

# Performance appraisal justice and employees' work engagement in the public sector: Making the most of performance appraisal design

Lorenza Micacchi  | Francesco Vidé  | Giorgio Giacomelli  |  
Marta Barbieri 

Government, Health and Not for Profit (GHNP), SDA Bocconi School of Management, Milan, Italy

## Correspondence

Giorgio Giacomelli, Government, Health and Not for Profit (GHNP), SDA Bocconi School of Management, Milan, Italy.  
Email: [giorgio.giacomelli@sdabocconi.it](mailto:giorgio.giacomelli@sdabocconi.it)

## Funding information

Department of Public Administration (*Dipartimento della Funzione Pubblica*) of the Italian Presidency of the Council of Ministers

## Abstract

Regarding organizational justice, performance appraisal (PA) systems are among the most crucial mechanisms shaping the employee work experience. Nonetheless, research has mostly neglected to explain the relationship between the characteristics of PA systems, employees' perceptions of PA justice, and work engagement in the public sector. Based on a combination of observational and experimental data from 11 public sector organizations in Italy, this article aims to unfold the nature of these relationships. Following the recent calls for diversified methodological designs, a combination of structural equation modeling with a discrete choice experiment is provided. The results show that designing PA systems with rater–ratee feedback sessions and calibration mechanisms leads to fairer perceived PAs, with PA justice being associated with public employees' work engagement.

## Abstract

Quando si parla di equità organizzativa, i sistemi di valutazione delle performance sono tra i meccanismi più determinanti a plasmare l'esperienza lavorativa dei dipendenti. Tuttavia, la ricerca ha spesso trascurato la relazione tra le caratteristiche dei sistemi di valutazione delle performance,

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Public Administration* published by John Wiley & Sons Ltd.

la percezione di equità del sistema stesso da parte dei dipendenti e il livello di engagement lavorativo del personale nel settore pubblico. Questo articolo si propone di esaminare la natura di tali relazioni attraverso una combinazione di analisi osservazionali e sperimentali condotte sulla base dei dati raccolti in 11 amministrazioni pubbliche in Italia. In linea con la crescente tendenza all'adozione di approcci metodologici misti, viene proposta un'analisi basata sulla combinazione di un modello di equazioni strutturali e un esperimento di scelta discreta. I risultati indicano che la progettazione di sistemi di valutazione delle performance che includono sessioni di feedback tra valutatori e valutati, unitamente all'adozione di meccanismi di calibrazione, favorisce una maggiore percezione di equità dei sistemi e, di conseguenza, un maggiore livello di engagement del personale.

## 1 | INTRODUCTION

Engagement has gained wide attention in management scholarship during the last thirty years and proved essential for managerial practice, as it relates to crucial individual and organizational outcomes, regardless of the sector and industry of interest (Borst et al., 2020). Employees who are highly engaged tend to perform better (Anitha, 2014), are more committed to their organization (Hakanen et al., 2008), and are less often absent from work (Schaufeli et al., 2009). At the same time, numerous sources of complexity are nowadays challenging public administrations, such as increasing service demands against shrinking resources and fiscal constraints (Cepiku et al., 2016) and decreasing effectiveness and legitimacy of governance institutions (Baimyrzaeva & Meyer, 2021). Meeting these challenges requires public organizations with strong administrative and management capacity, which entails motivated, knowledgeable, and experienced human resources (Christensen & Gazley, 2008). Under such conditions, fostering civil servants' engagement is paramount for public sector organizational effectiveness, as it fuels civil servants' commitment to increasing levels of quality service, which is “essential for building modern public administration” (Vigoda-Gadot et al., 2013, p. 523).

In this vein, performance appraisal (PA) systems might play a key role in driving employee behavior, ensuring positive individual outcomes (Kim & Rubianty, 2011), thus ultimately resulting in higher performance (Mani, 2001). However, the academic literature on the topic provides a controversial image of PA systems in the public sector as scholars have long highlighted mixed results about the relationship between PA and motivation and, eventually, the quality of public services (Marsden & Richardson, 1994). This dilemma might be explained by context-specific issues related to the use of PA in the public domain, such as the crowding-out effect of performance pay, the ethical issues in the performance assessment of civil servants, the lack of rating differentiation, and the low amount of related pays (Wenzel et al., 2019).

Over time, scholarship has dealt with the PA dilemma forming into two main research lines: (i) how the PA system is structured and (ii) what its effectiveness depends on. Concerning the former, previous academic contributions proposed different classifications of the structure of PA systems (e.g., Barbieri et al., 2021; Hajnal & Staronova, 2021) and suggested that much remains to be done in order to understand how the design of PA systems relates to employee perceptions (Gupta & Kumar, 2013). Therefore, analyzing the impact of structural characteristics of the PA system on PA justice represents a valuable area of research. In this respect, the public sector is a still an

under-investigated field, despite specificities that are likely to influence PA effectiveness and individual employee outcomes, such as civil servants' peculiar motivational profile (Quratulain et al., 2019).

Regarding PA effectiveness, scholars have tried to come to a finer-grained understanding of the construct by offering different theoretical frameworks and possible measurement criteria (Levy & Williams, 2004). Several studies focus on employee reactions and their perceptions of PA justice<sup>1</sup> (Kim & Holzer, 2016). However, these studies mainly relate employee perception of the PA system to individual characteristics, such as intrinsic and extrinsic motivation (Kim & Rubianty, 2011); other scholars link it to individual outcomes such as job satisfaction (Quratulain et al., 2019), but the relationship between employee perception of the PA system and engagement still lacks investigation, especially in the public sector. Nevertheless, this relationship might represent a promising area for research for public organizations, provided that civil servants are characterized by a “unique form of engagement” (Levitats et al., 2019) and the academic literature has investigated organizational justice among the possible antecedents of engagement (Fletcher et al., 2020), although without explicitly referring to PA justice.

Therefore, to contribute to the discussion on PA in the public sector, the present study aims to bridge previous contributions on the characteristics of the PA system and its effectiveness, as measured by employee perception of justice, discussing how they relate to civil servants' engagement.

On that account, the study addresses the following two research questions: (i) whether perceived PA justice is associated with work engagement, and whether this effect is mediated by satisfaction with the PA system, and (ii) what structural characteristics of the PA system relate to PA justice perceptions. In addressing the questions above, the authors leverage social exchange theory (SET), taking a perspective that views engagement as a “concept that is underpinned by the social exchange relationship between the employee and the employer” (Fletcher et al., 2016, p. 7).

The remainder of this article is structured as follows. Section 2 presents the theoretical background, which informs the hypotheses posited by the study. In Section 3, the methodology is discussed. The results of the analyses are presented in Section 4. In Sections 5 and 6, we discuss the emerging implications for practice on employee engagement in the public sector and then highlight possible avenues for further research in public management scholarship.

## 2 | THEORETICAL BACKGROUND

This section frames the research through the lenses of the SET and provides a critical review of the state of art of extant knowledge on the dependent variable (work engagement), the main antecedent (PA justice) and the mediator (satisfaction with the PA system), as well as their potential relationships.

### 2.1 | Organizational justice and work engagement: A social exchange view

SET is a widely adopted conceptual framework in organizational research (Cropanzano et al., 2017) and can be used to frame the relationships among the constructs used in our study. Over time, SET has been adapting to the changes of workplace relationships (Chernyak-Hai & Rabenu, 2018) and has increasingly expanded its original scope toward broader commitments. Departing from a fundamentally rational (i.e., transactional) logic (Blau, 1964), SET has been adopted more and more to frame the psychological dynamic mechanisms underpinning workplace relationships (Cropanzano & Mitchell, 2005) and has been shown to rest largely on an emotional rather than just a rational basis (Colquitt et al., 2013; Lawler, 2001). This intertwining between the emotional outcomes (what happens in people's head) and the perceived congruence of the outcomes of external processes with certain shared ethical norms (what happens in the environment), has been labeled by scholars (Cropanzano et al., 2000) as a cognitive appraisal model of organizational justice (or, later on, fairness theory). The caliber of the emotional dimension in these processes is

such as to suggest that “individuals’ own phenomenological understanding of unfairness is at least as much affective as it is cognitive or behavioral” (*ibidem*, p. 50). By generating unspecified obligations, social exchange relationships lean on both reasoning and feeling reciprocity between one’s behaviors and how things work in the workplace, a concept referred to as homeomorphic reciprocity (Lyons & Scott, 2012). In simple terms, employees are likely to experience positive (or negative) attitudes about their work when they perceive fair (or unfair) treatment, and to react accordingly (Cropanzano et al., 2017); such sense of fairness (or lack thereof) tends to build up over time in a dynamic process in which the relationship is not the outcome of a single stimulus-response, but rather a series of interdependent exchanges that resembles “climbing a ladder.” (Cropanzano & Mitchell, 2005, p. 890). In broader terms, being faced with work situations that could trigger socioemotional outcomes such as pride (as a result of a favorable outcome either after fair procedures or overcoming unfavorably biased ones), happiness, guilt, or anger (Weiss et al., 1999), accumulates over time in a way that such emotional concretions end up shaping how an individual relates and engages at work. In other words, including both emotional aspects such as dedication and a high intensity affect (excitement), work engagement might be a consequence of positive states such as enthusiasm, pride and cheerfulness. These emotional states have been linked by the academic literature to organizational justice (Colquitt et al., 2013; Weiss et al., 1999) since, according to appraisal theories, justice perceptions lead to positive state affects resulting in positive emotions (Cropanzano et al., 2000), thus potentially fostering a higher level of work engagement.

## 2.2 | Engagement in the public sector: Crucial yet under-investigated

Engagement has been considered one of the “magic words” of human resource management for the last three decades, after first entering scholarly vocabulary with Kahn’s (1990) seminal study. Kahn defined engagement by the extent to which people employ physical, cognitive, and emotional degrees of themselves at work. Starting from Kahn’s study, several definitions and measures have emerged. This self-role coupling (or uncoupling) process has subsequently been considered as the defining feature of the *need-satisfying* approach of engagement (Shuck, 2011), also referred to as *personal role* engagement (Fletcher et al., 2016). Besides this, Shuck (2011) classified three more perspectives on engagement emerging from the literature, namely the *satisfaction-engagement* approach, the *multi-dimensional* approach and the *burnout-antithesis* approach. The *satisfaction-engagement* approach focuses on engagement as the “individual’s involvement and satisfaction with, as well as enthusiasm for, work” (Harter et al., 2002, p. 269). The *multidimensional* approach refers to Saks’ contribution (Saks, 2006) and suggests engagement having multiple foci: cognitive, emotional, and behavioral. The *burnout-antithesis* approach relates to the literature on the effects of job-based emotional and interpersonal exhaustion (Maslach & Leiter, 2008) and initially considered burnout and engagement as opposite poles of a continuum. Eventually, Fletcher et al. (2016) renamed this research strand as the *work engagement* approach after the contribution of Demerouti and colleagues (2001), valuing engagement as a distinct concept “characterized by vigor, dedication and absorption” (Schaufeli et al., 2002, p. 74), rather than the opposite of job burnout. According to Schaufeli and colleagues, the dimension of vigor is linked to both energy and mental resilience, being characterized by “persistence even in the face of difficulty” (*ibidem*, p. 74), dedication is characterized as “a sense of significance, enthusiasm, inspiration, pride, and challenge” (*ibidem*, p. 74) while absorption relies on “being fully concentrated and deeply engrossed in one’s work” (*ibidem*, p. 74).

