Motivating and Demotivating Effects of Performance-Related Pay in Swedish Public Sector Organizations

Bengt Larsson¹,², Ylva Ulfsdotter Eriksson¹,², and Petra Adolfsson¹

Abstract
This study contributes to the evidence on motivational effects from performance-related pay (PRP) in the public sector. The theoretical point of departure is that the practical organization and administration locally affect the motivational effects of PRP. The analysis is based on surveys administered to employees (including managers) in Swedish public sector organizations at municipal, regional, and state levels. One of the main conclusions is that PRP is not motivating or demotivating per se, but can be both motivating and demotivating within the same organization. The (de/)motivational effect depends on the local level organization and practices of PRP, particularly the quality of the performance appraisal dialogue. While confirming the importance of justice perceptions, it also shows the effects of managers’ and employees’ preparations, knowledge of criteria, the quality of the performance appraisal dialogue, and the manager’s evaluation style, while controlling for justice perceptions and background variables.

Keywords
rewards, pay, motivation, performance appraisal, public sector

¹University of Gothenburg, Sweden
²Linnaeus University, Växjö, Sweden

Corresponding Author:
Bengt Larsson, Department of Sociology and Work Science, University of Gothenburg, Box 720, Göteborg 405 30, Sweden.
Email: bengt.larsson@socav.gu.se
Introduction

One of the main rationales for employers to use individualized performance-related pay (PRP) is to increase employee motivation. The assumption that linking employees’ rewards to their performance leads to “an endeavour to increase individual performance” is also common in economics and behavioral approaches (Papenfuß & Keppeler, 2020, p. 1123; cf. Weibel et al., 2009). However, the use of PRP has been questioned, and research in public sector organizations shows inconclusive results. The effects of PRP on motivation, commitment, or quality of services are at best occasional, generally rather marginal, and sometimes even negative (Bryson et al., 2017; Chatelain-Ponroy et al., 2018; Chen, 2018; Choi & Whitford, 2017; Lee, 2019; Perry et al., 2009; Schay & Fisher, 2013; Wenzel et al., 2019; Weske & Schott, 2018). Despite this, the attractiveness of PRP systems in the wake of the New Public Management trend show no tendency to abate (cf. Bellé, 2015; Fuller & Cooke, 2018; Oh & Lewis, 2009; Williams et al., 2019).

At least three strands of research problematize PRP in the public sector (Bellé, 2015). One argues that the reward-oriented self-interest upon which such systems are based conflicts with and may crowd out “other-regarding” or even altruistic values among public sector employees—so called public service motivation (PSM) (Papenfuß & Keppeler, 2020; cf. van Loon, 2017; Weske & Schott, 2018). A second strand points to institutional characteristics of public organizations, such as budget constraints and a lack of pay secrecy (cf. Bellé, 2015). Important as these approaches are, the present study instead relates to a third strand of research focusing on inadequacies in design and management practices of PRP (cf. Favero et al., 2016; Lee, 2019, Weske & Schott, 2018; van Loon, 2017).

The aim is to study which factors in the local level organization and practice of PRP increase or decrease motivation among public sector employees. More specifically, the study aims to analyze the motivational effects on employees from performance appraisal in three organizational contexts; that is, to what degrees the performance appraisal motivates or demotivates employees to do a better job or strive to increase their wage. The local level organization and practice of PRP is of particular relevance to study, since this is something that the organizations may actually influence, in contrast to the effects of intrinsic motivation values or budget constraints in general in the public sector. In addition, even though the practices and administration of PRP performed by managers seem to play an important role for motivation in some of the previous research, it is seldom at the center of analysis.

The study is based on surveys in organizations at the municipal, regional, and state levels in Sweden. The dependent variables were indexes over (self-reported) extrinsic motivational and demotivational effects from PRP practices. The explanatory variables were employee preparations, type of performance criteria used, the performance appraisal dialogue quality, and the manager’s evaluation style. In addition, the regressions controlled for feelings of being treated unfairly, as well as organization, position, education, gender, age, and tenure.
This study contributes to research on pay and motivation by confirming some previous results while also presenting novel insights into the importance of local-level PRP practices. On a general level, the study shows that PRP can be both motivating and demotivating within the same organization, depending on local practices and employee variation. More specifically, it shows the importance of justice perceptions, as well as of managers’ and employees’ preparations, knowledge of criteria, the quality of the performance appraisal dialogue, and the manager’s evaluation style. For practitioners, this means that it is important for local-level managers and employees to have good knowledge of criteria, and prepare before the performance appraisal; and also that performance appraisals include retrospective, evaluative, and forward-looking aspects and cover all performance criteria. A lack of knowledge and/or preparations from managers and employees, and experiences of injustice or bad quality in local-level appraisal dialogue will cause employee motivation to decrease.

The article begins with a presentation of theory, previous research and hypotheses, followed by a methods section that also includes information about PRP in the Swedish public sector. Thereafter, the findings are presented and some general conclusions are drawn.

**Theoretical Framework and Previous Research**

PRP systems are designed to be motivating by communicating goals to employees and by distributing praise and pay after evaluating performances (Wang & Teo, 2008). According to Self Determination Theory (SDT), motivation exists in a continuum from intrinsic to extrinsic (Gerhart & Fang, 2015; Ryan & Deci, 2000; Weske & Schott, 2018). As public sector workers are thought to put more value into intrinsic rewards, such as “a sense of achievement” (Bryson et al., 2017, p. 584), or PSM, it has been suggested that monetary rewards have less motivational effect for those workers (Papenfuß & Keppeler, 2020; van Loon, 2017; Weske & Schott, 2019). PRP might even be counterproductive, since its principles may conflict with the institutional values in the public sector and crowd out intrinsic motivation (Perry et al., 2009; Weibel et al., 2009).

