

University of Tübingen
Working Papers in
Business and Economics

No. 136

The Effect of Appraisal Interviews and Target
Agreements on Employee Effort - New Evidence Using
Representative Data

by

Patrick Kampkötter, Patrick Maier

Faculty of Economics and Social Sciences
www.wiwi.uni-tuebingen.de



The Effect of Appraisal Interviews and Target Agreements on Employee Effort - New Evidence Using Representative Data*

Patrick Kampkötter[†]

Patrick Maier[‡]

This version: 13 June 2020

Abstract

Performance measurement and evaluation systems are among the most common management instruments. An integral element of this process is the use of targets, typically set in appraisal interviews and formalized via written target agreements. In this paper, we investigate the relationship between performance management and evaluation systems and individual effort, proxied by the commonly used concept of work engagement. Using four waves of a new representative, linked employer-employee data set, the Linked Personnel Panel (LPP), we apply fixed effects estimations to account for unobserved heterogeneity. Our results show positive and statistically significant relationships between the presence of a performance management and evaluation process and employee engagement on the individual level. We are further able to differentiate between appraisal

*We thank Markus Arnold, Mario Bossler, Michael Haylock, Jan Riepe, Breno Sampaio, Matthias Seckler, Giuseppe Trevisan, conference participants at the IAB/ZEW HRM Workshop "Assessing the Impact of Human Resource Management Practices" in Nuremberg, the Workshop on Labour Economics 2018 in Trier, and seminar participants in Recife and Tübingen for helpful comments and suggestions. Mariella Misch provided excellent research assistance. Patrick Kampkötter thanks the German Research Foundation (DFG) for financial support through priority program SPP 1764 (KA 4591/1-2 and SL 46/2-1). Part of this research was conducted while Patrick Maier visited Universidade Federal de Pernambuco in Recife, Brazil. He would like to thank the Department of Economics for their hospitality.

[†]University of Tübingen, Department of Managerial Accounting, Nauklerstr. 47, 72074 Tübingen, email: patrick.kampkoetter@uni-tuebingen.de

[‡]University of Tübingen, Department of Managerial Accounting, Nauklerstr. 47, 72074 Tübingen, email: patrick.maier@uni-tuebingen.de

interviews and written target agreements which allows us to show a positive effect of appraisal interviews and an additional positive effect of target agreements. In addition, we find first evidence that these direct relationships are partially mediated by goal clarity and procedural fairness.

Key Words: Target Agreements; Performance Appraisals; Work Engagement; Goal Clarity; Procedural Fairness

JEL Classification: D23, J01, J33, M41, M52

1 Introduction

Performance management systems constitute an important management practice in modern organizations (Otley (1999), Franco-Santos et al. (2012)). An essential ingredient of this framework is the performance management and evaluation process (PMEP), which comprises formal target setting between supervisors and subordinates and performance evaluation (Ferreira and Otley (2009)), typically operated via annual appraisal interviews (AI). These AIs can be conceptualized as a “conversation about performance” (Gordon and Stewart (2009), p. 473) and might function by “providing feedback to employees, counseling and developing employees, and conveying and discussing compensation, job status, or disciplinary decisions” (Cederblom (1982), p. 219). During these AIs, supervisors often make use of written target agreements (TA), where formalized targets are often set at the beginning of a fiscal year and reviewed in the subsequent year.

According to recent evidence from the representative, matched employer-employee data for private establishments from Germany used in this paper, the PMEP is of high practical importance for firms. In 2018, 85% of establishments reported using appraisal interviews while 80% reported employing written TAs. In addition, 64% of all employees working in these establishments reported being covered by an AI and 42% reported having both an AI and a written TA. Nevertheless, recently a lively debate about the outcomes and effectiveness of performance reviews and target setting has emerged in many organizations. For instance in 2015, Volkmar Denner, CEO of the German company Robert Bosch, publicly announced to abandon compensation plans based on the achievement of individual targets.¹ Companies such as Commerzbank, SAP, and Infineon have undergone similar changes.²

Target setting aims at increasing the organization’s productivity by aligning employee incentives with organizational goals. The direct relationship between target setting and productivity is thus likely to be mediated through employee decision-making and effort (Bender et al. (2018)). However, despite decades of research in management accounting, applied psychology, and organizational economics, prior literature has provided mixed evidence with respect to the impact of feedback and target setting characteristics on employee effort and employee performance³ (e.g. Podsakoff and Farh (1989), Locke and Latham (1990), Fisher

¹Source: <http://www.manager-magazin.de/unternehmen/karriere/mitarbeitermotivation-schaft-die-boni-ab-a-1055113.html>.

²See, for instance, <https://www.reuters.com/article/us-sap-appraisals/europes-biggest-software-maker-sap-ditches-annual-reviews-idUSKCN10N0RO> or <https://www.handelsblatt.com/today/finance/variable-compensation-commerzbank-eliminates-most-individual-bonuses-in-revamp/23583676.html>.

³In this study, we regard employee performance as a consequence of employee effort.

et al. (2000), Christ et al. (2012)) and on employee perceptions such as goal clarity and procedural fairness (e.g. Lau and Buckland (2001), Marginson and Ogden (2005), Derfuss (2009), Hartmann and Slapničar (2009), Vouřem et al. (2016)).⁴ Classical goal-setting theory (e.g. Locke and Latham (1990), Locke and Latham (2002)) expects specific, challenging targets⁵ to boost performance and the literature examining participation in target setting largely finds a positive impact on performance as well (e.g. Fisher et al. (2000); Wentzel (2002); Sholihin et al. (2011)). Other studies, however, argue that challenging targets do not boost performance of all individuals (Eyring and Narayanan (2018)), that target setting is associated with costs often ignored by prior literature (Holzhacker et al. (2019)), and that these negative side effects often outweigh its benefits (Barsky (2008), Ordóñez et al. (2009)).