Despite a large amount of measures employed in the literature, scholars currently agree that engagement is a distinct construct consisting of cognitive, emotional, and behavioral components associated with individual role performance (Saks, 2006). Recent contributions have also highlighted the need to focus on the specificities of engagement in the public sector, reaffirming the construct’s multidimensionality (Levitats & Vigoda-Gadot, 2020) while suggesting the importance of the affective link between the employees and the organization, “even in the face of the public sector’s unique challenges” (Levitats et al., 2019, p. 842). In this vein, work

engagement as measured by the Utrecht Work Engagement Scale (Schaufeli et al., 2002) represents a key construct also for the public sector.

When it comes to the determinants of engagement, a wide range of academic contribution has pinpointed the importance of the different resources made available by the organization in crafting employees' working experience (Borst et al., 2019). In such perspective, attention should be paid to PA characteristics: the PA process should indeed provide employees with resources conducive to promote engagement "by fostering the psychological conditions that precede it" (Gruman & Saks, 2011, p. 127). In other words, according to the SET (Blau, 1964; Cropanzano & Mitchell, 2005), PA might represent a relevant resource provided by the organization, which might be reciprocated with positive attitudes from employees.

## 2.3 | Organizational justice and performance appraisal

Organizational justice can be defined as employees' personal and subjective beliefs on the fairness of managerial events, practices, and conduct (Cropanzano & Mitchell, 2005). In this vein, it mainly focuses on what employees perceive as fair, thus resulting in major interest as to why employees perceive organizational practices as just, as well as the consequences produced by such perceptions.

Relying on the SET (Blau, 1964; Cropanzano & Mitchell, 2005), the alignment between individual contribution and organizational rewards tied to the PA system has been assessed from different perspectives, addressing employee reactions to PA through distributive, procedural, and interactional justice constructs (Greenberg, 1986), which might result in employees engaging in mutual obligations toward their organization.

Regarding *distributive justice*, employees compare the results of PA with their contribution, as well as that provided by their colleagues, considering unfair those outcomes that are lower than expected (Kim & Rubianty, 2011). At the same time, the perceived justice of PA strongly depends on *procedural* variables, since employees tend to perceive higher rating accuracy when supervisors support employees during the entire PA process (Landy et al., 1978). Furthermore, the relationship between rater and ratee affects the perceived *interactional justice* of PA: if employees notice uncaring interpersonal treatment from their supervisors during the appraisal process, it could undermine the quality of rater–ratee interaction and their perceptions of PA fairness. From this perspective, perceived PA justice might represent a relevant explanatory variable when it comes to individual attitudes toward the organization, being especially key for the public sector in order to guarantee "the immersion of individuals into their duties" (Jang et al., 2021, p. 1).

## 2.4 | PA justice and engagement

Management scholars have extensively explored the relationship between organizational justice and several individual outcomes, such as trust and commitment (Colquitt et al., 2001), and organizational citizenship behavior (Cohen-Charash & Spector, 2001), showing that organizational justice should be accounted for when "evaluating the employment relationship quality and organizational contributions to the exchange relationship" (Ko & Hur, 2014, p. 178).

Among the different constructs related to organizational justice, public sector scholars have traditionally considered the impact of distributive justice on several different outcomes (Cho & Sai, 2013). Recently, scholars focused on the importance of procedural justice in affecting public employee identification with the organization (Quratulain et al., 2019), especially focusing on job satisfaction, turnover intention, and cooperation (Rubin & Kellough, 2012). Eventually, interactional justice has been addressed, which is the most recently defined as well as the least investigated (Choi, 2011), being mainly related to work attitudes such as organizational citizenship behavior (Chen & Jin, 2014), and employee misbehavior (de Schrijver et al., 2010).

Among the individual work attitudes analyzed, work engagement still presents room for further empirical analyses. Organizational justice indeed affects the quality of social interconnections between organization and individuals (Bhatnagar & Biswas, 2010), thus leading employees with high organizational justice to feel obliged to reciprocate the organization with higher engagement. From this perspective, as proven by Biswas et al. (2013), distributive and procedural justice perceptions might influence work engagement through social exchange mediators, such as perceived organizational support, which foster individuals to repay fair relationships with positive work behaviors and attitudes (Cropanzano & Mitchell, 2005).

As a result, scientific research has identified organizational justice as a potential antecedent of engagement (He et al., 2014), but few studies to date specifically explored this relationship, thus opening room for further research, especially when it comes to perceptions about PA (Nair & Salleh, 2015). This is particularly true for public sector organizations: among the few contributions on the topic, to the authors' knowledge the impact of perceived PA justice on engagement has only been addressed with regard to the Chinese administrative context and to merit pay systems (Meng & Wu, 2015), thus still lacking a more generalized approach. Yet, this relationship is likely to be insightful in the public domain, at least for two reasons: (i) performance evaluation can be particularly complex in public organizations (Lovrich et al., 1980), and (ii) among other individual work attitudes, engagement is a key driver for increasing attitudinal, behavioral, and performance outcomes within the public sector (Borst et al., 2020). With all this provided, the following hypothesis applies:

**Hypothesis 1.** PA justice has a positive direct effect on work engagement.

## 2.5 | Satisfaction with PA system

PA satisfaction represents a critical determinant for the effectiveness of PA systems itself (Mulvaney, 2019) as well as for several employee outcomes, such as organizational commitment and turnover intention (Kuvaas, 2006). Scholarly conceptualization of PA satisfaction can be summarized as threefold (Keeping & Levy, 2000): satisfaction with the appraisal interview, with performance ratings, and with the system. Concerning the latter, empirical studies prove that employees' satisfaction with the PA system is different from their attitudes toward the appraisal interview and discussion process (Mount, 1984), thus supporting the conceptualization of a distinctive and self-standing construct. Employees' satisfaction with the PA system refers to the "the level of satisfaction an employee has toward the agency's entire appraisal system and represents a [...] global measure of the appraisal system, including the instrument, interview, and subsequent actions following the interview" (Mulvaney, 2019, p. 204), being also investigated within the public sector (Roberts & Reed, 1996).

Based on the academic consensus about the importance of employee satisfaction with the system, scholars have investigated its possible antecedents. Among these, scholars have underlined the relevance of employees' beliefs that their performance is assessed in a "fair, valid, and accurate manner" (Kim & Holzer, 2016), thus referring to justice perceptions (Gruman & Saks, 2011). The higher the degree of distributive, procedural, and interactional justice, the higher the satisfaction level with the overall PA system, regardless of performance ratings and appraisal perceived usefulness (Dusterhoff et al., 2014). In this vein, the academic literature has mostly focused on employee perceptions about (i) whether their performance is fairly measured (i.e., *procedural justice*) and (ii) how much their performance is linked to rewards (i.e., *distributive justice*). More recent contributions have also underlined the importance of social interactions within the PA system with respect to employee dissatisfaction with PA (Thurston & McNall, 2010). However, based on the authors' best knowledge, no study seems to address the comprehensive effect of PA justice' components on employees' satisfaction with the PA system.

Moreover, recent studies isolated the effect of employee satisfaction with PA as a significant predictor of work engagement (Memon et al., 2019), since employees experiencing a fair and unbiased evaluation might be more

willing to reciprocate their organization with higher engagement. However, no contribution seems to address the link between satisfaction with the PA system and engagement within the public sector.

Therefore, this study aims at filling these research gaps, assessing the mediating effect of the overall satisfaction with the PA system on the relationship between organizational justice and work engagement, as well as investigating a still unexplored context such as the public sector. Based on these premises, the authors postulated the following hypothesis:

**Hypothesis 2.** PA justice has an indirect effect on engagement as mediated by satisfaction with the PA system.

## 2.6 | Characteristics of PA systems and employees' reactions

Besides employee perceptions on the PA system, for at least three decades managerial scholars have mainly focused on the impact of the structural characteristics of the system on its effectiveness (Levy & Williams, 2004). Relying on the theoretical framework proposed by the two authors, subsequent theoretical and empirical studies on the design of PA systems have been traditionally focused on the (i) *structural* (concerning the configuration of the PA system) and the (ii) *process* (regarding the appraisal process) proximal variables, both being positively related to rater and ratee behaviors.

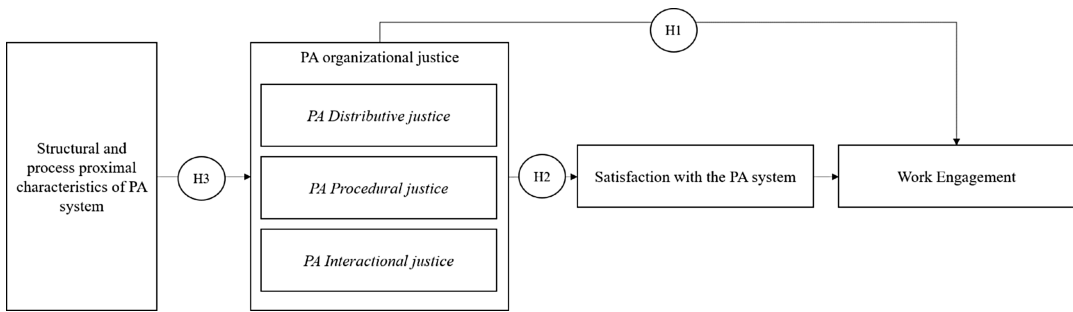
Despite the large number of studies dedicated to both structural and process variables, previous academic contributions have commonly addressed the impact of those variables on outcomes such as rating accuracy and the absence of errors in the appraisal (Murphy & Cleveland, 1995). More recently, in line with calls for further exploration of the PA social context (Levy & Williams, 2004), several studies have analyzed the impact of both structural and process PA characteristics on employee reactions, provided that the PA system should be not only well designed but should also “satisfy the needs of the parties involved in the process” (Longenecker & Nykodym, 1996, p. 2).

Regarding the *structural* characteristics, contributions have traditionally explored the effects of rating formats (Tziner & Kopelman, 2002) and performance dimensions (Taormina & Gao, 2009) on employee reactions. Furthermore, at least in the private sector, a large number of studies focused on rating sources, especially on multisource feedback, by investigating the effects on employee reactions. Overall, this research shares a private-sector focus, with only one study assessing the impact of multisource feedback on employee perceptions in the public domain (DeLeon & Ewen, 1997).