However, according to SDT, motivation can come from external incentives such as a (substantial) pay raise, although the effects are conditioned by certain factors, including: that the performance goals and expectations are clear and reasonable; that employee autonomy is supported; that employees trust managers; that employees are receptive to performance appraisal, and have an actual possibility to improve; that the appraisal is accurate and the feedback adequate; and that rewards are distributed fairly (Choi & Whitford, 2017; Favero et al., 2016; Gerhart & Fang, 2015; Glassman et al., 2010; Lee, 2019; Schay & Fisher, 2013; Oh & Lewis, 2009). Under such circumstances, PRP may strengthen employees’ perceptions that their efforts are appreciated, thereby increasing their job satisfaction, organizational commitment, and work intensity (cf. Choi & Whitford, 2017; Ogbonnaya et al., 2017).

It has been said that PRP is “capable of producing spectacularly good results” and increase employee commitment and motivation (Gerhart & Fang, 2015, p. 490;
Glassman et al., 2010; Ogbonnaya et al., 2017). On the other hand, PRP may also produce bad results, as many things can go wrong in the design and implementation (Gerhart & Fang, 2015, p. 490f.; Ogbonnaya et al., 2017; Wenzel et al., 2019). The focus on performance measures can be both too strong and too weak, criteria may be unclear, and it may be difficult to differentiate between performances. PRP may lead to employee competition and may be seen as “exploitative” by increasing work demands, or if perceived as being a control mechanism. The motivational effects of PRP may be diminished if employees perceive that the performance appraisal is unjust in its principles, practices, and outcomes, or if it is not communicated in a participatory way (Kim & Holzer, 2016; cf. Colquitt, 2001). In addition, employees may perceive rewards and the distribution of them in different ways and they may have diverging motivational profiles (De Gieter & Hofmans, 2015; Weske & Schott, 2018).

Positive effects from PRP have been found in the private sector, while the effects in the public sector are more uncertain. According to Wenzel et al. (2019, p. 233), existing research in the public sector does not indicate increased motivation or quality of services in general, and that performance appraisal unaccompanied by monetary rewards risks crowding out (intrinsic) motivation (Kim & Holzer, 2016, p. 32). Recent studies report varying results from different countries, probably both because of contextual factors and differences in methodological approaches. In the US, Schay and Fisher (2013) found that PRP strengthened the link between pay and performance among federal government employees, whereas Chen (2018, p. 135) found a decrease in intrinsic motivation in the public and non-profit sectors. Lee (2019) showed that the quality of the implementation process moderated the effects on PSM among public sector employees. The motivational effect was greater for those who found the process to be fair and received constructive feedback. In Britain, Bryson et al. (2017) found that PRP had positive effects on commitment in the private sector, while the effects were negative in public sector workplaces. In Germany, Wenzel et al. (2019) concluded that “intrinsic motivation was negatively affected by performance pay if [employees] perceived it as controlling.” In France, Chatelain-Ponroy et al. (2018) found a negative correlation between values of “publicness” and commitment to performance-related management among university employees. In the Netherlands, Weske and Schott (2018) found different motivational profiles among municipal employees, with extrinsic rewards affecting motivation to varying degrees. In Norway, a study of performance appraisal in hospitals found that dialogue around individual goal setting and constructive feedback is a decisive factor for motivation to increase (Vasset et al., 2011). In Sweden, research on PRP-related issues has focused on attitudes toward PRP pay and causes of justice perceptions and pay satisfaction rather than on motivation. However, justice perceptions have strong effects on pay satisfaction and organizational commitment, and feedback from line managers was found to be important (Malmrud et al., 2020; Stråberg, 2010).

To summarize, both theories and previous research on PRP point in somewhat diverging directions. Even though the effects of justice perceptions on motivation are quite robustly evidenced, it seems that the motivational effects depend on the design and practical organization and implementation of the system by local line managers.
Therefore, it is important to untangle such factors before trying to establish whether PRP has motivational effects or not in the public sector in general.

**Hypotheses**

Based on our focus on local-level practices and in relation to the theories and previous research results discussed above, the following hypotheses were developed:

**H1.** PRP may be both motivating and demotivating within public sector organizations, depending on local practices of administering PRP

**H2:** If the employee perceives being treated unfairly in the appraisal/wage determination, the motivational effect decreases and the demotivational effect increases.

Both of these hypotheses lay the groundwork for the following: H1 is general and undirected and functions as a departure for testing directed hypotheses on motivation and demotivation within the same organizations. H2, which is based on justice perception theory, has already been verified in previous research, so it is not a focal variable, but a control variable that is indirectly tested.

**H3:** The better known and more explicitly used the appraisal/wage criteria are, the more focused the appraisal will be on them, which increases the motivational effect and decreases the demotivational effect of PRP.

**H4:** The better quality of the feedback, the more motivational effect and the less demotivational effect it will have.

These two hypotheses are based on the above theories and results indicating the importance of clear goals and expectations, and constructive feedback on performance, allowing the employees to improve.

**H5.** The better prepared the employee is, the better the performance appraisal dialogue works, which increases the motivational effect and decreases the demotivational effect.

**H6:** The manager’s evaluation style affects the motivational effect on employees:

(a) If the manager takes a comprehensive approach to performance and listens to the employee’s self-evaluation, the motivational effect increases and the demotivational effect decreases;

(b) If the manager is biased or selective in the appraisal, the motivational effect decreases and the demotivational effect increases.

Both hypotheses concern what the participating actors bring to the appraisal dialogue, both in terms of knowledge and information about performances. They also relate to the theoretical assumptions that the evaluation and feedback need to be adequate and that the manager is perceived as worthy of the employees’ trust in the evaluation of performance.
Methods

Context: PRP in Swedish Public Sector Organizations

The Swedish system of pay determination has changed radically in recent decades. The Swedish post-war model of centralized national wage bargaining was abandoned in the 1980s and a system of coordinated sectoral bargaining with local adjustments was introduced in the 1990s (Baccaro & Howell, 2017). The discourse on pay turned from solidary values to organizational productivity, flexibility, and individually differentiated wages based on performance. Thus, the solidary principle of “equal pay for equal work,” based on nationally coordinated wage tariffs, in which wages were set based on job evaluations, qualifications, tenure, and general wage raise, was gradually supplanted by individualized principles rewarding contribution and effort.