It is important to note that prior studies using archival data are overwhelmingly based on cross sectional data sets, often from single firms, with only a small number of observations. This makes causal and generalizable interpretations difficult. We address this gap by using a large linked employer-employee data set, representative for all German establishments with more than 50 employees and their respective workforce. This allows us to move closer to causality than previous studies do by exploiting the longitudinal dimension of our data and using panel data methods. Thus, the first goal of this study is to provide more causal and generalizable evidence on the *average* effects of AIs and formalized TAs on employee effort. Second, we want to focus on effect heterogeneity by examining potential channels mediating the relationship between target setting and employee effort, in particular procedural fairness and goal clarity. In line with Lind and Tyler (1988), we define procedural fairness as referring to the perception of fairness with respect to the process via which outcomes are determined. The evidence with respect to the effect of (participation in) target setting on procedural fairness is again heterogeneous. Some studies (e.g. Sholihin et al. (2011)) find a positive impact of participation in budget setting and a sense of procedural fairness. Vouřem et al. (2016), for instance, argue that target characteristics such as subjectivity of performance measures are important, while Ordóñez et al. (2009) argue that targets might even lead to a feeling of unfairness. We consider goal clarity as our second potential mediator. In line with Sholihin and Pike (2013), we define goal clarity as a clearer understanding of organizational members about their goals. In a meta-analysis, Derfuss (2009) finds a positive relationship

⁴The literature uses different proxies for effort such as goal commitment (Sholihin et al. (2011)) or motivation (Locke and Latham (2002)). For the purpose of readability, we subsume these constructs under the term “effort”.

⁵Note that goal setting literature uses the term goal instead of target. For the sake of readability, we use these two terms interchangeably.

between participation in budget setting and goal clarity. In a recent lab experiment, Anderson and Stritch (2015) in turn find a positive relationship between goal clarity and performance. Thus, the second goal of our analysis is to examine whether the direct relationship between our PMEP and employee effort is mediated by procedural fairness and goal clarity.

Our analysis is based on the Linked Personnel Panel (LPP), a new matched employer-employee data set representative for all German private sector establishments with more than 50 employees and their respective workforce (Kampkötter et al. (2016)). Due to the panel structure of the data set, we are able to include various fixed effects in our model to account for omitted variable bias problems. As information on employee effort is difficult to collect particularly in representative data sets, we employ the concept of work engagement (Kahn (1990), Bakker (2017)), which is commonly applied in management and organizational psychology and has recently been introduced into the management accounting literature (Li and Sandino (2018)). This effort proxy has been empirically validated in various countries (Schaufeli and Bakker (2003)) and its implementation in a representative sample of the workforce allows us to shed some light on the nexus between individual effort and performance management practices. We are particularly interested in differentiating between the effects of a sole presence of an AI and the *additional* presence of a written TA in order to see if the effect of the PMEP on employee effort is driven entirely by the AI or if formalization via a written TA is of additional value for firms.

Our results show a positive and statistically significant effect of the presence of a PMEP on employee effort as proxied by work engagement. This effect is robust across various specifications. We further find first evidence that both AIs and TAs positively affect employee effort. While both the effects of AIs and TAs on work engagement are positive and statistically significant in all of our pooled OLS regressions, only the effect of AIs remain significant in our individual fixed effects specification. In this context, we explicitly discuss a problem commonly associated with the use of fixed effects in management practices research, namely that there is “not enough real time series variation (given measurement error) to identify any significant relationships” (Bender et al. (2018), p. 381). In our case, this means that only very few individuals switch from having no TA to having a TA without jointly switching from having no AI to having an AI (or vice versa). Hence, these HR measures in practice seem to be applied quite simultaneously. In our mediation analysis, we find that both procedural fairness and goal clarity *partially* mediate the direct effect of the PMEP on engagement.

We contribute to the literature in several ways. Our first and arguably most important contribution relates to the use of panel data methods to reduce the likelihood of endogeneity

problems such as omitted variable bias, which enables us to make more causal statements. As outlined above, most studies examining the effect of target setting on individual effort and the channels mediating this relationship rely on lab or field experiments (e.g. Liu and Zhang (2015), Li and Sandino (2018), Holzhaecker et al. (2019)), cross-sectional studies in single firms (e.g. Sholihin et al. (2011)), or small non-random samples of firms (e.g. Sholihin and Pike (2013)). While experiments are the best way to establish internal validity, they generally encounter problems related to external validity. Cross-sectional single firm case studies provide in-depth insights into the studied organization but are prone to issues regarding causality and generalizability. In the study most related to our paper, Sholihin et al. (2011) use data comprising 54 managers from a UK financial services institution and analyze both the direct effect of participative target setting on employee effort and the respective channels mediating this relationship. They find that procedural fairness and interpersonal trust fully mediate this direct relationship but call for future research using “larger samples from various organizations determined randomly” (p. 145) to further examine their propositions and findings. We are able to exploit four waves of a representative matched employer-employee data set comprising between 771 and 1,219 establishments per wave and between 6,500 and 7,500 employees randomly drawn from these establishments. Thus, we complement prior literature by being able to make more generalizable statements and to test whether the relationships found also hold for large, representative data.

Second, we add to the emerging literature on work engagement as a new proxy for employee effort at the workplace. This concept has recently been introduced into the management accounting literature by Li and Sandino (2018), who operationalize it via weekly salesperson attendance at a given store. As such sales information is usually not available in representative surveys spanning different industries, we employ a commonly used work engagement scale (Kahn (1990), Bakker (2017)). We argue that this scale measures effort in a more general way as concepts applied in previous studies using questionnaire data. Sholihin et al. (2011), for instance, analyze if participation in target setting affects goal commitment, which is defined as attachment to or determination to reach a goal (Locke and Latham (1990)) and willingness to put in effort to attain a goal (Renn et al. (1999)) and argue that this concept is related to employee effort and ultimately employee performance. While we agree with this statement, we think that goal commitment as a concept might focus too much on the attainment of the goal and might thus be too narrow to measure effort in a more general sense, as effort directed on goal areas might also crowd out effort directed at non-goal areas (Ordóñez et al. (2009)).

Third, we contribute to the literature examining the general effect of target setting on

employee effort and performance by showing a robust positive effect of our PMEP on employee effort. We argue that we also contribute to the literature examining participation in target setting. This is because, despite the fact that we do not know the exact contents of the AIs and TAs, this process is likely to be participative to some extent, as superior and subordinate meet in order to talk about the subordinate’s targets. There are numerous studies examining these links using lab experiments, single firm studies and literature reviews. In a literature review, Locke and Latham (2006) emphasize a positive, linear relationship between goal difficulty and task performance, as long as different goals are not conflicting, the respective person is committed to the goal and possesses the necessary ability to attain it. With regards to participation in target setting, Derfuss (2009) finds a positive relationship with employee behavior that is beneficial to the organization. However, the author also states that “many studies use small samples, and their conflicting findings might be due to statistical artifacts, such as sampling error” (p. 203). In a field experimental setting, Eyring and Narayanan (2018) provide evidence that challenging targets improve performance of above-median performers, but damp performers of below median performers. Finally, there is also evidence that the way that goals are set and communicated influences their effectiveness as well. Liu and Zhang (2015) find that performance is highest when the achievement of the target is revealed *ex post* (after the operation ends) rather than *ex ante* (before the operation starts) and when performance-contingent incentives are framed as a bonus rather than a penalty. Holzhacker et al. (2019) emphasize the potential costs of relative target setting, and analyze costs and benefits using data from an industrial services company. Thus, while some studies emphasize a positive impact of targets on effort and performance, others argue that the way that targets are set is important and that under certain conditions, the impact on effort and performance might even be negative. We thus use our representative data to provide evidence how targets set via two common performance management practices influence employee effort *on average*. For firms thinking about introducing (or abandoning) a PMEP, this may provide guidance about how this might influence employee effort.