A similar pattern applies to the *process* variables, with prominent focus on rater accountability, which deals with raters being held responsible for their decisions and behaviors (Frink & Ferris, 1998). Within this premise, different types of rater accountability entered into the picture, being classified by academic literature as *downward* accountability, that is, toward supervised ratee, and *upward* accountability, that is, toward the rater's supervisors or peers (Curtis et al., 2005).

With reference to the former, scholars have mostly focused on rater–ratee feedback sessions, which can contribute to strengthen their relationship (Kuvaas et al., 2017) and might be associated with positive employee attitudes toward the PA system. In fact, having the chance to voice their opinions and share viewpoints, employees might show a higher level of satisfaction and acceptance of the appraisal (Williams & Levy, 2000). This relationship has also been studied within the public sector (Roberts, 2003), being positively linked to interactional justice perception and, ultimately, to PA effectiveness.

However, raters should be accountable not only to their subordinates but also to their peers, in order to drive out errors and increase rating accuracy. For instance, calibration mechanisms (i.e., meetings where raters have the possibility to compare the given ratings) could be an important managerial tool—although less investigated by the academic literature—in order to avoid bias related to PA. Empirical contribution on this topic can be mainly found within the private sector and mostly related to PA effectiveness criteria such as rating quality (Mero et al., 2007). At



**FIGURE 1** Conceptual model.

the best of authors' knowledge, no contribution specifically addresses the impact of *upward* accountability, such as the one accomplished through calibration mechanisms, on employee reactions. However, this could represent a powerful tool in assuring a higher level of PA procedural and interactional justice, thus leading to satisfaction with the system, a stronger relationship with raters and/or supervisors, and positive individual work-related outcomes. In fact, through calibration meetings, raters compare with their peers the performance standards identified at the beginning of the appraisal process, as well as the final ratings assigned, thus reducing evaluation bias and being able to clearly justify their choices. In this perspective, ratees might be keener to accept their rater's evaluation and perceive PA as fair. In other words, it represents a promising area of research especially for public organizations, where the performance gap (i.e., the gap between the desired performance perceived as fair and the actual one) is a relevant predictor of employee productivity (Mani, 2001).

To conclude, both structural (rating source) and process proximal variables (rater–ratee feedback sessions, and rater upward accountability through calibration mechanisms) can predict positive individual outcomes (i.e., higher level of work engagement) through their effect on PA justice, thus being particularly relevant for performance evaluation in the public sector while representing a still unexplored field of research. From this, the authors propose the following research hypothesis:

**Hypothesis 3.** Structural and procedural characteristics of the PA system influence employee perceptions of PA justice.

In line with the aforementioned contribution, the authors tested the research hypotheses (as graphically summarized by the relational model in Figure 1) through a combination of observational and experimental analyses.

### 3 | DATA AND METHODS

The study employs a multi-method approach in order to address the two research questions with different quantitative methodologies, combining both an observational and an experimental design. The relationship between work engagement and PA, as well as that between the latter and its antecedents, is indeed complex, and caution in interpreting it is imperative due to a number of methodological and practical pitfalls. From a methodological point of view, the use of cross-sectional data alone to investigate such associations may be prone to the risks of omitted variable bias and reverse causality; cross-sectional analyses, in fact, cannot rule out the existence of potentially confounding results nor to determine the direction of the causal relationship (Bullock et al., 2010). On a practical level, the multi-faceted nature of the considered constructs implies that any managerial recommendations based on a simplistic interpretation on their interplay could be severely misleading.



Our multi-method approach draws from two main sources of data: after having gathered hints on the associations between PA justice, satisfaction with the PA system and work engagement through an initial observational analysis, we looked for a finer-grained understanding of the determinants triggering employees' perception of PA justice via an experimental design. The choice of these two approaches follows the characteristics of the phenomena to be observed, as well as the different nature of the research questions in the study.

First, structural equation modeling (SEM) was conducted to assess the relationship among PA justice and work engagement, and the hypothesized mediational effect of satisfaction with the PA system. However, applying SEM, despite “the virtue of modeling measurement error in sophisticated ways” (Bullock et al., 2010, p. 552), does not decisively rule out the methodological issues associated with cross-sectional designs, namely omitted variable bias and the inability to draw causal conclusions: following the recommendations suggested by Bullock et al. (2010), we probed the results of our analysis by addressing possible heterogeneous effects via a group-based analysis of the mediation model.

Second, a discrete choice experiment (DCE) was administered to a quota of the respondents to the survey, in order to elicit employees' preferences over PA systems' characteristics, and elucidate their causal contribution to a perception of PA justice.

### 3.1 | Characteristics of the sample

Both the observational and experimental analyses were administered via computer-assisted web interviewing. The questionnaire was sent to 4933 civil servants<sup>2</sup>. The respondents who did not confirm having received a formal individual appraisal review were dropped from the sample. Eventually, the interested sample population consisted of 4182 employees with 1483 valid responses collected (35.4% response rate).

The 1483 respondents to the survey were also invited to complete the DCE. Overall, more than 230 public employees who took part to the survey also completed the experiment. Provided that the anonymity of all responses—both for the survey and for the DCE—was guaranteed, a screening question was inserted in the DCE asking the participants whether they had completed the previous survey. Only employees who confirmed answering to the survey, were allowed to participate to the DCE. After controlling for attention check questions, the overall number of valid responses to the DCE was equal to 209. Table 1 provides a synthesis of the characteristics of the 1483 survey's respondents and of the 209 DCE respondents, compared to the characteristics of the investigated population (employees working for the considered organizations who received an individual PA).

Eventually, we tested the representativeness of (i) the survey sample with respect to the population and (ii) the DCE sample with respect to both the survey sample and the overall population. Both the survey and the DCE sample were not statistically significant different from the overall population, in terms of age and role. However, the DCE sample was statistically different from the survey sample and the population in terms of gender; nevertheless, we did not expect differences in respondents' gender to potentially bias the analysis on structural characteristics of the PA system. For instance, Gosselin et al. (1997) probed for gender differences in PA characteristics and found no evidence when it comes to those variables considered in our study (e.g., frequency of rating, source of appraisal). Table 2 reports the statistical results of these comparisons.

No missing data were reported for both the observational or experimental analysis, being a good proxy of the effectiveness of data collection (Klassen & Jacobs, 2001).

### 3.2 | Observational analysis

We started by running SEM analyses on data collected in different timespans, from September 2019 to March 2020, since the survey distribution was directly managed by the organizations involved. The survey was run as part of a

**TABLE 1** Characteristics of the sample.

Baseline characteristics	Survey sample		DCE sample		Population sample	
	n	%	n	%	n	%
<i>Gender</i>						
Male	490	33%	84	40%	1422	33%
Female	993	67%	125	60%	2760	67%
<i>Age</i>						
20–39	96	6%	16	7%	292	6%
40–54	776	52%	108	52%	2133	51%
>54	611	42%	85	41%	1757	43%
<i>Seniority</i>						
Regular	1433	97%	200	96%	4056	97%
Senior	50	3%	9	4%	126	3%

Abbreviation: DCE, discrete choice experiment.

**TABLE 2** Results of the chi-squared test for samples representativeness.

		$\chi^2$	df	p-Value
Comparison 1 Survey sample—Population	Comparison in terms of gender	0.611	1	0.434
	Comparison in terms of age	1.292	2	0.524
	Comparison in terms of seniority	0.653	1	0.419
Comparison 2 DCE sample—Population	Comparison in terms of gender	3.567	1	0.059
	Comparison in terms of age	0.244	2	0.885
	Comparison in terms of seniority	1.196	1	0.274
Comparison 3 DCE sample—Survey sample	Comparison in terms of gender	4.830	1	0.028
	Comparison in terms of age	0.482	2	0.786
	Comparison in terms of seniority	0.560	2	0.454

Abbreviation: DCE, discrete choice experiment.

research project funded by the Department of Public Administration of the Italian Presidency of the Council of Ministers. The key variables to be explained are PA justice, satisfaction with the PA system, and work engagement. All variables considered in the study were constructs grounded in literature and measured using multiple items.

To measure *PA justice*, we followed Kim and Rubianty (2011) and used their three-dimensional scale that singles out *procedural*, *distributive*, and *interactional justice*. All items elicited respondent perceptions on a 5-point Likert scale ranging from “strongly disagree” (1) to “strongly agree” (5).

With respect to *satisfaction with the PA system*, the three items of the “satisfaction with the overall system” measure proposed by Mulvaney (2019) were used, as derived from Keeping and Levy’s (2000) 14 items PA reaction instrument, who suggested the reflective nature of the construct. Each item was measured through a 6-point Likert scale from 1 (strongly disagree) to 6 (strongly agree).

The dependent variable (i.e., *work engagement*) was measured by adopting the UWES extended scale by Schaufeli et al. (2006), as already translated into Italian and tested by Balducci et al. (2010). We followed Utrecht group’s conceptualization based on existing evidence of a strong association of the three-dimensional scale

(i.e., vigor, dedication, and absorption) with PA justice (Gupta & Kumar, 2013; Meng & Wu, 2015). All items used a 7-point Likert scale ranging from “never” (0) to “always” (6).

All the aforementioned measures, whenever not directly available in Italian, were translated following a forward and back translation process (Eremenco et al., 2005). Eventually, an expert panel revised the final version of the questionnaire in order to account for possible specificities of the Italian normative and administrative context.

To test our hypotheses, we fitted an SEM associating PA justice (independent variable) and work engagement (dependent variable), with satisfaction with the PA system posited to act as mediator. All statistical analyses of this phase of the study were performed with IBM AMOS 27 software.

### 3.3 | Discrete choice experiment

We then administered a DCE to employees who previously responded to the survey. DCE has been recently gaining traction in public administration scholarship (Battaglio et al., 2021) and entails presenting participants with a series of choice sets, each consisting of two or more scenarios differing along a set of attributes. For each choice set, respondents are asked to choose one scenario. This allows for the estimation of relative impact that each attribute has on participants' stated preferences. DCEs are particularly well suited to represent real-world complex decisions involving trade-offs (Ryan et al., 2012). In fact, with these types of choices alternative scenarios are bundles of attributes that differ along multiple dimensions, rather than being a single factor. Following the approach adopted in previous studies conducted by scholars in the field (Bellé & Cantarelli, 2018b), our DCE adopts a paired conjoint design with a forced choice that entails presenting respondents with two alternatives (i.e., a choice-set), differing along multiple attributes and asking them to pick the preferred option. Extant scholarly research highlights several advantages of conjoint design, compared to alternative experimental designs (Battaglio et al., 2021), such as: (i) the greater realism of information-rich decision environment, (ii) the efficiency in estimating the simultaneous and relative effects of different attributes on a single behavioral outcome, and (iii) a mitigated risk of raising social desirability bias.

We manipulated three attributes on two levels, thus generating eight unique job scenarios. Using a fold-over approach (Street et al., 2005), we generated eight choice sets by pairing each scenario with its mirror design obtained by shifting every attribute to its alternative level, guaranteeing minimum overlap across options and thus maximizing the information drawn out from respondents. Respondents had to choose between the two job scenarios proposed for each choice set; in other words, no opt-out scenario was included in the analysis. We displayed to the respondents the eight different choice sets. In line with academic literature recommendations on DCE, both orthogonality and level balance were ensured.