Individualized PRP has gone furthest in the public sector (Kjellberg, 2019). Today, 60% of public sector employees are covered by agreements that set no figures on local wage raise as long as the local partners agree. For the other 40%, the central agreement specifies a wage pot, but with no individual guarantees regarding the wage raise. Consequently, practically all public employees are covered by agreements giving space for individually differentiated PRP in the yearly salary review process. The figures in the private sector are lower: approximately one-third of all employees and two-thirds of private white-collar employees are covered by agreements that leave room for PRP (Ulfsdotter Eriksson et al., 2020a).

The overarching PRP principles are defined by the collective agreements, and there is 100% coverage of collective agreements in the Swedish public sector. The agreements generally state that wages should be individual and differentiated, based on responsibilities, skills, task complexity, and performance and contribution. The former categories mainly differentiate between occupational groups, whereas performance and contribution are used to differentiate pay between employees with similar jobs or positions.

Collective agreements and pay policies both stress the importance of employees perceiving the wage process as just, that they understand why they receive a certain wage and how they may improve it. Even if the local performance evaluation process and appraisal criteria differ, they have similar basic principles. Employees are evaluated on performance, behaviors, traits, and individual goals. The criteria formulated in the wage policy of the three organizations studied (Table 1) are fairly representative of Swedish public sector organizations overall. These general criteria are often exemplified or detailed for different positions and occupational groups.

The motivating aspect of PRP is strongly emphasized in the Swedish context, as it is in other countries. Such ideas are both implicated in the sectoral collective agreements in the public sector and highlighted by the employer associations (Ulfsdotter Eriksson et al., 2020a, 2020b).

Procedure and Sample

This study is based on web surveys distributed in 2016–2017 to employees (including managers) in three large organizations that represent the three levels of the public
Table 1. Official Performance/Wage Criteria in the Three Studied Organizations.

| Organization    | Main criteria and specifications                                                                 | Grading                                      |
|-----------------|---------------------------------------------------------------------------------------------------|----------------------------------------------|
| Municipality    | • Professionalism: Knowledge and quality of work; Goal- and result-oriented                        | ✓ Exceeds the requirements                   |
|                 | • Learning and Development: Develops works methods; develops competence vis-à-vis business needs  | ✓ Meets the requirements                     |
|                 | • Engagement and Behavior: An approach that creates security, enthusiasm, openness and good behavior| ✓ Need to develop                            |
| County hospital | • Professionalism: Improved results and good quality                                              | ✓ Exemplary                                  |
|                 | • Development: The business being brought forward through competence and innovation               | ✓ Good                                       |
|                 | • Collaboration: For respectful relationships                                                     | ✓ Need to be improved                        |
| Government agency | • Professionalism, result-oriented and contributes to business goals                               | ✓ Excellent                                  |
|                 | • Understands the role and collaborates with others for the benefit of the whole                  | ✓ Good                                       |
|                 | • Is innovative and ready to change                                                                | ✓ Acceptable                                 |
|                 | • Shares knowledge and experiences                                                                  | ✓ Insufficient                               |
|                 | • Talks and writes objectively and is easy to understand                                            |                                             |
|                 | • Takes responsibility for one’s development                                                       |                                             |

sector in Sweden: municipal, regional, and state. The dataset consists of 4,313 responses, covering a broad variety of operational areas and occupations, and background variables such as age, sex, education, occupation, and tenure.

The municipal organization is a mid-size Swedish municipality (75,000–150,000 inhabitants, employing 7,500–13,000 people) with responsibilities in public services such as childcare, primary and secondary education, but also water supply, rescue services, and waste disposal. The survey was distributed and 1,980 filled-in questionnaires were received. The exact population cannot be revealed since that would make it possible to identify the municipality, and we had to promise to keep this confidential to acquire access to the organization, but the response rate was quite low (15%–25%). The response analysis showed adequate representativeness for general categories of employees and operating areas, with slightly lower rates for low-educated and foreign-born staff.

The regional organization is a large public hospital, with all the functions and occupations regularly operated by such. The survey was distributed to 4,999 randomly selected employees, and 1,838 responses were received (37% response rate). The response analysis showed good representativeness for general occupational categories such as physicians, nurses, and assistant nurses.
The state agency has a quite a broad set of operational activities, including case handling and contacts with citizens. A large proportion of its staff is highly educated in law and social sciences. The survey was sent to all 1,127 employees in one main unit, covering a wide range of operational areas, and 495 responses were received (44% response rate). The response analysis showed good representativeness in terms of gender, age, and proportion of managers and staff, with slightly higher rates for highly educated and experienced staff.

**Analytical Approach and Measures**

Since the survey was sent to three organizations, the dataset might violate the assumption of OLS regression that data is independent and not “nested.” Therefore, we checked the need and benefits of using multilevel modeling. However, the grouping level clustering effect tested through the intraclass correlations coefficients (ICC) was very low (and insignificant): only 3.4% of the variance on the motivation index and 1% of the variance on the demotivation index were explained at the organizational level. Since we did not have any hypotheses including independent variables at that level, we chose not to conduct multilevel analyses but instead controlled for organization in OLS (cf. Garson, 2013, p. 27). We safeguarded by also performing separate OLS regressions on each organization, which did not reveal any major differences between the organizations (see Supplementary Material, Tables 5–10). The regression tables report the Beta coefficients, although we also present the β and t values in the text for the important variables. The effect sizes are discussed mainly at the model level (that is, adj. $R^2$). In addition, we considered the risk of common source bias (CSB), and performed a Harman’s One-Factor test, which did not indicate serious CSB problems as the first factor accounted for 23% of the variation (the conventional threshold for CSB being 50%). However, we still tried to be cautious in drawing causal conclusions and discussing correlations—for example, by comparing respondents from different levels in the organization and performing both joint and separate regressions for the three organizations (cf. George & Pandey, 2017).

The dependent variables of motivation and demotivation are operationalized as self-reported outcome-focused (extrinsic) motivation. As such, these are rather narrow constructs of motivation and demotivation, which do not take into consideration behavioral measures or process-related (means-focused and intrinsic) aspects of motivation (cf. Touré-Tillery & Fishbach, 2014). However, the interconnected organizational and individual goals—that the employees become motivated to improve their work to increase their pay—are often assumed in both policy texts guiding PRP, and in general economics and behavioral approaches (Papenfuß & Keppeler, 2020, p. 1123; cf. Weibel et al., 2009).