Fourth, by differentiating between AIs and TAs, we also contribute to the literature analyzing whether a formalization of targets is of additional value and provide evidence that it is indeed. Locke and Latham (1990) e.g. state that specific, or more formalized, targets induce higher levels of effort and performance as do “do-your-best” ones. In line with Hartmann and Slapničar (2009), we hypothesize that targets might possess high formality in case they are explicated by superiors in a quantitative and written fashion. By discussing the performance of the last year and key areas of improvement, targets for the next period are set in a relatively

informal way during an AI and might then be formalized via a written TA. Therefore, we analyze if it is of additional value to move from a more informal way of setting targets to a more formal one. First, this might be important for firms that discuss implementing a PMEP and question whether an AI suffices or if a written TA adds additional value. Second, this information might also be interesting for firms that already use AIs and think about further formalizing the target setting process via a written TA.

Finally, we contribute to the debate about potential channels via which target setting affects effort and performance. In particular, we consider two potential mediators: procedural fairness and goal clarity. While most authors (e.g. Libby (1999, 2001); Wentzel (2002); Sholihin et al. (2011); Sholihin and Pike (2013)) find a positive link between (participation in) target setting and procedural fairness, Vouřem et al. (2016) examine whether procedural and distributive fairness perceptions are influenced by the degree of subjectivity of performance measures and find an inverted U-shape of this relationship. Ord3n3ez et al. (2009) argue that targets or goals might even lead to a feeling of unfairness. This is because employees are heterogeneous with respect to their level of ability. The same targets might thus be too easy for some individuals and too hard for others but when tailoring goals to individuals, some might feel treated unfairly in case they have the feeling that rewards do not fairly match effort and performance. With regards to the impact of procedural fairness on effort and performance, there is no controversy in the literature such that most authors (e.g. Libby (1999, 2001); Sholihin and Pike (2009); Zapata-Phelan et al. (2009)) find a positive effect. As outlined above, Derfuss (2009) finds a positive relationship between participation in budget setting and (among other dependent variables) goal clarity in a literature review, while Anderson and Stritch (2015) find a positive relationship between goal clarity and performance in an experimental study. Thus, we contribute to the literature by providing evidence on whether, on average, the positive effects of targets on procedural fairness found by most authors dominate the negative side effects emphasized by Ord3n3ez et al. (2009) and whether this translates into higher employee effort. Furthermore, we provide evidence whether goal clarity mediates the direct relationship between targets and employee effort.

The paper proceeds as follows: In Chapter 2, we review the literature and develop our hypotheses. In Chapter 3, we describe the data and our dependent and independent variables of interest. Chapter 4 presents our empirical results. Finally, Chapter 5 concludes.

2 Background and Hypotheses Development

2.1 The impact of PMEP on effort

Prior literature has provided conflicting results with respect to the way in which targets influence employee effort and performance. Proponents of classical goal setting theory (e.g. Locke and Latham (1990), Locke and Latham (2002), Locke and Latham (2006)) argue that there is a “positive linear relationship” between the difficulty of the target and task performance. They argue that challenging targets are more motivating because they induce a feeling of success when targets are met and because such targets help employees to grow in the workplace (Locke and Latham (2006)). Furthermore, attaining targets is often linked to bonus payments (e.g. Kampkötter et al. (2017)). According to principal-agent theory, this link between pay and performance aligns employee incentives with company goals and therefore induces effort (Jensen and Murphy (1990)). We further argue that the PMEP in our setting is rather participative, as the employee has at least the possibility to voice concerns during the AI. There is an array of literature on participation in target setting and the relationship to employee effort and performance. Sholihin et al. (2011), for instance, find a positive impact of participation in target setting on goal commitment, which they use as their proxy for employee effort. In a recent literature review, Derfuss (2009) finds a moderately strong relationship between participation in budgeting and employee behaviors beneficial to the organization.

However, there is also evidence suggesting that under certain circumstances, targets might lead to lower effort and performance. In a lab experiment, Seijts and Latham (2001) find that “do your best” outcome goals have a larger effect on performance than specific, difficult outcome goals, while they find the opposite for learning goals. This is the case as employees might focus too much on the attainment of the desired outcome than on learning, which is necessary to reach this outcome. Thus, even proponents of classical goal setting theory admit that setting specific, challenging targets might not have a positive impact on employee performance under every contingency. Li and Sandino (2018) provide evidence that challenging tasks might discourage below-median performers, while Ordóñez et al. (2009) highlight further potentially harmful side effects of target setting, among them crowding-out of intrinsic motivation by extrinsic motivation. Therefore, the impact of targets on employee effort is not unambiguous and the question about how targets affect employee effort *on average* in a large and representative sample remains unanswered.

Since the bulk of literature on target setting and participation in target setting finds a positive impact of the presence of a PMEP on employee effort, we expect the average effect

to be positive but emphasize that it is ultimately an empirical question. We thus formulate the following hypothesis:

Hypothesis 1: *There is a positive effect of the presence of a PME P on employee effort.*

2.2 The effect of appraisal interviews and target agreements on employee effort

Next, we analyze the question whether formalization of targets is of value by differentiating between AIs and TAs. Locke and Latham (1990) state that specific, more formalized targets induce higher levels of effort as “do your best” targets. Hartmann and Slapničar (2009) further posit that formal targets are superior to informal ones as they provide higher feedback quality. They arguably specify the performance dimensions being evaluated and their link to rewards better, ultimately increasing goal orientation and motivation. We argue that the two components of the PME P considered by us possess different degrees of formality. As described in more detail in Section 3, an employee can only have a written TA in case she also receives an AI. AIs are themselves not a completely informal way to set targets, as respondents in our data set are asked to only consider pre-scheduled AI meetings. Nevertheless, explicating these targets in a written form via a TA implies *additional* formality.

