional Work in Changing Public Services*, Amsterdam: Amsterdam University Press.
- Pollitt, C. and Bouckaert, G. (2000), *Public Management Reform. A Comparative Analysis*, Oxford: Oxford University Press.
- Rhodes, M., Biondi, L., Gomes, R., Melo, A., Ohemeng, F., Perez-Lopez, G., Rossi, A. and Sutiyono, W. (2012), Current state of public sector performance management in seven selected countries, *International Journal of Productivity and Performance Management*, 61, 3: 235–71.
- Riccio, J., Friedlander, D. and Freedman, S. (1994), *GAIN: Benefits, Costs, and Three-Year Impacts of a Welfare-to-Work Program*, New York, NY: MDRC.
- Riccucci, N., Meyers, M., Lurie, I. and Han, J. (2004), The implementation of welfare reform policy: the role of public managers in front-line practices, *Public Administration Review*, 64, 4: 438–49.
- Ridzi, F. (2004), Making TANF work: organizational restructuring, staff buy-in, and performance monitoring in local implementation, *Journal of Sociology and Social Welfare*, XXXI, 2: 27–49.
- Rosenthal, P. and Peccei, R. (2006), The social construction of clients by service agents in reformed welfare administration, *Human Relations*, 59, 12: 1633–58.

- Ryu, S., Wenger, J. and Wilkins, V. (2012), When claimant characteristics and prior performance predict bureaucratic error, *The American Review of Public Administration*, 42, 6: 695–714.
- Schram, S. (2012), Welfare professionals and street-level bureaucrats. In M. Grey, J. Midgley and S. Webb (eds), *The SAGE Handbook of Social Work*, London: Sage, pp. 67–81.
- Soss, J., Fording, R. and Schram, S. (2011), The organization of discipline: from performance management to perversity and punishment, *Journal of Public Administration Research and Theory*, 21, suppl2: i203–32.
- Stevens, M. (2008), Workload management in social work services: what, why and how? *Practice: Social Work in Action*, 20, 4: 207–21.
- Talbot, C. (2010), *Theories of Performance. Organizational and Service Improvement in the Public Domain*, Oxford: Oxford University Press.
- van Berkel, R. (2014), Quasi-markets and the delivery of activation – a frontline perspective, *Social Policy & Administration*, 48, 2: 188–203.
- van Berkel, R., De Graaf, W. and Sirovatka, T. (2011), *The Governance of Active Welfare States in Europe*, Basingstoke: Palgrave.
- Weibel, A., Rost, K. and Osterloh, M. (2010), Pay for performance in the public sector – benefits and (hidden) costs, *Journal of Public Administration Research and Theory*, 20, 2: 387–412.