After performing a PCA (not shown), we created two separate indexes for motivation and demotivation (see Table 2). The *Motivation* index included two items: “The dialogue made me motivated to do an even better job in the future,” and “The dialogue made me motivated to work to increase my wage.” For both statements, the respondents could choose from the following options: “Do not know/Have no opinion” (re-coded as
Table 2. Descriptive Statistics for Variables.

|                      | n (n yes) | Mean | SD   | Min-Max | A   |
|----------------------|-----------|------|------|---------|-----|
| **1. Dependent variables** |           |      |      |         |     |
| Motivation Index     | 3,651     | 4.20 | 1.91 | 2–8     | 0.89|
| Demotivation Index   | 3,668     | 4.56 | 2.11 | 2–8     | 0.81|
| **2. Independent variables** |       |      |      |         |     |
| **(A) Focus variables** |           |      |      |         |     |
| Employee’s preparations before talk | | | | | |
| No preparations (ref) | (553)     | 0.13 | 0.33 | 0/1     |     |
| Checked my notes from last year | (1,224)   | 0.28 | 0.45 | 0/1     |     |
| Read through the PA/wage criteria | (2,824)   | 0.65 | 0.48 | 0/1     |     |
| Wrote arguments and/or examples | (2,248)   | 0.52 | 0.50 | 0/1     |     |
| Kept notes throughout the year | (426)     | 0.10 | 0.30 | 0/1     |     |
| Prep. (one or more of the above) | (3,563)   | 0.83 | 0.38 | 0/1     |     |
| Try to find out what others earn | (952)     | 0.22 | 0.42 | 0/1     |     |
| **Type of performance criteria used** | | | | | |
| No/unknown criteria (ref) | (1,023)   | 0.24 | 0.42 | 0/1     |     |
| Previously existing criteria | (2,144)   | 0.50 | 0.50 | 0/1     |     |
| Elaborated/revised criteria | (1,146)   | 0.27 | 0.44 | 0/1     |     |
| Perf. appraisal dialogue quality | 3,667     | 11.93| 3.75 | 5–20    | 0.83|
| **Manager’s evaluation style** | | | | | |
| Comprehensive and listening (Index) | 4,161     | 15.23| 3.87 | 6–24    | 0.83|
| Selective and biased (Index) | 4,168     | 9.52 | 2.46 | 4–16    | 0.67|
| **(B) Control variables** | | | | | |
| Feelings of being treated unfairly | (1,973)   | 0.54 | 0.50 | 0/1     |     |
| **Organization** | | | | | |
| Government authority (ref) | (495)     | 0.12 | 0.32 | 0/1     |     |
| Regional hospital | (1,838)   | 0.43 | 0.49 | 0/1     |     |
| Municipality | (1,980)   | 0.46 | 0.50 | 0/1     |     |
| **Position** | | | | | |
| Employee + Man. without W-S (ref) | (3,978)   | 0.93 | 0.25 | 0/1     |     |
| Wage setting manager | (297)     | 0.07 | 0.25 | 0/1     |     |
| **Education** | | | | | |
| Tertiary (ref) | (3,206)   | 0.75 | 0.43 | 0/1     |     |
| Secondary | (932)     | 0.22 | 0.41 | 0/1     |     |
| Primary | (111)     | 0.03 | 0.16 | 0/1     |     |
| **Gender** | | | | | |
| Man (ref) | (955)     | 0.22 | 0.42 | 0/1     |     |
| Women/other | (3,310)   | 0.78 | 0.42 | 0/1     |     |
| **Age and experience** | | | | | |
| Age | 4,303     | 46.63| 11.14| 15–75   |     |
| Years in organization | 4,297     | 15.51| 12.11| 0–55    |     |
| Total | 4,313     |      |      |         |     |
“missing” and excluded from the analysis), “Do not agree at all” (1), “Agree to a low degree” (2), “Agree to a high degree” (3), and “Fully agree” (4). The Demotivation index was similarly constructed from two items: “The dialogue made me resigned concerning the possibility to develop in my work,” and “The dialogue made me resigned concerning my possibility to affect my wage.”

The independent variables used were categorized as focus variables and control variables (see Table 2), which are introduced in a set of subsequent models (Aneshensel, 2002). The first set of focus variables concerns Employee preparations before the performance appraisal dialogue, and these are categorical, related to the following questions: “I check my notes from last year”; “I keep notes throughout the year over my work efforts”; “I check the performance/wage criteria”; “I write down my arguments and/or examples of my work”; and “I try to find out what others earn.” As the first four variables had positive coefficients and the last had a negative coefficient, a new categorical variable (“Preparations”) was constructed, indicating whether the respondent had done one or more of the first four preparations. This was based on the assumption that doing preparations is qualitatively different from not doing any preparations. This variable was used in Models 2 to 5.

The variable Performance criteria used has three classes: (A) Those who did not know whether any criteria were used in the appraisal/wage determination, or thought that no such criteria existed/were used; (B) those for whom since long-established general or local criteria existed/were used; and (C) those for whom recently elaborated criteria existed/were used. In the municipality, only some respondents had received locally/occupationally specified criteria, while in the hospital, only some had new general criteria instead of previous local ones. The point of merging these in the same category is the assumption that a recent elaboration of criteria increases the focus on them from both managers and employees. As no such changes had been made in the state agency, there were no respondents from that organization in this class.

The third variable is an additive index indicating the quality of the performance appraisal dialogue, PA dialogue quality. The index was based on a PCA (not shown) and shows good scalability (see Table 2). It includes the variables: “The dialogue dealt with how well I performed my work”; “The dialogue dealt with my work in relation to the salary criteria”; “The dialogue dealt with what I can do to increase my wage”; “The dialogue dealt with how I can develop in my work”; and “The dialogue dealt with my opportunities to get tasks with more responsibilities.” For all statements, the respondents could choose from among the following responses: “Do not know/Have no opinion” (re-coded as “missing,” and excluded from the analysis), “Do not agree at all” (1), “Agree to a low degree” (2), “Agree to a high degree” (3), and “Fully agree” (4).