With respect to AIs, our study is closely related to Kampkötter (2017). By employing representative German data on the employee level from the German Socio-economic Panel (SOEP), he estimates the impact of performance appraisals on job satisfaction. Performance appraisals as a concept are closely related to AIs as both comprise a developmental and an evaluative function (e.g. Boswell and Boudreau (2002)). The developmental function aims at improving an employee’s effectiveness by enhancing her skills, attitudes, and experiences (e.g. via identification of strengths, weaknesses, and training needs, or goal setting). Evaluation, in contrast, consists of comparing the employee’s performance to a certain standard and is often linked to decisions such as pay increases, promotion, or termination decisions. Despite the similarities, we abstain from calling our measure a performance appraisal for several reasons. As pointed out by Aguinis et al. (2013) and recently by Bayo-Moriones et al. (2019), performance appraisals and performance management are two interrelated yet distinct concepts in the sense that performance management is more general. Likewise, we argue that the AI as part of our PME P is more general than a performance appraisal as the focus is not only on past performance, but also future potential. Indeed, the specific question

used by Kampkötter (2017) focuses more closely on the evaluative function, while our measure is more balanced between the evaluative and developmental functions. The author finds an overall positive effect of performance appraisals on job satisfaction, in particular when performance appraisals are linked to monetary outcomes. However, the question whether (performance) AIs are successful in increasing employee performance remains unexplored in this study. Furthermore, the data used only provides information on the employee but very crude information on the establishment the employee works in, an issue we are able to tackle using linked employer-employee data. Considering that Kampkötter (2017) provides evidence that performance appraisals on average lead to higher job satisfaction, an employee attitude shown to lead to higher employee and organizational productivity (Krekel et al. (2019)), and that AIs are used to set targets, we expect a positive impact of AIs on employee effort.

In contrast to the extensively studied topic of performance appraisals, research specifically focusing on written TAs for employees is scarce.⁶ It is likely that TAs on the individual level influence performance on the establishment level indirectly via beneficial employee behavior, in particular via increased effort provision, but this has not been shown so far. Also, the role of mediators such as goal clarity and procedural fairness has not been studied yet.

We thus examine whether there is a positive effect of AIs, as our less formal way of target setting, on employee effort and whether a higher degree of formality as implied by written TAs provides *additional* value. We therefore formulate the two following hypotheses:

Hypothesis 2a: *There is a positive effect of the presence of an AI on employee effort.*

Hypothesis 2b: *There is an additional positive effect of the presence of a written TA on employee effort.*

2.3 Mediation analysis

It is likely that our PMEP does not only affect employee effort directly but also indirectly via affecting other employee behaviors which then in turn affect effort. In particular, we consider procedural fairness and goal clarity as potential mediators.

With respect to procedural fairness, our study is most closely related to Sholihin et al. (2011) who use data comprising 54 managers from a UK financial services institution and analyze both the direct effect of participative target setting on employee effort and the respective channels mediating this direct relationship. They find that the direct effect is fully mediated

⁶Using German establishment-level data, Kampkötter et al. (2017) find that establishments using TAs achieve 5% higher sales, implying a positive impact of TAs on organizational performance.

by procedural fairness and interpersonal trust meaning that the direct effect becomes insignificant when these mediators are accounted for. We believe that, due to our methodological advantages, our study serves as a complement as we analyze firms in different industries and of different size.

There are various theories dedicated to understanding organizational justice.⁷ Fairness Heuristics Theory (e.g. Lind and Tyler (1988); Lind (2001); van den Bos et al. (2001)), for instance, argues that in most work situations, individuals are at risk of being exploited. Due to this immanent risk of exploitation, they ask the question if the authority is to be trusted (Cropanzano et al. (2001)). As it is impossible to accurately calculate trustworthiness for each relationship, individuals use heuristics to facilitate the decision. Procedures such as participation or voice signal in-group membership (van den Bos et al. (2001)), ultimately increasing procedural fairness perception. Leventhal (1980) posits that individuals fairness perceptions are influenced by six rules, in particular accuracy, bias suppression, consistency, correctability, ethicality of procedures, and representativeness. Ordóñez et al. (2009), in contrast, state that literature ignores potential negative effects of target setting on fairness perceptions. In particular they argue that setting targets might in fact lead to a feeling of unfairness rather than fairness. This is because employees possess heterogeneous ability, making the same goal easily attainable for some individuals and too difficult to achieve for others. However, tailoring goals to individuals might in turn lead to a feeling of unfairness as some individuals might feel that rewards do not fairly match effort and performance. Despite the negative side effects emphasized by Ordóñez et al. (2009), we follow Sholihin and Pike (2013) in arguing that the PMEP considered by us both give employees some degree of voice, in the sense that it is a rather participative way to set goals, and fulfills many of the six rules put forward by Leventhal (1980). Therefore, we expect a positive effect of the presence of our PMEP on employees' perception of procedural fairness.

There is an array of literature examining the link between procedural fairness and beneficial employee behaviors (e.g. Korsgaard et al. (1995)), but most studies focus on the impact of procedural fairness on employee behaviors like group commitment (e.g. Colquitt (2001)) and not effort per se. Exceptions specifically analyzing the impact of procedural fairness on goal commitment as a concept related to effort are Wentzel (2002) and Sholihin et al. (2011). Wentzel (2002) expects a positive relationship between procedural fairness and goal commitment for two reasons: First, attaining the goal should be in the self-interest of the employee in case procedures are fair (instrumental perspective) and second, compliance with the group

⁷For an overview, see Cropanzano et al. (2001).

policy, in this case the goal, should affirm group membership (relational perspective). Both Wentzel (2002) and Sholihin et al. (2011) find empirical support for this hypothesis. We follow the authors and expect a positive impact of procedural fairness on effort, here proxied by work engagement, and formulate the following hypothesis:

Hypothesis 3: *The direct effect of the presence of a PMEP on employee effort is mediated by procedural fairness.*