Based on the results of the SEM we selected three attributes, each identifying specific characteristics of the PA system (*rating source*, *upward rater accountability*, and *rater-ratee feedback sessions*), and measured their effect on respondents' stated preferences with regard to the PA system that they would consider as fairest. As an operation of the target construct *rating source*, we manipulated the case of one's job being evaluated by one's direct supervisor (i.e., *hierarchical rating*) or by one's supervisor, peers, and reports (i.e., *multisource feedback*). As an operation of the target construct of *upward rater accountability* (i.e., *calibration mechanisms*), we manipulated the presence or absence of meetings among evaluators aimed at both (i) calibrating different evaluation criteria and (ii) comparing the assigned scores. We manipulated the target construct *rater-ratee feedback sessions* on two levels: no feedback meeting with one's evaluator to comment on the evaluation results vs. at least two feedback meetings per year (devoted to assigning goals and commenting on PA). Table 3 reports the attributes and the attribute levels for the DCE.

The choice made by the participants in a given choice set (the dependent variable) is either A or B, depending on the selected level of the attributes. We consequently conducted a conditional fixed effects logistic regression, as this model is shown to be consistent with random utility theory (McFadden, 1973). The conditional logistic regression groups data and then estimates the likelihood relative to each group. To estimate conditional logit, we used *clogit*

**TABLE 3** Design of the DCE. “Which PA system, between the following, would you perceive as fairer?” (*Outcome variable*).

Attributes	System A	System B
Rating source	Your performance appraisal rating is expressed by your supervisor	Your performance appraisal rating is expressed by your supervisor, your colleagues and those reporting to you (if any)
Upward rater accountability (i.e., calibration)	The system does not account for meeting among raters in order to compare the assigned performance ratings.	The system does account for two meeting among raters: the first one to discuss performance standards and the second one to compare the assigned performance ratings.
Rater–ratee feedback sessions	During the year, there is no feedback session with your rater(s) to discuss your performance.	During the year, there are at least two feedback sessions with your rater(s): one at the beginning of the year and one after the final evaluation.

Abbreviations: DCE, discrete choice experiment; PA, performance appraisal.

command in Stata 17, which is computationally equivalent to a fixed-effects logit. This command considers each respondent's group of answers when making fitting the conditional logistic regression.

## 4 | RESULTS

The following paragraphs present the results emerging from the observational and experimental analysis.

### 4.1 | Observational analysis

Data were analyzed through the SPSS 27 Statistical Package. Data normality was checked through a manual Q–Q plots while the authors also accounted for skewness and kurtosis for each item.

Provided that some of the items were translated in Italian for the first time for the sake of the present research, an EFA with *promax* rotation was performed, which confirmed the considered items loaded on five different factors corresponding to the variables in the model (see [Appendix](#) for further details). Items loading on more than one factor as well as those with correlation coefficient below 0.5 were removed. Furthermore, scale reliability and variance-extracted measures showed encouraging results. Concerning internal validity, McDonald's omega (Dunn et al., 2014) was satisfactory for all measures, namely *procedural justice* (0.814), *distributive justice* (0.836), *interactional justice* (0.932), *satisfaction with the system* (0.895), and *work engagement* (0.928). Eventually, both construct reliability and convergent validity were assessed; respectively, values ranged between 0.807 and 0.928 and between 0.500 and 0.727 (AVE).

Afterward, a confirmatory factor analysis (CFA) was performed including PA justice measures, satisfaction with the system and work engagement. In detail, two different models were tested and compared: the first model accounted for the first-order factors (i.e., *procedural*, *distributive*, and *interactional*), whereas the second one included the second-order factor of PA justice, to account for the potential correlation between the three first-order factors related to PA justice. The correlation between the three latent variables was very high, ranging from 0.667 (correlation between distributive and interactional justice) to 0.741 (correlation between procedural and distributive justice). Thus, the pattern of correlations suggested to account for a second-order factor explaining these correlations. Concerning the first-order model, we observed an acceptable fit with a CFI index of 0.976, a GFI index of 0.956 (both acceptable over 0.9), RMSEA equal to 0.045 (good under 0.08), and a

**TABLE 4** Results of the CFA.

	Standardized coeff.	S.E.	C.R.	p-Value
Relational justice → PA justice	0.882 <sup>a</sup>			
Distributive Justice → PA justice	0.824	0.037	25.236	<0.001
Procedural justice → PA justice	0.883	0.037	19.837	<0.001
<i>Relational justice first-order factor</i>				
Relational item 2 → Relational justice	0.912 <sup>a</sup>			
Relational item 3 → Relational justice	0.897	0.028	33.647	<0.001
Relational item 4 → Relational justice	0.912	0.027	36.300	<0.001
Relational item 5 → Relational justice	0.890	0.029	33.369	<0.001
Relational item 6 → Relational justice	0.769	0.028	30.206	<0.001
<i>Distributive justice first-order factor</i>				
Distributive item 1 → Distributive justice	0.842 <sup>a</sup>			
Distributive item 2 → Distributive justice	0.731	0.030	28.844	<0.001
Distributive item 3 → Distributive justice	0.828	0.035	27.042	<0.001
Distributive item 4 → Distributive justice	0.685	0.032	26.585	<0.001
<i>Procedural justice first-order factor</i>				
Procedural item 1 → Procedural justice	0.597 <sup>a</sup>			
Procedural item 2 → Procedural justice	0.737	0.052	24.812	<0.001
Procedural item 4 → Procedural justice	0.666	0.060	19.776	<0.001
Procedural item 7 → Procedural justice	0.811	0.065	22.201	<0.001
<i>Satisfaction with the PA system first-order factor</i>				
Satisfaction item 1 → Satisfaction with the system	0.791 <sup>a</sup>			
Satisfaction item 2 → Satisfaction with the system	0.896	0.027	38.796	<0.001
Satisfaction item 3 → Satisfaction with the system	0.894	0.029	38.676	<0.001
<i>Engagement first-order factor</i>				
Engagement item 2 → Engagement	0.768 <sup>a</sup>			
Engagement item 3 → Engagement	0.935	0.036	39.692	<0.001
Engagement item 4 → Engagement	0.896	0.038	37.915	<0.001
Engagement item 5 → Engagement	0.789	0.035	35.612	<0.001
Engagement item 7 → Engagement	0.824	0.038	34.198	<0.001

Abbreviations: CFA, confirmatory factor analysis; PA, performance appraisal.

<sup>a</sup>Items constrained for identification purposes.

CMIN/DF of 3.985, ( $\chi^2(176) = 701.318$ ). Second, we tested whether the imposition of a second-order model results in a significant decrease in fit, compared to the first-order model. The second-order model fit indices were equal to 0.976 (CFI), 0.956 (GFI), 0.045 (RMSEA), and a CMIN/DF of 3.938, ( $\chi^2(180) = 708.968$ ). Provided that the first-order model is nested within the second one, a chi-squared difference test might be performed, in line with previous recommendations from the academic literature (Brown, 2006). The computed chi-squared difference value between the second-order and the first-order was equal to 7.65(4) and was not statistically significant at the 0.05 significance level ( $p = 0.105$ ). A non-statistically significant difference means that both the first- and the second-order model fits the data equally well; therefore, the second-order model, characterized by a lower number of parameters estimated, might be preferred for parsimony. Eventually, we let error covariance correlate, according to

modification index, and the CMIN/DF became equal to 2.794, the CFI to 0.986, and the RMSEA to 0.035. Results of the CFA are reported in Table 4.

Through CFA, we also tested for model invariance across organizations operating at different administrative levels, confirming both configural and metric invariance (see Appendix for further details).

Afterward, the SEM was run. We performed a mediation analysis, following the recommendations of Holmbeck (1997), testing the hypothesized model against possible rival models (Table 5). First, the direct relationships between the independent (PA justice) and the dependent variable (work engagement) was modeled (M1). The model describing the direct relationships between PA justice and work engagement fit the data quite poorly (CFI < 0.95; RMSEA > 0.05). Next, we examined whether satisfaction with the PA system mediated the direct relationships between PA justice and work engagement. Against this first model, two other models were then tested and compared: the full (M2) and the partial (M3) mediation models. In the full mediation model, the independent variable was only indirectly related to the dependent variable through the mediator. In the partial mediating model, an additional direct path was allowed between the independent and the dependent variables. According to Holmbeck (1997), evidence for full mediation is found if the full mediation model fits the data at least equally well as the partial mediational model. The full mediation model (M2) yielded a good fit to the data, significantly better than the baseline model (M1), with both CFI and TLI > 0.98, and RMSEA < 0.05. We then proceeded to examine whether adding a direct path between PA justice and work engagement would be associated with an increased model fit: the partial mediation model (M3) actually fitted the data significantly better than the full mediation model, with  $\Delta\chi^2$  (df) significant at  $p < 0.05$ . In short, model testing indicated that satisfaction with the PA system acted as a partial mediator in the relationship between PA justice and work engagement.

**TABLE 5** Comparison of structural equation models.

		$\chi^2$	df	$\chi^2/\text{df}$	Delta $\chi^2$	CFI	RMSEA	TLI
Model 1	PA justice → Work engagement	1737.601	178	9.762	-	0.931	0.077	0.919
Model 2	PA justice → Satisfaction with the system → Work engagement	507.188	177	2.865	-1230.41	0.985	0.035	0.983
Model 3	PA justice → Satisfaction with the system → Work engagement; PA justice → Work engagement	491.69	176	2.794	-15.498	0.986	0.035	0.983

Abbreviation: PA, performance appraisal.

**TABLE 6** SEM direct, indirect, and total effects.

	Coeff.	S.E.	z	$p > z$	[95% C.I.]
<i>Direct effects</i>					
PA justice → Engagement	0.283	0.086	3.291	<0.01	[0.111 0.443]
<i>Indirect effects</i>					
PA justice → Satisfaction with the system → Engagement	0.107	0.059	1.796	=0.072	[0.033 0.237]
<i>Total effects</i>					
PA justice → Engagement	0.390	0.029	13.448	<0.01	[0.328 0.441]

Abbreviations: PA, performance appraisal; SEM, structural equation modeling.

**TABLE 7** Results of the SEM.

	Standardized coeff.	S.E.	CR	p-Value
PA justice → PA satisfaction with the system	0.861	0.041	25.484	<0.001
PA justice → Work engagement	0.283	0.074	4.006	<0.001
PA Satisfaction → Engagement	0.124	0.058	1.803	0.070
<i>Control variables</i>				
Gender → Engagement	0.075	0.053	3.021	0.003

Abbreviations: PA, performance appraisal; SEM, structural equation modeling.

In the structural model, we estimated all the hypothesized relationships, simultaneously controlling path coefficients for respondents' gender.