The fourth focus variable concerns the manager’s evaluation style. This is an additive index based on a PCA (not shown) indicating two underlying dimensions: (A) Comprehensive and listening consists of six items: “My manager knows enough about my work to be able to value it”; “My manager takes all performance/wage criteria into account when assessing me”; “I can influence my manager’s evaluation if I argue for my cause”; “I can easily highlight the value of what I do for my manager”; “It is easy for my manager to discern my performance from the team/group”;
Selective and biased consists of four items: “I do important things that are not valued in the salary criteria or by my manager”; “My manager listens more to those who set the tone in my department/unit”; “My manager assesses personal attributes more than work efforts”; “My manager mainly values the activities we document in systems.” For all statements, the respondents could choose from among the following responses: “Do not know/Have no opinion” (re-coded as “missing,” and excluded from the analysis), “Do not agree at all” (1), “Agree to a low degree” (2), “Agree to a high degree” (3), and “Fully agree” (4).

The control variables used was sorted into six groups. The variable Feelings of being treated unfairly is a categorical variable based on the item “Do you perceive yourself to be disadvantaged/treated unfairly in the setting of wages?” and is used as a proxy to control both for feelings of injustice and pay dissatisfaction. We also had controls for Organization (municipality, regional hospital, government agency), Position (employees and managers without wage-setting duties; wage-setting managers), Education (tertiary, secondary, primary), Gender (man, woman; or other), Age (years), and Tenure (years of employment).

Results

The preparatory PCA (not shown) indicated that motivation and demotivation are two separate underlying dimensions in the four items concerning the effects of the performance appraisal dialogue. That is, a person may be motivated or not to do a better job and receive a wage raise, but that is slightly different from the effects of becoming resigned concerning the possibility to develop in one’s work or affect one’s wage. Already from this, we have support for the first part of H1 (that PRP may be both motivating and demotivating), and we now turn to the second part of this general hypothesis with the analysis of what effects the local practices of administering PRP have. To some extent, the effects of different variables on motivation and demotivation mirror each other, as expected in the hypotheses. Table 3 presents the results of the regression analysis of the construct of motivation, and Table 4 on demotivation. In order to highlight the similarities and differences between the motivation and demotivation effects, we discuss each model in both tables jointly.

Models 1 and 2 presents two versions of the effects of the focus variables relating to Employee’s preparations, and type of Performance criteria used. As can be seen from the R² the total effects of these variables are quite low. Model 1 explains just 3% of the variance on the motivation index, and Model 2 just above 6%. For the demotivation index, Model 1 explains 3% and Model 2 about 4%. In Model 1, there are some very weak effects from preparations such as checking notes from last year and performance criteria, keeping notes over or writing arguments or examples of performances. The joint categorical variable preparations in Model 2 shows a weak positive correlation with motivation ($\beta 0.10$, $t 5.56$), but no significant correlation with demotivation. Of course, this could be caused by reverse causation. Employees who have low expectations of PRP or who do not care a lot about pay may choose not to prepare. However,
**Table 3. Motivation From Performance Appraisal Dialogue. OLS (Beta Coefficients).**

|                   | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|-------------------|---------|---------|---------|---------|---------|
| (A) Focus variables |         |         |         |         |         |
| Employee preparations |       |         |         |         |         |
| No preparations (ref) |       |         |         |         |         |
| Checked my notes from last year | 0.24** |         |         |         |         |
| Checked performance/wage criteria | 0.44*** |         |         |         |         |
| Wrote arguments and/or examples | 0.13 |         |         |         |         |
| Kept notes throughout the year | 0.05 |         |         |         |         |
| Preparations (any of the above) | 0.48*** | 0.09 | 0.12 | 0.14* |         |
| Try to find out what others earn | −0.54*** | −0.40*** | −0.20*** | −0.16** | −0.10 |
| Performance criteria used |       |         |         |         |         |
| No/unknown (ref) |       |         |         |         |         |
| Previously existing | 0.59*** | −0.01 | −0.07 | −0.07 |         |
| Elaborated/revised | 1.07*** | 0.11 | 0.06 | 0.03 |         |
| PA dialogue quality (index) |       |         |         |         |         |
| Comprehensive and listening (Index) | 0.35*** | 0.28*** | 0.27*** |         |         |
| Selective and biased (Index) | 0.10*** | 0.09*** |         |         |         |
| Manager’s evaluation style |       |         |         |         |         |
| Comprehensive and listening (Index) | 0.30*** | 0.28*** | 0.27*** |         |         |
| Selective and biased (Index) | −0.06*** | −0.04*** |         |         |         |
| (B) Control variables |       |         |         |         |         |
| Feelings of being treated unfairly |       | −0.43*** |         |         |         |
| Organization |       |         |         |         |         |
| Government authority (ref) |       |         |         |         |         |
| Regional hospital |       | 0.16* |         |         |         |
| Municipality |       | 0.19* |         |         |         |
| Position |       |         |         |         |         |
| Staff (ref) |       |         |         |         |         |
| Wage setting manager |       |         |         |         |         |
| Education |       |         |         |         |         |
| Tertiary (ref) |       |         |         |         |         |
| Secondary |       | 0.19*** |         |         |         |
| Primary |       | 0.31* |         |         |         |
| Gender |       |         |         |         |         |
| Man (ref) |       |         |         |         |         |
| Women/other |       |         |         |         |         |
| Age and experience |       |         |         |         |         |
| Age |       |         |         |         |         |
| Tenure |       |         |         |         |         |
| Intercept | 3.885 | 3.313 | −0.080 | −0.166 | 0.153 |
| $R^2$ Adjusted | 0.03 | 0.06 | 0.50 | 0.54 | 0.55 |
| N | 3,445 | 3,445 | 3,445 | 3,445 | 3,445 |

*p < .05. **p < .01. ***p < .001 (index range 2–8) (exclude pairwise).

Preparations that are oriented toward finding out what others earn have a contrary effect or marginally decreasing the motivation in Model 1 ($\beta −0.09, t −5.27$). The correlations with demotivation are opposite and much more marked ($\beta 0.78, t 9.01$). Trying to find out what others earn seems to reduce motivation and increases demotivation. With this exception, there is some weak support for H5.