As compared to the literature on procedural fairness, studies that specifically examine goal clarity are relatively scarce in the management accounting literature, although clarity about the organization’s goals is essential for performance management. Sholihin and Pike (2013) e.g. state that “the existence of prespecified goals is likely to provide clearer understanding (goal clarity) for organizational members and indicate how they will be evaluated” (p. 32). Further, they argue that “goal specificity⁸ and clarity informs employees of their responsibilities and performance targets” and that the “existence of specific goals will guide employees in deciding where they should direct their attention and effort” (p. 32). Since our PMEP consists of the joint presence of an AI and a written TA, we argue that goals set via this process are highly specific, inform employees about their responsibilities and performance targets and should thus increase goal clarity. With respect to the goal clarity-performance link, Anderson and Stritch (2015) expect that higher goal clarity leads to higher performance by referring to classical goal setting theory (e.g. Dossett et al. (1979)) and indeed find a positive impact in a lab experiment. However, in some circumstances, goal clarity might actually be detrimental to performance, as it might lead to tunnel vision (Seijts and Latham (2001); Anderson and Stritch (2015)). Ordóñez et al. (2009) further argue that goals directed at goal areas might crowd out effort directed at non-goal areas without leading to a greater overall effort. We argue that engagement as our proxy for effort is rather general in the sense that it does not differentiate between effort directed at goal and non-goal areas, such that we can examine whether or not this is the case. We follow the bulk of the literature and expect both a positive relationship between the presence of our PMEP and goal clarity and between goal clarity and employee effort and formulate the following hypothesis:

⁸Note that Sholihin and Pike (2013), in line with Fang et al. (2005), define goal specificity as “the extent to which the goals are clearly defined by a supervisor.” We therefore consider goal specificity as being an antecedent to goal clarity.

Hypothesis 4: *The direct effect of the presence of a PMEP on employee effort is mediated by goal clarity.*

3 Data

In order to examine the nexus between the use of AIs, TAs, and work engagement, we use a new, representative matched employer-employee data set covering German private sector establishments with more than 50 employees, the Linked Personnel Panel (LPP) (for a detailed description of the design of the data set and the sources of the applied constructs see Kampkötter et al. (2016)).⁹ Surveyed establishments are randomly drawn from the IAB establishment panel, a representative annual survey of nearly 16,000 German establishments. We can make use of the four waves 2012, 2014, 2016/17 and 2018/19 of the LPP linked employer-employee data set. In detail, the employer survey covers between 769 and 1,219 establishments per wave. Establishment managers provide information on HRM practices and other firm characteristics. From these establishments, a random sample of employees working within the surveyed establishments (roughly between 6,500 and 7,500 individuals per wave) are interviewed at home via telephone (CATI) or web interface (CAWI) about job characteristics and perceptions, personal characteristics, attitudes towards their organization and behavioral variables. This feature of the data enables us to examine the link between the presence of AIs and TAs on the individual level and engagement, while simultaneously being able to control for organizational characteristics on the level of the establishment. Furthermore, the longitudinal structure of the data allows us to employ panel data methods, which enables us to move closer towards causality.

Our analysis is based on two items from the LPP employee survey. The item we use to measure the presence of an AI asks the interviewee the following question: *“Did you have an appraisal interview with your superior last year (e.g. on your professional growth or staff assessment)? Please consider only appraisal interviews for which an appointment was made.”* This question is then used as a filter question for the item measuring the incidence of a formal TA, implying that, by construction, an employee can only be covered by a TA in case she is also covered by an AI. It is important to stress the formal character of AIs, since respondents should only consider meetings for which a formal appointment was made. The item measuring

⁹The data set is open to any researcher and is available via the Research Data Centre (FDZ) of the German Federal Employment Agency at the Institute for Employment Research (IAB). The DOI is: 10.5164/IAB.LPP1617.de.en.v1. For more details, see Haylock and Kampkötter (2019).

the presence of a TA (in conjunction with an AI) is based on the following question: “*Did your superior agree with you on the objectives fixed in writing during the appraisal interview?*”.

We use these two questions to construct two alternate specifications for our main explanatory variables of interest. We exploit this twofold strategy because we encounter a problem common to fixed effects analyses, namely that we do not have enough within-variation to separate the effect of the presence of the TA from the effect of the presence of the AI (Bender et al. (2018)). This problem stems from the fact that an employee can only have a TA in case she also has an AI. As a result, in a specification including both an AI and a TA dummy, the TA dummy has to be interpreted as an interaction term. Therefore, the control group for the TA variable are all individuals who report having an AI but no TA. In order to identify the effect of TAs using individual fixed effects, we therefore need a sufficient number of employees who report having an AI in both periods and either switch from having no TA to having one or switch from having a TA to having none. However, most individuals either do not switch at all within the four waves available to us, or they *jointly* switch in both the AI and the TA variable. Therefore, we cannot disentangle the effect of the AI from the one of the TA in our individual fixed effects specification in a meaningful way. For our first specification, we thus construct a dummy variable taking the value one if an employee is covered by both an AI and a TA, and zero otherwise. In case the dummy variable takes the value one, we define the employee as being subject to a full PMEP. The coefficient of this variable can be interpreted as the effect of the *joint* presence of AIs and TAs. Our second specification differentiates between the presence of an AI and the presence of a TA by constructing two dummies: One that takes the value one in case an employee is covered by an AI and zero otherwise and one that takes the value one if an employee is covered by *both* an AI *and* a TA. Consequently, the TA dummy measures the *additional* impact of the written TA *in addition* to the effect of the AI.

As a proxy for individual effort, we apply the widely used, internationally validated nine-item work engagement scale UWES-9 by Schaufeli and Bakker (2004). Respondents were asked to indicate to which extent they agree with nine statements regarding their job such as the following on a five point Likert scale: *At my work, I feel bursting with energy.*¹⁰ The reported scores of every single item are then added up and divided by 9, such that the resulting *engagement score* represents an equally weighted average with values between 1 and 5. We further standardize this engagement score in order to make a quantitative interpretation in standard deviations possible. Cronbach’s Alpha of our engagement index is 0.915, suggesting

¹⁰A complete list of the items used can be found in the Appendix 6.1.

a high degree of internal consistency of this construct.

Our two potential mediators are conceptualized as follows. Goal clarity is based on two items from the organizational climate questionnaire by Patterson et al. (2005). Specifically, respondents are asked to state on a five point Likert scale to which extent they agree with the following statements: *“The superiors clearly communicate requirements and objectives”* and *“Everyone who works here is well aware of the long-term plans and direction of this company”*. Again, the individual answers are added, the total score is divided by the number of items, and the resulting index is standardized. Procedural fairness is operated by one item from the justice scale by Kim and Leung (2007), which asks respondents to state on a five point Likert scale to which extent they agree with the following statement: *“The rules and procedures to make decisions are fair.”* We again standardize this variable.