The second-order model of justice shows a positive and statistically significant direct effect on work engagement ( $\beta = 0.283$ ;  $p < 0.01$ ). Moreover, also the satisfaction with the system has a positive and statistically significant effect on work engagement ( $\beta = 0.124$ ;  $p < 0.10$ ). Furthermore, when only the direct relationship between PA justice and work engagement is accounted, the coefficient is equal to 0.389 ( $p < 0.01$ ). Nevertheless, once the mediator—that is, satisfaction with the system—is considered in the model, the direct path between the independent and the dependent variable still remains statistically significant at the 0.01, although the magnitude of the effect slightly decreases. Table 6 reports data on the total, direct, and indirect effect.

Following previous recommendations (MacKinnon et al., 2004), bootstrapping was applied to account for bias-corrected confidence intervals for the indirect effect, with 1,000 bootstrap samples and 95% percentile confidence interval. According to Hayes (2009), bootstrapping represents one of the most powerful method useful to test hypotheses about intervening variables since it treats the sample as a representation of the population that is repeatedly resampled during analysis as a “mean of mimicking the original sample process” (Hayes, 2009, p. 412). When it comes to control variables, gender has a positive and statistically significant association with engagement ( $\beta = 0.075$ ;  $p < 0.01$ ), as women seem to show a higher level of engagement compared to their male colleagues. Table 7 reports the full results for the SEM.

Moreover, we considered correlations between the variables included in the model (see Appendix for further details): as we observed a high correlation between the second-order factor independent variable and the mediator, with this possibly being indicative of redundant concepts, we performed several tests aimed at checking for the discriminant validity of the two variables. Following the procedures recommended in the literature (Cheung et al., 2023; Rönkkö & Cho, 2022) and the corresponding criteria and thresholds, we confirmed discriminant validity by (i) checking unidimensionality, (ii) testing that correlation between the two constructs were significantly less than unit, (iii) comparing AVE with MSV, and (iv) applying the heterotrait-monotrait approach.

Eventually, to control for heterogeneous effects, in the partial mediation model we checked whether the effect of the dependent variable and the mediator vary among different groups within the sample, following academic recommendations on the topic (Bullock et al., 2010). Specifically, we tested for differences in terms of gender, and age. No relevant differences both in terms of coefficient magnitude and statistically significant might be underlined across different groups, although some groups were not compared due to large differences in sample size (e.g., under 39 vs. other two groups). For the same reason, that is, large differences in sample size, no comparison for the organizational role was made.

## 4.2 | Experimental analysis

The results of the SEM show that procedural justice represents the most relevant variable in predicting the overall level of PA justice. Therefore, in selecting the three attributes to include in our experimental design, we considered

**TABLE 8** Results of the DCE.

Variables	Coeff.	S.E.	z	p > z	[95% C.I.]	Odd ratios	% Change in odds
Rating source (multisource)	.058	.0517	-1.14	0.255	[-0.160 0.042]	0.942	-6
Calibration (presence)	.353	.0520	6.79	0.000	[0.251 0.455]	1.423	42
Rater–ratee feedback (presence)	.598	.0520	11.51	0.000	[0.496 0.700]	1.819	81
Observations	3344						
$\chi^2$	192.05						
Prob > $\chi^2$	0.000						

Abbreviation: DCE, discrete choice experiment.

those variables that (i) the academic literature on PA system commonly relates to both structural and procedural proximal variables and (ii) mainly relates to the process of PA, since it is the variable which seems to affect the most overall perceptions of PA justice. In this vein, the DCE has accounted for the following features: rating source, feedback sessions, and rater accountability mechanisms.

Hence, we use conditional logit in order to analyze the effects of *rater–ratee feedback sessions* (vs. their absence), the presence of a *multisource rating source* (vs. its absence), and the presence of calibration mechanisms, that is, *upward rater accountability*, (vs. their absence). Table 8 reports the coefficient of the regression for each attribute, standard errors (SE), standardized coefficients (z), *p*-values (*p* > *z*); odd ratios ( $e^b$ ), and percentage change in odds when the attribute varies from one level to the other.

The results suggest that, when considering structural characteristics of the PA system, the presence of *rater–ratee feedback sessions* is the key variable. The presence of two or more meetings between the rater and the ratee increase the odds of perceiving the PA system as fairer by 81 percentage points, compared to the alternative (i.e., no feedback sessions). Also the presence of mechanisms of *rater accountability* has a positive and statistically significant effect on employee perceptions: when the system provides opportunities for the raters to gather and compare the performance evaluations of their employees, PA justice perceptions increase by 42 percentage points compared to the alternative scenario (i.e., no calibration). On the other hand, the presence of *multiple rating sources* (i.e., multisource feedback) does not seem to affect the odds of the respondents in perceiving the PA system as fair.

## 5 | DISCUSSION

The analysis confirms many of the formulated hypotheses on the considered factors supporting work engagement in the public sector, corroborating the central role of PA when it comes to individual employee outcomes. In doing so, it further extends previous contributions on this topic, while providing new insights to the academic literature.

First, the study further complements previous academic evidence (Meng & Wu, 2015), suggesting the importance of the three aspects associated to PA justice. In detail, results confirm that all three constructs are related to overall perceptions of PA justice although the highest coefficient might be underlined for procedural justice. This result might suggest that, at least in the Italian administrative and cultural contexts, having unbiased and accurate procedures and standards for individual evaluation represents the key element that drive employees' perceptions about what is fair in the organization. Second, the present contribution confirms the role of justice in fostering a positive cognitive–affective state as work engagement.



On the one hand, these results complement previous contributions suggesting the importance of perceiving fair treatment at work in fostering work engagement in public services (van Gelderen & Bik, 2016). In detail, relying on social exchange relationships (Saks, 2006), the present study is coherent with previous studies assessing the impact of both distributive and procedural justice (Krehbiel & Cropanzano, 2000) on discrete emotions such as happiness and pride (Weiss et al., 1999)—affective states positively correlated to work engagement—as well as with contributions focusing on the positive emotions associated with interactional justice (Foa & Foa, 1980). In this perspective, the contribution shed light on the link between PA justice and work engagement, confirming that fairness perceptions—for both distributive, procedural, and interactional justice—might be translated into a positive psychological state (Weiss et al., 1999).

On the other hand, in considering the mediating role of satisfaction with the PA system, the study provides a nuance to the academic debate on the cognitive versus affective foundation of SET (Cropanzano & Mitchell, 2005). In fact, while the positive impact of justice might be interpreted in the light of possible discrete emotions resulting from high level of fairness perceptions triggering the affective state of work engagement, a more cognitive state—satisfaction, related to what people believe about the PA system, seems to (i) have a positive and statistically significant effect on work engagement and (ii) account for a part of PA justice's relationship with the latter. In other words, perceiving the system as fair in the outcome, the procedures and in its relational aspects might increase work engagement both directly, fostering positive discrete emotions such as pride and enthusiasm, and indirectly, by crafting the employee's beliefs that the system is effective. Furthermore, our results highlight a high correlation between perceived PA justice and satisfaction with the PA system. An explanation of this result can be related to the peculiarities of the Italian public sector context. On the one hand, this might be linked to cultural features, and especially to a medium-high level of uncertainty avoidance that characterizes Italy (Hofstede, 1991) which might result, at the organizational level, in a strong preference toward structured and standardized practice, including the PA system. As a result, employees might perceive their PA system as fairer, primarily from a procedural standpoint, as long as it provides for well-structured and clear procedures. On the other hand, this can be interpreted also in the light of the administrative Italian public sector tradition (Sotiropoulos, 2004), characterized by a higher level of formalism: with regard to satisfaction with the PA system, Italian civil servants might value perceived fairness, rooted in clarity and neutrality of procedures, more than their counterparts in other countries characterized by a different administrative tradition.

Eventually, the study specifically provides further evidences to the puzzling context of the academic literature on fairness perceptions in the PA process (Gupta & Kumar, 2013; Meng & Wu, 2015; Vidè et al., 2022). Nevertheless, differently from previous contributions, it does not look at the differential effects of these constructs, assuming these aspects to be highly correlated—although arising from different managerial practices, and dealing in a broader sense with the “personal evaluation about the ethical and moral standard of a managerial conduct” (Cropanzano et al., 2007, p. 34).

Ultimately, in dealing with the research gap in managerial literature, this study explores the effects of *structural* (rating source) and *procedural* (rater–ratee feedback sessions and calibration mechanisms) proximal variables. This choice is informed by the results of the SEM: rater–ratee feedback sessions, calibration mechanisms, and rating source refers to the process of evaluation, thus being potentially related to procedural justice (which is the construct mostly correlated to PA justice as a whole). Moreover, the DCE confirmed the relevance of process proximal variables (especially rater–ratee feedback), whereas the rating source (*multisource feedback*) does not increase the odds of perceiving the PA system as fair. This result seems to suggest that, at least in the Italian administrative context, the structural characteristics of PA systems (i.e., rating source) do not drive the perceived justice of PA, differently from process proximal variables. However, multisource feedback is not widely used in the Italian public sector organizations included in the sample: thus, a possible explanation of the results could be that civil servants might not value multiple rating sources as a relatively stronger driver for PA justice compared to other factors, based on their actual limited experience of it.

## 6 | CONCLUSION

Through the combination of both observational and experimental analyses, the present study aims to provide novel insights into the relationship between PA justice and work engagement, and to shed new light on how to effectively design PA systems (in terms of both structural and procedural proximal variables), with the aim to ensure better individual outcomes. By doing so, the study contributes to context-based HR management literature (Cooke, 2018), discussing public sector specificities in a still under-investigated administrative context, such as the Italian one, when it comes to PA effectiveness and work engagement. The study advances the academic scholarship as it follows.

First, the results of the study respond to the scholars' calls for further research on the effectiveness of PA appraisal (Gupta & Kumar, 2013), while trying to increase the understanding of public sector specificities. On the one hand, the extant research on the characteristics of PA systems and their effects on PA effectiveness mainly relates to private organizations, thus leaving public organizations quite unexplored. On the other, public sector scholarship has, over time, analyzed the relationship between PA justice and several relevant individual-level outcomes (Kim & Rubianty, 2011; Quratulain et al., 2019), but it has substantially neglected its association with engagement. However, the latter represents a key working variable in the public domain, due to the peculiar combination of inherent complexity factors and peculiar motivational drivers. Second, this contribution underlines the need for further research on the impact of PA effectiveness, as measured by employees' PA justice perceptions, within the public sector. In this vein, a fair-perceived PA system could be a valuable tool in fostering employee positive work-related attitudes, ultimately resulting in increased productivity, effectiveness, and efficiency in public service. Therefore, the present study contributes to the debate on PA role in the public sector, which has been characterized by mixed and conflicting results, suggesting the need for a broader investigation of different administrative and cultural contexts.