As Model 2 also tests H3, part of their explanatory power concerns the use of criteria in the performance appraisal dialogue. There are weak positive correlations with
motivation (β 0.15, t 7.33) of usage of previously existing criteria and weak negative correlations with demotivation (β −0.11, t −5.10), compared to no usage of criteria. These effects are somewhat stronger if the criteria were recently elaborated: both for motivation (β 0.25, t 11.58) and demotivation (β −0.14, t −6.41). Thus, there seems to be some support for H3, although caution is again required for this conclusion.

Model 3 increases the effect size radically for the motivation index and explains close to 50% of the variance, whereas the increase on the demotivation index is only
to 13%. This indicates that the quality of the appraisal dialogue has a strong and significant positive correlation with motivation (β 0.69, t 54.6), and a negative correlation with demotivation (β −0.32, t −19.03). Thus, the more encompassing the dialogue—in terms of being both retrospective, evaluative and forward-looking—the more motivated the employees felt to improve in their work and increase their wage (and the less resigned to develop in their work to affect their wage). Thus, this model supports H4, but we still explain more of motivation than of demotivation. In addition, as the introduction of the PA quality index reduces the effects from preparations and criteria used, this could indicate support for the reasoning in H3 and H5: that the effect of employee preparations and explicitly usage of criteria is mediated by an improvement of the performance appraisal dialogue. However, given that the data is cross-sectional, caution is required concerning this interpretation, which is a bit speculative.

By entering the variables relating to the Manager’s evaluation style in Model 4, the explanatory effect increases on both indexes, only slightly on motivation (to around 54%), but quite drastically on demotivation (to 29%). If the manager was perceived as taking a comprehensive and listening approach in the appraisal dialogue, motivation increases (β 0.20, t 12.55), while demotivation decreases (β −0.15, t −7.48). If the manager was experienced as selective and biased in the appraisal, this decreased motivation slightly (β −0.07, t −5.47) and increased demotivation substantially (β 0.36, t 22.23). These results support H6. While the dialogue quality had an important effect on motivation, it seems that the manager’s evaluation style is of greater importance for feelings of demotivation. As the manager’s evaluation style is related to (though not identical to) both the usage of criteria and the quality of the PA dialogue, it is unsurprising that the entering of that variable decreases the latter two. However, the quality of the dialogue and the manager’s evaluation style have significant independent effects.

When the control variables are entered in Model 5, we only saw slight increases in the explanatory effect on both the motivation index (to 55%) and the demotivation index (to 33%). An important control variable is the proxy for perceived injustice; that is, Feelings of being treated unfairly. This has a negative correlation with motivation (β −0.11, t −8.542), while the effect of appraisal dialogue quality remains strong and significant (β 0.53, t 34.95), as did the effect of the manager having a comprehensive and listening style, albeit at a lesser level (β 0.18, t 11.11). The correlation of the injustice proxy with demotivation was positive, as expected (β 0.19, t 11.86), while the effect of the appraisal dialogue quality was still significant (β −0.12, t −6.44), as was the effect of the manager having a comprehensive and listening style (β −0.10, t −5.20), whereas the correlation with the manager having a selective and biased evaluation style was still positive (β 0.31, t 19.12). There were also some small effects of organization, education, age, and employee position. The latter was that wage-setting managers experienced slightly higher degrees of motivation (β 0.04, t 3.01) and slightly lower degrees of demotivation (β −0.04, t 3.07) than did regular employees and line managers without wage-setting tasks. Thereby, H2 seem to be supported, as did H4 and H6 in Model 5.
Conclusion

Previous research offers different reasons why PRP work less well in the public sector. Some studies have focused on the tension between extrinsic and intrinsic PSM motivation (Papenfuß & Keppeler, 2020; cf. van Loon, 2017; Weske & Schott, 2018), and others on the institutional characteristics such as budget constraints and a lack of pay secrecy (cf. Bellé, 2015). The present study departed from the suggestion that the design of PRP systems and their practical organization and administration affect their motivational effects (cf. Favero et al., 2016; Lee, 2019, van Loon, 2017; Weske & Schott, 2018). In line with the suggestion that more studies need to focus on the role of context, local workplaces, and employee variation (De Gieter & Hofmans, 2015; Gerhart & Fang, 2015, p. 514f.), we focused on local line-manager practices and employee experiences of interaction with managers.

A first conclusion is that PRP systems are not motivating or demotivating per se—at least not in the public sector and given the contextual factors of the Swedish system. On the contrary, as various theories have indicated, they can have both motivating and demotivating effects within the same organization, depending on how well local line managers’ practices work. In addition, the results confirm previous research showing that perceived injustice in the appraisal and wage-setting principles and practices decrease motivation and increase demotivation (cf. Colquitt, 2001; Kim & Holzer, 2016; Malmrud et al., 2020; Stråberg, 2010).

However, the main contribution of the present study is that it shows the importance of employee preparations, the usage of explicit criteria in performance appraisal/wage determination, the quality of the performance appraisal dialogue and feedback, and the evaluation style of the manager in the appraisal. To some extent, our findings are in line with theoretical reasoning and empirical indications scattered in various previous research (Choi & Whitford, 2017; Favero et al., 2016; Gerhart & Fang, 2015; Glassman et al., 2010; Lee, 2019; Malmrud et al., 2020; Oh & Lewis, 2009; Schay & Fisher, 2013). However, these variables have seldom been at the center of the analysis.