Furthermore, the data allows us to account for a rich set of control variables on the establishment and individual level. Establishment-level controls comprise industry (5 categories), region (north, east, south, west), a set of dummies capturing ownership structure, and a dummy capturing whether or not the establishment is independent. Individual-level controls include sex (0/1), age (8 dummies), supervisory position (0/1), full-time position (0/1), white-collar employee (0/1), monthly net income, type of employment contract (fixed term/permanent), permanent relationship (0/1), highest level of school and occupational or university education (7 dummies), household size, and survey method (CAWI/CATI). Establishment-level controls include industry (5 dummies), regional area (4 dummies), and establishment size (5 dummies). In all regressions, standard errors are clustered on the establishment level. Detailed descriptive statistics on our main dependent and independent variables are displayed in Tables A.1, A.2, and A.3 in the Appendix.

In 2019, the latest year comprised in our data, the mean (unstandardized) engagement index is 3.44, while the median is 3.56. Therefore, we observe a higher probability mass at larger values of the engagement index, indicating that employees in our sample are, on average, rather engaged. The distribution is also rather stable over time (mean values range between 3.68 and 3.76 from 2012 to 2016). In total, our data comprises 16,506 employee-year observations that are non-missing with respect to the AI variable. 8,622 (52.24 %) of these employee-year observations reported having an AI. Out of these, 5,875 (68.14%) also reported having a written TA.

4 Empirical Strategy and Results

4.1 The direct impact of a PMEP on employee effort

In order to investigate Hypothesis 1, i.e. whether there is a positive effect of the presence of a full PMEP on employee effort, we use OLS regressions employing various fixed effects in order to be able to make more causal statements. In addition, all of our specifications include a rich set of covariates on the establishment and individual level, as can be seen in Table 1.

In column (1), we regress work engagement on PMEP by including all controls on the establishment and individual level as well as establishment size and year fixed effects. The positive PMEP coefficient is statistically and economically significant: the magnitude of the coefficient implies that the engagement score of employees covered by a PMEP is, on average, 0.203 standard deviations higher compared to employees without a PMEP. Therefore, column (1) provides first support to Hypothesis 1, indicating that target setting via a PMEP seems to have a positive impact on employee effort. In column (2), we tackle the question whether the impact of our PMEP on employee engagement is driven by the use of variable incentive pay since achieving pre-negotiated targets might be tied to a variable pay component. Thus, it might not be the PMEP *per se* that induces larger employee engagement, but rather the link to variable pay. The simple correlation coefficient between the PMEP variable and the use of variable compensation is 0.24, suggesting that employees who report having a PMEP also tend to have a variable pay component. However, the size of this correlation is not large enough to suggest that the two variables capture the same effect. Results in column (2) are consistent with these descriptives: The coefficient of variable pay is positive and statistically significant, suggesting that employees who have a variable pay component show, on average, a higher work engagement. More important, the magnitude of the PMEP coefficient changes only marginally, suggesting that the relationship between the PMEP and employee engagement is not just driven by incentive pay.

In columns (3) and (4), we additionally include establishment and year fixed effects to reduce the likelihood of omitted variable bias and to take a further step towards causality. In column (3), we separately include establishment and year fixed effects. Thereby, we are able to account for time-constant unobserved heterogeneity on the level of the establishment while simultaneously controlling for general market trends through time fixed effects. Results are robust as the coefficient of PMEP is positive and statistically significant. In fact, the magnitude of the effect even increases from 0.193 to 0.245, suggesting that the effect gets larger when taking the within-firm rather than the across-firm perspective (this difference is

Table 1: Direct Effect of PMEP on Work Engagement

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Engagement Index (std.)					
PMEP	0.203*** (0.0182)	0.193*** (0.0184)	0.245*** (0.0207)	0.253*** (0.0210)	0.0783*** (0.0285)	0.0597** (0.0239)
Variable Pay		0.0548*** (0.0196)	0.0706*** (0.0221)	0.0636*** (0.0227)	-0.00861 (0.0305)	-0.0133 (0.0254)
Establishment Controls	yes	yes	yes	yes	yes	yes
Employee Controls	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes			yes
Size FE	yes	yes				yes
Establishment FE			yes			
Establishment \times Year FE				yes	yes	
Lagged Engagement Individual FE					yes	yes
Constant	-0.324*** (0.0629)	-0.346*** (0.0636)	-0.418*** (0.118)	-0.466*** (0.0433)	-0.187*** (0.0552)	0.223* (0.124)
Observations	16,506	16,498	16,498	16,026	4,296	16,498
Number of Employees					12,057	
R-squared (within)	0.076	0.077	0.189	0.225	0.631	0.028

The dependent variable Engagement Index is an index containing the weighted average of nine items and is standardized. All underlying items are measured on a 5-point Likert scale (between 1 and 5). In all columns, ordinary least squares regressions are applied. Employee controls comprise female (0/1), age (8 dummies), supervisory position (0/1), white-collar employee (0/1), full-time position (0/1), monthly net income, type of employment contract (fixed term/permanent), permanent relationship (0/1), highest level of training qualification (7 dummies), household size, and survey method (CAWI/CATI). Establishment-level controls include industry (5 dummies), regional area (4 dummies), and establishment size (5 dummies). Standard errors clustered on the establishment level in parentheses. The symbols *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

also statistically significant, $p = 0.01$). Here, we also account for unobserved establishment level heterogeneity such as time-constant performance management or leadership culture. In column (4), we include an interaction between establishment and year fixed effects. We thereby allow unobserved, establishment-specific characteristics to vary over time. Results remain virtually unchanged, both with respect to the magnitude and significance of the PMEP coefficient.

Columns (5) and (6) account for unobserved individual heterogeneity that both determines an employee’s engagement level and coverage by a PMEP. In column (5) we apply a lagged dependent variable (LDV) model by including lagged individual engagement as an explanatory variable. This specification accounts for the possibility that engagement is rather stable over time and is not caused by the presence of a PMEP. Results indicate that this is partly the case, as the size of the effect decreases from 0.259 to 0.0783 standard deviations. However, even though the magnitude of the coefficient is reduced, it is still sizeable and statistically significant, indicating that the presence of a PMEP does indeed have a positive impact on employee engagement even if past engagement levels are controlled for. In column (6), we finally conduct an individual fixed effects regression, now explicitly controlling for unobserved individual-level heterogeneity. The coefficient is positive and statistically significant, thus confirming our previous results: a within-person change in PMEP leads to an increase in employee engagement of around 0.06 standard deviations.¹¹ To conclude, our results provide support for Hypothesis 1, as we find a robust positive impact of the PMEP on individual work engagement. On average, the positive effects of targets on employee effort as emphasized by classical goal setting theory thus seem to outweigh potential negative side effects in our representative sample.