Third, the study addresses gaps in managerial literature, bringing together employee perceptions of the PA system and its structural characteristics, in line with a recent call for further research on this topic (Barbieri et al., 2021). Furthermore, in applying the lens of SET to the link between justice and work engagement, it contributes to that strand of the literature framing social exchanges as not only cognizing and thinking, but also emoting and feeling (Lawler, 2001), focusing on a still under-investigated sector (i.e., the public sector) and relationship (PA justice and work engagement).

When it comes to the implications for the community of practice, the research suggests that public managers should account for and monitor employee perception of PA justice, so as to ensure a higher level of work engagement and, in the end, sustain higher performance. Coming to a better understanding of how the PA system should be structured seems to be a relevant line of intervention, in increasing PA justice. Public management might rely on leadership programs and training, with an eye to how to give and receive feedback effectively. Furthermore, public organizations might foster a participatory approach toward PA among senior and middle managers through calibration mechanisms, in order to control for the level of subjectivity characterizing performance evaluation.

All this provided, the study presents several limitations. From a methodological standpoint, the study is potentially exposed to the usual risks associated with observational designs, with these risks being of particular concern when performing mediation analyses. As pinpointed by previous academic contributions on the topic (Bullock et al., 2010), these methods might in fact be flawed by biased estimated, due to the presence of omitted variables that might affect the relationship between the mediator and the dependent variable, thus suggesting the need of employing experimental methods of mediation analysis. In this vein, further studies could explore this mediation paths relying on an experimental design. Moreover, as the SEM analysis is based on cross-sectional data, the risk of common method bias cannot be fully excluded. In addition, the study employs a measure of PA effectiveness (i.e., satisfaction with the PA system) based on employees' perceptions. Further studies could analyze the impact of both characteristics of the system and PA justice perceptions on other relevant measures such as PA usefulness and PA accuracy.

Our data also suggest a strong correlation between PA justice and satisfaction with PA system: this might raise concerns about conceptual redundancy. However, the results of the discriminant validity tests were reassuring, even

if in some cases borderline with the thresholds recommended in the literature. Still, particular attention should be paid to issues of construct validity when performing mediation analyses.

When it comes to the DCE, on the base of previous academic contributions and the results of the observational analysis, the authors rely on structural and process proximal variables associated with procedural justice. However, other studies might include more investigated features such as PA purpose, dealing with the effect on both perception of PA justice and individual outcome measures. Furthermore, from a methodological standpoint, the adoption of artificial scenarios detracts from external validity: extreme caution is indeed required when interpreting the results of the study, as there is no evidence that they may extend beyond our DCE to real-life settings. Nonetheless, DCEs have proven effective in predicting how individuals behave in reality and are hence more and more employed in public administration research (Bellé & Cantarelli, 2018a). Moreover, a limited number of respondents completed the experiment, compared to the overall sample that took part to the survey: this aspect might be linked to higher cognitive load associated with the DCE design.

Besides, experiments might be exposed to some other potential flaws, such as the difficulty to manipulate one variable at the time and the possibility that the results apply only to those people affected by the intervention, not to the entire sample.

Eventually, from a theoretical point of view, the present contribution relies on SET to explain the relationship between PA justice and work engagement. Nevertheless, this link might be explained also according to other theoretical frameworks. As an example, justice might represent a job resource, which might have an impact on the employee psychological state and his/her attitude toward work. Therefore, future studies might rely on the JDR model to explain this relationship, as well as they might integrate this more heuristic theoretical framework with other theories, such as self-determination theory (Van Den Broeck et al., 2008). As a result, future analyses might explicitly account for the interplay of need satisfaction in the relationship between PA justice and work engagement.

## ACKNOWLEDGMENTS

The study was conducted as part of a research project supported by the Department of Public Administration (*Dipartimento della Funzione Pubblica*) of the Italian Presidency of the Council of Ministers. The authors wish to thank Prof. Giovanni Valotti, research supervisor of the project from which the study originated, for his helpful suggestions, and the anonymous reviewers for their valuable comments.

## CONFLICT OF INTEREST

The authors declare that they have no involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

## DATA AVAILABILITY STATEMENT

The data underlying the findings of this study have been collected in a research project supported by the Department of Public Administration of the Presidency of the Council of Ministers of Italy. Restrictions apply to the availability of such data, which were used under license for this study.

## ORCID

Lorenza Micacchi  <https://orcid.org/0000-0002-7075-222X>

Francesco Vidé  <https://orcid.org/0000-0001-9416-1080>

Giorgio Giacomelli  <https://orcid.org/0000-0002-7798-0508>

Marta Barbieri  <https://orcid.org/0000-0002-2551-285X>

## ENDNOTES

<sup>1</sup> Within the academic literature, scholars have addressed this construct referring alternatively to (PA) justice and PA fairness perceptions. For the sake of the present study, the authors refer to the PA justice nomenclature.

<sup>2</sup> The sample has been defined according to a stratified random sampling method accounting for the institutional differences among Italian public organizations and relying on the policy indications provided by the Italian Department of Public Administration at the Presidency of the Council of Ministers. The 11 organizations include: 3 municipalities, 1 organization operating at the supra-municipality level (so-called “Città Metropolitana”), 2 regions, 4 agencies operating at the central level and 1 park authority.

## REFERENCES

- Anitha, J. (2014) Determinants of employee engagement and their impact on employee performance. *International Journal of Productivity and Performance Management*, 63(3), 308–323.
- Baimyrzaeva, M. & Meyer, C.T. (2021) Megatrends and the future of work for public sector leaders and educators. *International Journal of Public Leadership*, 17(2), 178–195. Available from: <https://doi.org/10.1108/ijpl-07-2020-0064>
- Balducci, C., Fraccaroli, F. & Schaufeli, W.B. (2010) Psychometric properties of the Italian version of the Utrecht Work Engagement Scale (UWES-9): a cross-cultural analysis. *European Journal of Psychological Assessment*, 26(2), 143.
- Barbieri, M., Micacchi, L., Vidè, F. & Valotti, G. (2021) The performance of performance appraisal systems: a theoretical framework for public organizations. *Review of Public Personnel Administration*, 43(1), 104–129. Available from: <https://doi.org/10.1177/0734371X211043560>
- Battaglio, R.P., Bellé, N. & Cantarelli, P. (2021) Self-determination theory goes public: experimental evidence on the causal relationship between psychological needs and job satisfaction. *Public Management Review*, 24(9), 1411–1428. Available from: <https://doi.org/10.1080/14719037.2021.1900351>
- Bellé, N. & Cantarelli, P. (2018a) Randomized experiments and reality of public and nonprofit organizations: understanding and bridging the gap. *Review of Public Personnel Administration*, 38(4), 494–511.
- Bellé, N. & Cantarelli, P. (2018b) The role of motivation and leadership in public employees' job preferences: evidence from two discrete choice experiments. *International Public Management Journal*, 21(2), 191–212.
- Bhatnagar, J. & Biswas, S. (2010) Predictors & outcomes of employee engagement: implications for the resource-based view perspective. *Indian Journal of Industrial Relations*, 46(2), 273–286.
- Biswas, S., Varma, A. & Ramaswami, A. (2013) Linking distributive and procedural justice to employee engagement through social exchange: a field study in India. *International Journal of Human Resource Management*, 24(8), 1570–1587. Available from: <https://doi.org/10.1080/09585192.2012.725072>
- Blau, P.M. (1964) *Exchange and power in social life*. New York, NY: Wiley. Available from: <https://doi.org/10.4324/9780203792643>
- Borst, R.T., Kruyen, P.M. & Lako, C.J. (2019) Exploring the job demands–resources model of work engagement in government: bringing in a psychological perspective. *Review of Public Personnel Administration*, 39(3), 372–397.
- Borst, R.T., Kruyen, P.M., Lako, C.J. & Vries, M.S. (2020) The attitudinal, behavioral, and performance outcomes of work engagement: a comparative meta-analysis across the public, semipublic, and private sector. *Review of Public Personnel Administration*, 40(4), 613–640. Available from: <https://doi.org/10.1177/0734371X19840399>
- Brown, T.A. (2006) *Confirmatory factor analysis for applied research*. New York, NY: Guilford Publications.
- Bullock, J.G., Green, D.P. & Ha, S.E. (2010) Yes, but what's the mechanism? (don't expect an easy answer). *Journal of Personality and Social Psychology*, 98(4), 550–558. Available from: <https://doi.org/10.1037/a0018933>
- Cepiku, D., Mussari, R. & Giordano, F. (2016) Local governments managing austerity: approaches, determinants and impact. *Public Administration*, 94(1), 223–243. Available from: <https://doi.org/10.1111/padm.12222>
- Chen, H. & Jin, Y.H. (2014) The effects of organizational justice on organizational citizenship behavior in the Chinese context: the mediating effects of social exchange relationship. *Public Personnel Management*, 43(3), 301–313. Available from: <https://doi.org/10.1177/0091026014533897>
- Chernyak-Hai, L. & Rabenu, E. (2018) The new era workplace relationships: is social exchange theory still relevant? *Industrial and Organizational Psychology*, 11(3), 456–481.
- Cheung, G.W., Cooper-Thomas, H.D., Lau, R.S. & Wang, L.C. (2023) Reporting reliability, convergent and discriminant validity with structural equation modeling: a review and best-practice recommendations. *Asia Pacific Journal of Management*, Epub ahead of print. 1–39.
- Cho, Y.J. & Sai, N. (2013) Does organizational justice matter in the federal workplace? *Review of Public Personnel Administration*, 33(3), 227–251. Available from: <https://doi.org/10.1177/0734371X12458126>
- Choi, S. (2011) Organizational justice and employee work attitudes: the federal case. *American Review of Public Administration*, 41(2), 185–204. Available from: <https://doi.org/10.1177/0275074010373275>
- Christensen, R.K. & Gazley, B. (2008) Capacity for public administration: analysis of meaning and measurement. *Public Administration and Development*, 28(4), 265–279.
- Cohen-Charash, Y. & Spector, P.E. (2001) The role of justice in organizations: a meta-analysis. *Organizational Behavior and Human Decision Processes*, 86(2), 278–321.