The results show that a major motivational factor is the quality of the performance appraisal dialogue/salary talk. To be motivating, the dialogue should evaluate the previous year’s performances in relation to the criteria and the employee’s goals, but also touch on the possibilities for the employee to develop, acquire more responsibilities, and improve their wage. In addition, we found that employee preparations and the criteria used in the appraisal seem to be important for forming a high-quality performance dialogue. Naturally, if explicit criteria exist and are used in the appraisal, this improves the dialogue compared to if the employee did not know about the criteria used or did not find them being used. More interesting is that if the criteria had been recently elaborated, this seemed to increase the motivation and decrease motivation further, by increasing the dialogue quality. This also seems reasonable given that such elaborations would raise the awareness and possibly adequacy of the criteria used, from both managers and employees.
Employee preparations before the dialogue generally benefit motivation and decrease demotivation, seemingly by improving the performance appraisal dialogue quality. Employees who checked notes from last year’s dialogue, reflected on the criteria list, wrote arguments or examples or took notes of performances throughout the year experienced more motivational effect, if also in this case through the intermediation of an improved dialogue quality. However, there were no such effects on demotivation, which may be accounted for by reverse causation. Employees having bad experiences or low expectations on PRP and performance appraisal may choose not to prepare. In addition, preparations that were oriented toward trying to find out what others earn had negative effects on motivation and increased demotivation. This fits previous research stating that a lack of pay secrecy is an institutional obstacle to motivational effects in the public sector (cf. Bellé, 2015). However, this may also be an effect of reverse causality, in that those who suspect or feel they are treated unfairly compared to others tend to be more interested in what their colleagues earn.

Finally, we find that the manager’s perceived evaluation style has effects on motivation and demotivation, which are independent if related to the quality of the performance appraisal dialogue and (general) injustice perceptions. If the line manager was found to be biased or selective in the appraisal, the motivational effect decreased; whereas it increased when the manager was perceived as being comprehensive and listening.

Theoretically, the results confirm aspects of both justice-perceptions (Colquitt, 2001; Malmrud et al., 2020), and SDT (Gerhart & Fang, 2015; Ryan & Deci, 2000). Employee motivation increased with information and fair feedback, and through a good dialogue regarding performances and future possibilities. Thus, the results indicate that non-monetary rewards and feedback (praise) are important aspects of the motivational effect of PRP (Lee, 2019). However, this lies against the backdrop of small monetary rewards from PRP in the Swedish public sector, and previous research indicates that monetary rewards have to be substantial to be motivating (Glassman et al., 2010; Schay & Fisher, 2013). On the other hand, as the present study focused on self-reported (extrinsic) motivation in relation to the goals of doing a better job in the future and getting a pay raise, the results do not contradict research finding that these kinds of PRP pay systems may crowd out intrinsic motivation in the public sector (Papenfuß & Keppeler, 2020; cf. van Loon, 2017; Weske & Schott, 2018).

**Limitations**

There are some obvious limitations regarding the generalizability of these results. The data is cross-sectional and based on self-reported motivation. This makes causal arguments uncertain, and there is a risk that CSB may slightly inflate correlations, even though the Harman’s One-Factor test did not indicate serious problems (cf. George & Pandey, 2017). In addition, there might be some halo effects that are not controlled for adequately (cf. Favero et al., 2016). However, it is a strength that we surveyed three different organizations separately, targeting different levels in the organizations, and
that the separate regressions for each organization produced similar results. Another limitation is that the dependent variables of motivation and demotivation were operationalized as self-reported outcome-focused (extrinsic) de/motivation in relation to two presumed main goals from employers in using PRP: that the employees get motivated to improve their work to increase their pay. Thus, this is a rather narrow construct that does not take into consideration behavioral measures or process-related (means-focused and intrinsic) aspects of motivation (cf. Touré-Tillery & Fishbach, 2014). Finally, analyses of workplace-related effects could be expected to be contextually contingent (Gerhart & Fang, 2015; van Loon, 2017). As we only surveyed Swedish public sector organizations, we cannot control for the effects of national cultures or industrial relations institutions. Therefore, the results need to be further developed and tested in future research.

Implications

A main implication of this study is that the local-level administration and practice in connection to PRP is one strand of research that should be pursued more in the future. Even though the main mechanisms of motivation are well theorized, as are some of the general institutional characteristics of the public sector, it seems that we need more of local contextually empirical studies to understand in what concrete situations employees experience motivation and demotivation from PRP. That is, we suggest that more studies explore the organizational and managerial PRP-practices, and particularly from a comparative perspective, since also national contexts may play a role in this (cf. De Gieter & Hofmans, 2015, p. 200; Gerhart & Fang, 2015, p. 514f.; Papenfuß & Keppeler, 2020, p. 1134f.).

In addition, the above-mentioned results are of practical importance for public sector organizations, since the factors studied are of the kind that they actually can improve—in contrast to the effects of intrinsic motivation values or budget constraints in general in the public sector. The results point to the centrality of the quality of the performance appraisal dialogue, that it evaluates employee performances fairly and systematically, communicates feedback and listens to the employees, and identifies reasonable ways for the employee to improve and develop further. In performance appraisal and PRP systems, for the dialogue quality to be high, both employees and line-managers must be committed to the preparations before the performance appraisal, and be willing to listen to each other.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The study was funded by Vetenskaprådet, The Swedish Research Council.
References

Aneshensel, C. S. (2002). *Theory-based data analysis for the social sciences*. Pine Forge Press.

Baccaro, L., & Howell, C. (2017). *Trajectories of neoliberal transformation. European industrial relations since the 1970s*. Cambridge University Press.

Bellé, N. (2015). Performance-related pay and the crowding out of motivation in the public sector: A randomized field experiment. *Public Administration Review, 75*(2), 230–241.

Bryson, A., Forth, J., & Stokes, L. (2017). How much performance pay is there in the public sector and what are its effects. *Human Resource Management Journal, 27*(4), 581–597.

Chatelain-Ponroy, S., Mignot-Gérard, S., Musselin, C., & Sponem, S. (2018). Is Commitment to performance-based management compatible with commitment to University “Publicness”? academic values in French Universities. *Organization Studies, 39*(10), 1377–1401.

Chen, C.-A. (2018). ‘A little is better than zero’ or ‘Pay enough or Don’t pay at all’? Evidence on the size of pay-for-performance across the sectors. *Public Personnel Management, 47*(2), 119–143.

Choi, S., & Whitford, A. B. (2017). Employee satisfaction in agencies with merit-based pay: Differential effects for three measures. *International Public Management Journal, 20*(3), 442–466.

Colquitt, J. A. (2001). On the dimensionality of organisational justice: A construct validation of a measure. *Journal of Applied Psychology, 86*(3), 386–400.