4.2 Separating the impact of appraisal interviews and target agreements on individual performance

In this section, we split the PMEP into its components. The aim is to analyze if the observed positive effect of the PMEP on employee effort is driven by AIs alone or whether additional formalization via written TAs in the performance management process provides additional value.

In column (1) of Table 2, we check whether there is a positive relationship between the presence of AIs and employee engagement. Results show a positive and statistically significant coefficient of 0.193 standard deviations.¹² It is noteworthy that the coefficient in this

¹¹In order to tackle issues related to selection bias and reverse causality, we estimate the IV method proposed by Lewbel (2012) by using Stata’s `ivreg2h` command developed by Baum and Schaffer (2012) as a robustness check for our baseline specification in column (1). These issues might arise in case engaged individuals self-select into having a PMEP or in case they are chosen for having a PMEP based on their previous level of engagement. Our results remain qualitatively the same, such that the coefficient of PMEP is still positive and statistically significant.

¹²Note that we directly control for variable pay in this specification. Results do not change when omitting variable pay from the regression equation.

Table 2: Direct Effects of AIs and TAs on Work Engagement

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Engagement Index (std.)					
Appraisal Interview (AI)	0.193*** (0.0193)	0.124*** (0.0253)	0.144*** (0.0268)	0.150*** (0.0272)	0.114*** (0.0358)	0.0959*** (0.0301)
Target Agreement (TA)		0.111*** (0.0238)	0.162*** (0.0265)	0.169*** (0.0267)	0.0154 (0.0335)	0.0104 (0.0268)
Variable Pay	0.0537*** (0.0194)	0.0483** (0.0196)	0.0675*** (0.0220)	0.0603*** (0.0227)	-0.0107 (0.0305)	-0.0152 (0.0253)
Establishment Controls	yes	yes	yes	yes	yes	yes
Employee Controls	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes			
Size FE	yes	yes				yes
Establishment FE			yes			
Establishment FE × Year FE				yes	yes	
Lagged Engagement Individual FE					yes	yes
Constant	-0.391*** (0.0658)	-0.359*** (0.0636)	-0.443*** (0.119)	-0.505*** (0.0434)	-0.221*** (0.0557)	0.205 (0.124)
Observations	16,528	16,498	16,498	16,026	4,296	16,498
Number of Employees						12,057
R-squared (within)	0.077	0.078	0.190	0.227	0.632	0.030

The dependent variable Engagement Index contains the weighted average of nine items and is standardized. All underlying items are measured on a 5-point Likert scale (between 1 and 5). In all columns, ordinary least squares regressions are applied. Employee controls comprise female (0/1), age (8 dummies), supervisory position (0/1), white-collar employee (0/1), full-time position (0/1), monthly net income, type of employment contract (fixed term/permanent), permanent relationship (0/1), highest level of training qualification (7 dummies), household size, and survey method (CAWI/CATI). Establishment-level controls include industry (5 dummies), regional area (4 dummies), and establishment size (5 dummies). Standard errors clustered on the establishment level in parentheses. The symbols *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

specification is relatively similar to the coefficient of the PMEP variable in column (2) of Table 1. Therefore, when not controlling for TAs, the AI variable seems to pick up the entire effect of the presence of a PMEP. This can have two reasons: Either the effect of a PMEP on work engagement is entirely driven by AIs such that formalization via written TAs does not

increase engagement or there exists an omitted variable bias problem in column (1) of Table 2, such that part of the effect of AIs on engagement is actually due to the additional presence of a written TA. We test this conjecture in column (2) by including the TA dummy. The results show that the latter seems to be true. Both the AI and the TA variables show positive and statistically significant coefficients. The magnitude of the coefficient of the AI variable is reduced to 0.124, while the TA variable reports a coefficient of 0.111, indicating that approximately half of the effect of AIs as reported in column (1) can actually be attributed to the additional presence of a written TA.

Consistent with subsection 4.1, we also run two different specifications including establishment fixed effects. In column (3), we include establishment fixed effects and year fixed effects while we interact these two fixed effects in column (4). In analogy to Table 1, the results remain qualitatively the same. Results up to this point thus provide support for hypotheses 2a and 2b. There seems to be both a positive impact of AIs on employee effort as proxied by work engagement and an *additional* positive impact of formalization via a written TA.

In the last step, we again include lagged engagement in column (5) and individual fixed effects in column (6). As can be seen in Table 2, the effect of AIs on work engagement is still positive and statistically significant. However, the TA variable now turns statistically insignificant. The same is true for the individual fixed effects regression in column (6). One potential explanation refers to the joint presence of AIs and TAs in many firms, i.e. AIs and TAs are often introduced jointly for employees. If this is true, estimating the isolated effects of AIs and TAs in lagged dependent variable and fixed effects specifications is almost impossible. The reasoning is the following: Fixed effects (and also lagged dependent variable specifications) require a certain degree of variation within individuals, implying the need for a sufficient amount of *switchers*. In order to identify a significant effect for our TA variable, we would need a sufficient number of individuals who switch from having no TA in one period to having one in the next period or vice versa. At this point it is crucial to remember that by construction of the data set and also plausibly in firms, an individual can only have a TA if she also has an AI. A switch from zero to one in the TA variable can thus capture two different events: Either the employee obtains an AI and a TA *jointly* or the employee has already had an AI in the previous period and *in addition* obtains a TA in the actual period. In case AIs and TAs are introduced jointly, one cannot isolate the effects of an introduction of AIs and TAs. The same reasoning applies in case the TA is abolished. Descriptive results indeed suggest that a large fraction of individuals who switch from having no (an) AI to having one (none) simultaneously switch from having no (a) TA to having one (none). Out

of 503 (449) employees included in the individual fixed effects analysis who report switching from zero to one (from one to zero) with respect to our AI variable, 221 (213) also report switching from zero to one (one to zero) with respect to the TA variable. Therefore, the effect of the TA variable can only be identified if the data contains enough individuals that report having an AI in two consecutive periods and switch from having no TA to having one or vice versa. Out of the 2,004 individuals who report having an AI in two consecutive periods, only 194 report switching from having no TA to having one, while 230 employees report switching from having a TA to having none. Hence, it is highly likely that this lack of variation causes our lagged dependent variable and fixed effects estimates for TAs to be insignificant. At least, it nicely shows how challenging it can be to causally analyze isolated effects of simultaneously applied performance management practices using firm data.