- Colquitt, J.A., Scott, B.A., Rodell, J.B., Long, D.M., Zapata, C.P., Conlon, D.E. et al. (2013) Justice at the millennium, a decade later: a meta-analytic test of social exchange and affect-based perspectives. *Journal of Applied Psychology*, 98(2), 199.
- Colquitt, J.A., Wesson, M.J., Porter, C.O.L.H., Conlon, D.E. & Ng, K.Y. (2001) Justice at the millennium: a meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, 86(3), 425–445. Available from: <https://doi.org/10.1037/0021-9010.86.3.425>
- Cooke, F.L. (2018) Concepts, contexts, and mindsets: putting human resource management research in perspectives. *Human Resource Management Journal*, 28(1), 1–13. Available from: <https://doi.org/10.1111/1748-8583.12163>
- Cropanzano, R., Anthony, E.L., Daniels, S.R. & Hall, A.V. (2017) Social exchange theory: a critical review with theoretical remedies. *Academy of Management Annals*, 11(1), 479–516.
- Cropanzano, R., Bowen, D.E. & Gilliland, S.W. (2007) The management of organizational justice. *Academy of Management Perspectives*, 21(4), 34–48. Available from: <https://doi.org/10.5465/AMP.2007.27895338>
- Cropanzano, R. & Mitchell, M.S. (2005) Social exchange theory: an interdisciplinary review. *Journal of Management*, 31(6), 874–900. Available from: <https://doi.org/10.1177/0149206305279602>
- Cropanzano, R., Weiss, H.M., Suckow, K.J. & Grandey, A.A. (2000) Doing justice to workplace emotion. In: *Emotions in the workplace. Research, theory and practice*. London, UK: Quorum Books.
- Curtis, A.B., Harvey, R.D. & Ravden, D. (2005) Sources of political distortions in performance appraisals: appraisal purpose and rater accountability. *Group and Organization Management*, 30(1), 42–60. Available from: <https://doi.org/10.1177/1059601104267666>
- de Schrijver, A., Delbeke, K., Maesschalck, J. & Pleysier, S. (2010) Fairness perceptions and organizational misbehavior: an empirical study. *American Review of Public Administration*, 40(6), 691–703. Available from: <https://doi.org/10.1177/0275074010363742>
- DeLeon, L. & Ewen, A.J. (1997) Multi-source performance appraisals: employee perceptions of fairness. *Review of Public Personnel Administration*, 17(1), 22–36.
- Demerouti, E., Bakker, A.B., Nachreiner, F. & Schaufeli, W.B. (2001) The job demands-resources model of burnout. *Journal of Applied Psychology*, 86(3), 499.
- Dunn, T.J., Baguley, T. & Brunson, V. (2014) From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399–412.
- Dusterhoff, C., Cunningham, J.B. & MacGregor, J.N. (2014) The effects of performance rating, leader-member exchange, perceived utility, and organizational justice on performance appraisal satisfaction: applying a moral judgment perspective. *Journal of Business Ethics*, 119(2), 265–273. Available from: <https://doi.org/10.1007/s10551-013-1634-1>
- Eremenco, S.L., Cella, D. & Arnold, B.J. (2005) A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Evaluation & the Health Professions*, 28(2), 212–232.
- Fletcher, L., Bailey, C., Alfes, K. & Madden, A. (2020) Mind the context gap: a critical review of engagement within the public sector and an agenda for future research. *The International Journal of Human Resource Management*, 31(1), 6–46.
- Fletcher, L., Bailey, K., Alfes, K. & Madden, A. (2016) Employee engagement in the public sector: a narrative evidence synthesis. *Academy of Management Proceedings*, 1, 13106. Available from: <https://doi.org/10.5465/ambpp.2016.13106abstract>
- Foa, E.B. & Foa, U.G. (1980) Resource theory. In: *Social exchange*. Boston, MA: Springer, pp. 77–94.
- Frink, D.D. & Ferris, G.R. (1998) Accountability, impression management, and goal setting in the performance evaluation process. *Human Relations*, 51(10), 1259–1283. Available from: <https://doi.org/10.1177/001872679805101003>
- Gosselin, A., Werner, J.M. & Hallé, N. (1997) Ratee preferences concerning performance management and appraisal. *Human Resource Development Quarterly*, 8(4), 315–333.
- Greenberg, J. (1986) Determinants of Perceived Fairness of Performance Evaluations. *Journal of Applied Psychology*, 71(2), 340–342. Available from: <https://doi.org/10.1037/0021-9010.71.2.340>
- Gruman, J.A. & Saks, A.M. (2011) Performance management and employee engagement. *Human Resource Management Review*, 21(2), 123–136.
- Gupta, V. & Kumar, S. (2013) Impact of performance appraisal justice on employee engagement: a study of Indian professionals. *Employee Relations*, 35(1), 61–78. Available from: <https://doi.org/10.1108/01425451311279410>
- Hajnal, G. & Staronova, K. (2021) Changing patterns of individual performance appraisal systems for civil service in European Union countries: toward a developmental or an incentivizing model. *International Journal of Public Sector Management*, 34(7), 748–764.
- Hakanen, J.J., Schaufeli, W.B. & Ahola, K. (2008) The job demands-resources model: a three-year cross-lagged study of burnout, depression, commitment, and work engagement. *Work and Stress*, 22(3), 224–241. Available from: <https://doi.org/10.1080/02678370802379432>
- Harter, J.K., Schmidt, F.L. & Hayes, T.L. (2002) Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: a meta-analysis. *Journal of Applied Psychology*, 87(2), 268. Available from: <https://doi.org/10.1037/0021-9010.87.2.268>

- Hayes, A.F. (2009) Beyond Baron and Kenny: statistical mediation analysis in the new millennium. *Communication Monographs*, 76(4), 408–420. Available from: <https://doi.org/10.1080/03637750903310360>
- He, H., Zhu, W. & Zheng, X. (2014) Procedural justice and employee engagement: roles of organizational identification and moral identity centrality. *Journal of Business Ethics*, 122(4), 681–695. Available from: <https://doi.org/10.1007/s10551-013-1774-3>
- Hofstede, G. (1991) *Cultures and organizations: software of the mind*. London, UK: McGraw-Hill.
- Holmbeck, G.N. (1997) Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology*, 65(4), 599.
- Jang, J., Lee, D.W. & Kwon, G.H. (2021) An analysis of the influence of organizational justice on organizational commitment. *International Journal of Public Administration*, 44(2), 146–154. Available from: <https://doi.org/10.1080/01900692.2019.1672185>
- Kahn, W.A. (1990) Psychological conditions of personal engagement and disengagement at work. *The Academy of Management Journal*, 33(4), 692–724.
- Keeping, L.M. & Levy, P.E. (2000) Performance appraisal reactions: measurement, modeling, and method bias. *Journal of Applied Psychology*, 85(5), 708–723. Available from: <https://doi.org/10.1037/0021-9010.85.5.708>
- Kim, S.E. & Rubianty, D. (2011) Perceived fairness of performance appraisals in the federal government: does it matter? *Review of Public Personnel Administration*, 31(4), 329–348. Available from: <https://doi.org/10.1177/0734371X11428903>
- 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.
- Klassen, R.D. & Jacobs, J. (2001) Experimental comparison of web, electronic and mail survey technologies in operations management. *Journal of Operations Management*, 19(6), 713–728.
- Ko, J. & Hur, S. (2014) The impacts of employee benefits, procedural justice, and managerial trustworthiness on work attitudes: integrated understanding based on social exchange theory. *Public Administration Review*, 74(2), 176–187. Available from: <https://doi.org/10.1111/puar.12160>
- Krehbiel, P.J. & Cropanzano, R. (2000) Procedural justice, outcome favorability and emotion. *Social Justice Research*, 13(4), 339–360. Available from: <https://doi.org/10.1023/A:1007670909889>
- Kuvaas, B. (2006) Performance appraisal satisfaction and employee outcomes: mediating and moderating roles of work motivation. *The International Journal of Human Resource Management*, 17(3), 504–522.
- Kuvaas, B., Buch, R. & Dysvik, A. (2017) Constructive supervisor feedback is not sufficient: immediacy and frequency is essential. *Human Resource Management*, 56(3), 519–531.
- Landy, F.J., Janet, L. & Murphy, K.R. (1978) Correlates of perceived fairness and accuracy of performance evaluation. *Journal of Applied Psychology*, 63(6), 751–754.
- Lawler, E.J. (2001) An affect theory of social exchange. *American Journal of Sociology*, 107(2), 321–352.
- Levitats, Z. & Vigoda-Gadot, E. (2020) Emotionally engaged civil servants: toward a multilevel theory and multisource analysis in public administration. *Review of Public Personnel Administration*, 40(3), 426–446.
- Levitats, Z., Vigoda-Gadot, E. & Vashdi, D.R. (2019) Engage them through emotions: exploring the role of emotional intelligence in public sector engagement. *Public Administration Review*, 79(6), 841–852.
- Levy, P.E. & Williams, J.R. (2004) The social context of performance appraisal: a review and framework for the future. *Journal of Management*, 30(6), 881–905. Available from: <https://doi.org/10.1016/j.jm.2004.06.005>
- Longenecker, C.O. & Nykodym, N. (1996) Public sector performance appraisal effectiveness: a case study. *Public Personnel Management*, 25(2), 151–164. Available from: <https://doi.org/10.1177/009102609602500203>
- Lovrich, N.P., Shaffer, P.L., Hopkins, R.H. & Yale, D.A. (1980) Do public servants welcome or fear merit evaluation of their performance? *Public Administration Review*, 40(3), 214–222.
- Lyons, B.J. & Scott, B.A. (2012) Integrating social exchange and affective explanations for the receipt of help and harm: a social network approach. *Organizational Behavior and Human Decision Processes*, 117(1), 66–79.
- MacKinnon, D.P., Lockwood, C.M. & Williams, J. (2004) Confidence limits for the indirect effect: distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39(1), 99–128. Available from: [https://doi.org/10.1207/s15327906mbr3901\\_4](https://doi.org/10.1207/s15327906mbr3901_4)
- Mani, B.G. (2001) Performance appraisal systems, productivity, and motivation: a case study. *Public Personnel Management*, 31(2), 141–159. Available from: <https://doi.org/10.1177/009102600203100202>
- Marsden, D. & Richardson, R. (1994) Performing for pay? The effects of 'merit pay' on motivation in a public service. *British Journal of Industrial Relations*, 32(2), 243–261.
- Maslach, C. & Leiter, M.P. (2008) Early predictors of job burnout and engagement. *Journal of Applied Psychology*, 93(3), 498–512.