De Gieter, S., & Hofmans, J. (2015). How reward satisfaction affects employees’ turnover intentions and performance: an individual differences approach. *Human Resource Management Journal, 25*(2), 200–216.

Favero, N., Meier, K. J., & O’Toole, L. J., Jr. (2016). Goals, trust, participation, and feedback: Linking internal management with performance outcomes. *Journal of Public Administration Research and Theory, 26*(2), 327–343.

Fuller, S., & Cooke, L. P. (2018). Workplace variation in fatherhood wage premiums: Do formalization and performance pay matter? *Work, Employment and Society, 32*(4), 768–788.

Garson, D. (2013). *Hierarchical Linear Modeling: Guide and Applications*. Sage.

George, B., & Pandey, S. K. (2017). We know the Yin—But where is the Yang? Toward a balanced approach on common source bias in Public Administration Scholarship. *Review of Public Personnel Administration, 37*(2), 245–270.

Gerhart, B., & Fang, M. (2015). Pay, intrinsic motivation, extrinsic motivation, performance, and creativity in the workplace: Revisiting long-held beliefs. *Annual Review of Organizational Psychology and Organizational Behavior, 2015*(2), 489–521.

Glassman, M., Glassman, A., & Zugelser, M. T. (2010). Evaluating pay-for-performance systems: critical issues for implementation. *Compensation & Benefits Review, 42*(2), 231–238.

Kim, T., & Holzer, M. (2016). Public employees and performance appraisal: A study of antecedents to employees’ perception of the process. *Review of Public Personnel Administration, 36*(1), 31–56.
Kjellberg, A. (2019). Sweden: Collective bargaining under the industry norm. In T. Müller, K. Vandaele, & J. Waddington (Eds.), Collective bargaining in Europe: Towards an endgame. ETUI.

Lee, H.-W. (2019). Performance-based human resource management and federal employee’s roles of goal-clarifying intervention, appraisal fairness, and feedback satisfaction. Review of Public Personnel Administration, 39(3), 323–348.

Malmrud, S., Falkenberg, H., Eib, C., Hellgren, J., & Sverke, M. (2020). Just what I see? Implications of congruence between supervisors’ and employees’ perceptions of pay justice for employees’ work-related attitudes and behaviors. Frontiers in Psychology, 11(2069), 1–16.

Ogbonnaya, C., Daniels, K., & Nielsen, K. (2017). Does contingent pay encourage positive employee attitudes and intensify work? Human Resource Management Journal, 27(1), 94–112.

Oh, S. S., & Lewis, G. B. (2009). Can performance appraisal systems inspire intrinsically Motivated employees? Review of Public Personnel Administration, 29(2), 158–167.

Papenfuß, U., & Keppeler, F. (2020). Does performance-related pay and public service motivation research treat state-owned enterprises like a neglected Cinderella? A systematic literature review and agenda for future research on performance effects. Public Management Review 22(7), 1119–1145.

Perry, J. L., Engbers, T. A., & Jun, S. Y. (2009). Back to the future? Performance-related pay, empirical research, and the perils of persistence. Public Administration Review, 69(1), 39–51.

Ryan, R. K., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemporary Educational Psychology, 25(1), 54–67.

Schay, B. W., & Fisher, S. F. (2013). The challenge of making performance-based pay systems work in the public sector. Public Personnel Management, 42(3), 359–384.

Stråberg, T. (2010). Employee perspectives on individualized pay. Attitudes and fairness perceptions. Stockholm University.

Touré-Tillery, M., & Fishbach, A. (2014). How to measure motivation: A guide for experimental social psychologist. Social and Personality Psychology Compass, 8(7), 328–341.

Ulfsdotter Eriksson, Y., Larsson, B., & Adolfsson, P. (2020a). Under the surface of individual and differentiated pay in Sweden: A Zero-sum game of performance-based Pay?, British Journal of Industrial Relations. Advanced online publication. https://doi.org/10.1111/bjir.12561

Ulfsdotter Eriksson, Y., Larsson, B., & Adolfsson, P. (2020b). Employees of greatness: Signifying values in performance appraisal criteria. Nordic Journal of Working Life Studies. Advanced online publication. https://doi.org/10.18291/njwls.122589

Vasset, F., Marnburg, E., & Furunes, T. (2011). The effects of performance appraisal in the Norwegian municipal health services: A case study. Human Resources for Health, 9(22), 1–12.

van Loon, N. M. (2017). Does context matter for the type of performance-related behavior of public service motivated employees? Review of Public Personnel Administration, 37(4), 405–429.

Wang, K. Y., & Teo, S. T. T. (2008). Performance appraisal. In S. R. Clegg, & J. R. Bailey (Eds.), International Encyclopedia of Organization Studies. SAGE.

Weibel, A., Rost, K., & Osterloh, M. (2009). Pay for performance in the public sector—benefits and (Hidden) costs. Journal of Public Administration Research and Theory, 20(2), 387–412.
Wenzel, A.-K., Krause, T. A., & Vogel, D. (2019). Making performance pay work: The impact of transparency, participation, and fairness on controlling perception and intrinsic motivation. *Review of Personnel Administration, 39*(2), 232–255.

Weske, U., & Schott, C. (2018). What motivates different groups of public employees working for dutch municipalities? Combining autonomous and controlled types of motivation. *Review of Public Personnel Administration, 38*(4), 415–430.

Williams, M., Zhou, Y., & Zou, M. (2019). The rise in pay for performance among higher managerial and professional occupations in Britain: Eroding or enhancing the service relationship? *Work, Employment and Society*. Advance online publication. https://doi.org/10.1177/0950017019841552.

**Author Biographies**

**Bengt Larsson**, is professor of sociology at the Department of sociology and work science, University of Gothenburg, Sweden, and at the Department of social studies, Linnaeus University, Sweden.

**Ylva Ulfsdotter Eriksson**, is associate professor of sociology at the Department of sociology and work science, University of Gothenburg, Sweden, and at the Department of social studies, Linnaeus University, Sweden.

**Petra Adolfsson**, is associate professor of Business administration at the School of Business, Economics and Law, University of Gothenburg, Sweden.