- McFadden, D. (1973) Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.) *Frontiers in econometrics*. New York: Academic Press, pp. 105–142.
- Memon, M.A., Salleh, R., Mirza, M.Z., Cheah, J.-H., Ting, H. & Ahmad, M.S. (2019) Performance appraisal satisfaction and turnover intention: the mediating role of work engagement. *Management Decision*, 58(6), 1053–1066.
- Meng, F. & Wu, J. (2015) Merit pay fairness, leader-member exchange, and job engagement: evidence from Mainland China. *Review of Public Personnel Administration*, 35(1), 47–69. Available from: <https://doi.org/10.1177/0734371X12453057>
- Mero, N.P., Guidice, R.M. & Brownlee, A.L. (2007) Accountability in a performance appraisal context: the effect of audience and form of accounting on rater response and behavior. *Journal of Management*, 33(2), 223–252. Available from: <https://doi.org/10.1177/0149206306297633>
- Mount, M.K. (1984) Satisfaction with a performance appraisal system and appraisal discussion. *Journal of Occupational Behaviour*, 5(4), 271–279.
- Mulvaney, M.A. (2019) Examining the role of employee participation, supervisor trust, and appraisal reactions for a pay-for-performance appraisal system. *Public Organization Review*, 19(2), 201–225. Available from: <https://doi.org/10.1007/s11115-017-0400-8>
- Murphy, K.R. & Cleveland, J.N. (1995) *Understanding performance appraisal: social, organizational, and goal-based perspectives*. Thousand Oaks, CA: SAGE. Available from: <https://doi.org/10.5860/choice.32-5748>
- Nair, M.S. & Salleh, R. (2015) Linking performance appraisal justice, trust, and employee engagement: a conceptual framework. *Procedia-Social and Behavioral Sciences*, 211, 1155–1162.
- Quratulain, S., Khan, A.K. & Sabharwal, M. (2019) Procedural fairness, public service motives, and employee work outcomes: evidence from Pakistani public service organizations. *Review of Public Personnel Administration*, 39(2), 276–299. Available from: <https://doi.org/10.1177/0734371X17718029>
- Roberts, G.E. (2003) Employee performance appraisal system participation: a technique that works. *Public Personnel Management*, 32(1), 89–98.
- Roberts, G.E. & Reed, T. (1996) Performance appraisal participation, goal setting and feedback: the influence of supervisory style. *Review of Public Personnel Administration*, 16(4), 29–60.
- Rönkkö, M. & Cho, E. (2022) An updated guideline for assessing discriminant validity. *Organizational Research Methods*, 25(1), 6–14.
- Rubin, E.V. & Kellough, J.E. (2012) Does civil service reform affect behavior? Linking alternative personnel systems, perceptions of procedural justice, and complaints. *Journal of Public Administration Research and Theory*, 22(1), 121–141. Available from: <https://doi.org/10.1093/jopart/mur008>
- Ryan, M., Kolstad, J.R., Rockers, P.C. & Dolea, C. (2012) *How to conduct a discrete choice experiment for health workforce recruitment and retention in remote and rural areas: a user guide with case studies*. The World Bank.
- Saks, A.M. (2006) Antecedents and consequences of employee engagement. *Journal of Managerial Psychology*, 21(7), 600–619. Available from: <https://doi.org/10.1108/02683940610690169>
- Schaufeli, W.B., Bakker, A.B. & Rehen, W.V. (2009) How changes in job demands and resources predict burnout, work engagement, and sickness absenteeism. *Journal of Organizational Behavior*, 30, 839–862. Available from: <https://doi.org/10.1002/job>
- Schaufeli, W.B., Bakker, A.B. & Salanova, M. (2006) The measurement of work engagement with a short questionnaire: a cross-national study. *Educational and Psychological Measurement*, 66(4), 701–716.
- Schaufeli, W.B., Salanova, M., González-Romá, V. & Bakker, A.B. (2002) The measurement of engagement and burnout: a two sample confirmatory factor analytic approach. *Journal of Happiness Studies*, 3(1), 71–92.
- Shuck, B. (2011) Integrative literature review: four emerging perspectives of employee engagement: an integrative literature review. *Human Resource Development Review*, 10(3), 304–328.
- Sotiropoulos, D.A. (2004) Southern European public bureaucracies in comparative perspective. *West European Politics*, 27(3), 405–422.
- Street, D.J., Burgess, L. & Louviere, J.J. (2005) Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing*, 22(4), 459–470.
- Taormina, R.J. & Gao, J.H. (2009) Identifying acceptable performance appraisal criteria: an international perspective. *Asia Pacific Journal of Human Resources*, 47(1), 102–125. Available from: <https://doi.org/10.1177/1038411108099292>
- Thurston, P.W. & McNall, L. (2010) Justice perceptions of performance appraisal practices. *Journal of Managerial Psychology*, 25(3), 201–228. Available from: <https://doi.org/10.1108/02683941011023712>
- Tziner, A. & Kopelman, R.E. (2002) Is there a preferred performance rating format? A non-psychometric perspective. *Applied Psychology*, 51(3), 479–503. Available from: <https://doi.org/10.1111/1464-0597.00104>
- Van den Broeck, A., Vansteenkiste, M., De Witte, H. & Lens, W. (2008) Explaining the relationships between job characteristics, burnout, and engagement: the role of basic psychological need satisfaction. *Work & Stress*, 22(3), 277–294.
- Van Gelderen, B.R. & Bik, L.W. (2016) Affective organizational commitment, work engagement and service performance among police officers. *Policing*, 39(1), 206–221. Available from: <https://doi.org/10.1108/PIJPSM-10-2015-0123>

- Vid , F., Micacchi, L., Barbieri, M. & Valotti, G. (2022) The renaissance of performance appraisal: engaging public employees through perceived developmental purpose and justice. *Review of Public Personnel Administration*, Epub ahead of print. 0734371X221116584, 1–29.
- Vigoda-Gadot, E., Eldor, L. & Schohat, L.M. (2013) Engage them to public service: conceptualization and empirical examination of employee engagement in public administration. *The American Review of Public Administration*, 43(5), 518–538.
- Weiss, H.M., Suckow, K. & Cropanzano, R. (1999) Effects of justice conditions on discrete emotions. *Journal of Applied Psychology*, 84(5), 786–794.
- 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 Public Personnel Administration*, 39(2), 232–255.
- Williams, J.R. & Levy, P.E. (2000) Investigating some neglected criteria: the influence of organizational level and perceived system knowledge on appraisal reactions. *Journal of Business and Psychology*, 14(3), 501–513. Available from: <https://doi.org/10.1023/A:1022988402319>

**How to cite this article:** Micacchi, L., Vid , F., Giacomelli, G., & Barbieri, M. (2023). Performance appraisal justice and employees' work engagement in the public sector: Making the most of performance appraisal design. *Public Administration*, 1–26. <https://doi.org/10.1111/padm.12952>

## APPENDIX

### CORRELATION MATRIX

	Factor loading					
	1	2	3	4	5	6
Factor 1: PA justice (second-order factor)	1.00					
Factor 2: Engagement	0.388	1.00				
Factor 3: Satisfaction with the PA system	0.861	0.367	1.00			
Factor 4: Procedural justice (first-order factor)	0.883	0.343	0.76	1.00		
Factor 5: Distributive justice (first-order factor)	0.824	0.320	0.709	0.727	1.00	
Factor 6: Relational justice (first-order factor)	0.882	0.312	0.691	0.708	0.661	1.00



## RESULTS OF THE EXPLORATORY FACTOR ANALYSIS (EFA)

Items	Factor loading				
	1	2	3	4	5
<i>Factor 1: Relational PA justice</i>					
I trust my supervisor to fairly assess my performance and contributions.	Dropped				
I trust my supervisor clearly communicate conduct expectations	<b>0.591</b>	0.051	0.203	0.031	0.026
My supervisor keeps me informed about how well I am doing	<b>0.900</b>	0.005	-0.049	0.027	0.005
My supervisor provides constructive feedback on my job performance	<b>1.054</b>	-0.032	-0.023	-0.068	-0.036
My supervisor provides timely feedback on my job performance	<b>0.932</b>	-0.025	-0.005	-0.015	-0.034
My supervisor provides coaching, training opportunities, or other assistance to help me improve my skill and performance	<b>0.706</b>	0.040	-0.031	0.055	0.058
I understand how my supervisor will evaluate my performance	Dropped				
My supervisor rates my performance fairly and accurately	Dropped				
My supervisor is held accountable for rating employee performance fairly and accurately	Dropped				
<i>Factor 2: Work engagement</i>					
At my work, I feel bursting with energy	Dropped				
At my job, I feel strong and vigorous	-0.018	<b>0.772</b>	0.034	0.057	-0.031
When I get up in the morning, I feel like going to work	0.07	<b>0.930</b>	-0.04	0.034	-0.034
I am enthusiastic about my job	0.030	<b>0.895</b>	-0.055	-0.002	0.001
My job inspires me	0.008	<b>0.796</b>	0.032	-0.072	0.064
I am proud of the work that I do	Dropped				
I feel happy when I am working intensely	-0.031	<b>0.834</b>	0.008	-0.023	0.007
I am immersed in my work	Dropped				
I get carried away when I am working	Dropped				
<i>Factor 3: Procedural PA justice</i>					
I understand the basis for my most recent performance rating	-0.059	0.015	<b>0.797</b>	-0.099	0.001
The standards used to appraise my performance are appropriate	-0.024	-0.013	<b>0.765</b>	-0.043	0.105
I participate in setting standards and goals used to evaluate my job performance	Dropped				
I understand what I must do to receive a high performance rating	0.045	0.021	<b>0.633</b>	0.184	-0.085
I have sufficient opportunities (such as challenging assignments or projects) to earn a high performance rating	Dropped				
I know how my performance rating compares to others in my organization who have similar jobs	Dropped				
Objective measures are used to evaluate my performance	0.148	-0.014	<b>0.684</b>	0.056	0.058

(Continues)

Items	Factor loading				
	1	2	3	4	5
<i>Factor 4: Distributive PA justice</i>					
Recognition and rewards are based on performance	0.031	-0.041	0.106	<b>0.728</b>	-0.007
I understand how my pay relates to my job performance	-0.021	-0.001	0.228	<b>0.650</b>	-0.107
My organization takes steps to ensure that employees are appropriately paid and rewarded	0.096	0.021	-0.100	<b>0.612</b>	0.183
If I perform well, it is likely I will receive a cash award or pay increase	-0.041	0.010	-0.141	<b>0.855</b>	0.016
<i>Factor 5: Satisfaction with the performance appraisal system</i>					
The performance appraisal system is effective in assessing how an employee worked during the evaluation period	-0.054	0.021	0.042	-0.027	<b>0.823</b>
Overall, I think that the organization has an excellent performance appraisal system	0.045	-0.003	-0.040	0.076	<b>0.848</b>
The appraisal system provides a fair and unbiased measure of the level of my performance	0.060	-0.019	0.108	0.007	<b>0.767</b>
Overall, I am satisfied by my work	Dropped				

Note:  $N = 1483$ . The extraction method was principal axis factoring with an oblique rotation (Promax with Kaiser normalization). Factor loading above 0.50 are in bold. Items loading on more than one factor and items loading below 0.50 on each factor were removed.

#### DETAILS ABOUT MODEL INVARIANCE TESTING

First, we tested for configural invariance: the configural model reported a good fit to the data (CMIN/DF of 1.873; CFI of 0.980; GFI of 0.941; and RMSEA of 0.024). Concerning regression weights, some limited differences might be underlined across the groups when it comes to the coefficient powers while no differences about significance level might be highlighted. Then, we tested for metric invariance, specifying equality constraints on factor loadings. The metric invariance model resulted also in a good fit (CMIN/DF of 1.839; CFI of 0.979; GFI of 0.940; and RMSEA of 0.024). We used the chi-squared differences test to check whether the metric invariance model represented a significantly worse fit to the data than the configural invariance model. The analysis suggested a  $\Delta\chi^2$  of 40.455(32) not statistically significant ( $p = 0.145$ ), pinpointing a not statistically significant reduction in fit after constraining parameters thus highlighting metric invariance across groups.