In sum, our analyses provide support for both hypotheses 2a and 2b. Results indicate a positive impact of AIs on employee effort. This effect is robust across all specifications. Results also provide evidence that formalization via written TAs further increases employee effort, by showing a positive and statistically significant coefficient in all but the specifications including lagged engagement and individual fixed effects. We are fairly certain that this is due to a lack in within-variation. However, we acknowledge that our results with respect to Hypothesis 2b are somewhat weaker as the ones regarding Hypothesis 2a.

4.3 Mediation analysis

In this section, we present our test of hypotheses 3 and 4, i.e. whether the direct effect of the PMEP on employee effort is mediated by procedural fairness and goal clarity.¹³ To examine these two potential channels, we follow the mediation analysis approach put forward by Baron and Kenny (1986) and estimate three different equations. First, the potential mediator (goal clarity, procedural fairness) is regressed on the independent variable, here PMEP. In a second step, the dependent variable (engagement) is regressed on the potential mediator. In a third step, the dependent variable is regressed on both the mediator and the independent variable. Full mediation is achieved if the respective coefficients of interest are statistically significant in the first two regressions and if a previously significant relationship between the independent and the dependent variable in the first regression becomes insignificant when including the mediator in the third regression. A variable partially mediates the relationship between an

¹³Note that in this analysis, we do not differentiate between AIs and TAs. However, looking at the two performance management instruments separately, we find the same patterns as in the analysis presented above.

independent and a dependent variable if it significantly decreases the direct path between the independent and the dependent variable rather than completely eliminating it. Very often statistical relationships, such as the relationship between the presence of PMEP and employee engagement have multiple causes, such that full mediation is rather unlikely.

Table 3 shows the results of our mediation analysis. As we consider column (6) of Table 1 to be our most reliable specification, we also use individual fixed effects for our mediation analysis.¹⁴

Column (1) is equivalent to column (6) in Table 1 and again shows the positive effect of the PMEP on employee engagement. Columns (2) and (3) depict the first step of the actual mediation analysis, the regression of procedural fairness and goal clarity on the PMEP. As can be seen from the coefficient of the PMEP variable in column (2), there seems to be a significantly positive association between the presence of a PMEP and goal clarity. The PMEP coefficient in column (3) also indicates a positive association with procedural fairness. This coefficient is, however, only significant at the 10 percent level. Column (4) depicts the second step of the mediation analysis, namely the regression of the dependent variable, engagement index, on the potential mediators, goal clarity and procedural fairness.¹⁵ Both the coefficients of the goal clarity index and the procedural fairness variable are positive and statistically significant, suggesting that both goal clarity and procedural fairness are positively associated with employee engagement. Column (5) depicts the third step, the regression of the dependent variable, employee engagement, on both potential mediators, goal clarity and procedural fairness, and the PMEP as our main independent variable. The results show that all variables of interest report a positive and statistically significant coefficient. Furthermore, the effect size and statistical significance of the PMEP variable decreases.

In sum, these results are consistent with partial mediation and we find evidence for hypotheses 3 and 4. The PMEP seems to increase employees' feeling of procedural fairness by giving them some degree of voice. The increase in perceived procedural fairness then results in a higher level of work engagement or effort, respectively. Furthermore, we find that the PMEP increases goal clarity. This shows that performance management is a useful tool for firms as it helps to make the organizational goals more visible to the workforce. Again, this increase in goal clarity results in an increase in overall employee effort. However, results indicate that procedural fairness and goal clarity do not fully, but only partially mediate the

¹⁴Note that the results are robust to and even more significant in all the other specifications applied above.

¹⁵In unreported further analyses, we also regress employee engagement on the two potential mediators separately. Results are qualitatively robust, such that in both regressions, the coefficient of interest is positive and statistically significant.

Table 3: Mediation Analysis

Variables	(1) Engagement Index (std.)	(2) Goal Clarity Index (std.)	(3) Procedural Fairness (std.)	(4) Engagement Index (std.)	(5) Engagement Index (std.)
PMEP	0.0597** (0.0239)	0.0834** (0.0354)	0.0657* (0.0355)		0.0461* (0.0236)
Goal Clarity Index				0.0859*** (0.0142)	0.0846*** (0.0143)
Procedural Fairness				0.102*** (0.0123)	0.102*** (0.0123)
Variable Pay	-0.0133 (0.0254)	0.0553 (0.0347)	0.0256 (0.0362)	-0.0199 (0.0245)	-0.0196 (0.0246)
Establishment Controls	yes	yes	yes	yes	yes
Employee Controls	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes
Size FE	yes	yes	yes	yes	yes
Individual FE	yes	yes	yes	yes	yes
Constant	0.223* (0.124)	0.0389 (0.168)	-0.123 (0.173)	0.235* (0.127)	0.233* (0.121)
Number of Employees	12,057	12,251	12,188	11,992	11,974
Observations	16,498	16,839	16,740	16,395	16,351
R-squared (within)	0.028	0.019	0.011	0.069	0.070

The dependent variables are constructed as follows. Engagement Index contains the equally weighted average of nine items. Procedural fairness contains one item. Goal Clarity Index contains the equally weighted average of two items. All items are measured on a 5-point Likert scale (between 1 and 5). The resulting variables are standardized. In all columns, ordinary least squares regressions are used. Employee controls comprise female (0/1), age (8 dummies), supervisory position (0/1), white-collar employee (0/1), full-time position (0/1), monthly net income, type of employment contract (fixed term/permanent), permanent relationship (0/1), highest level of training qualification (7 dummies), household size, and survey method (CAWI/CATI). Establishment-level controls include industry (5 dummies), regional area (4 dummies), and establishment size (5 dummies). Standard errors clustered on the establishment level in parentheses. The symbols *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

impact of the PMEP on work engagement, as the direct effect is still marginally significant.

5 Discussion

Classical goal setting theory (e.g. Locke and Latham (1990, 2002, 2006)) has long emphasized that there is a positive link between target setting mechanisms and employee effort. However, recent contributions (Barsky (2008); Ordóñez et al. (2009); Liu and Zhang (2015); Eyring and Narayanan (2018); Holzhaecker et al. (2019)) provide evidence that the way goals are set is important and that goal setting might also have negative side effects that can actually outweigh its potential benefits. Most contributions in the management accounting literature studying the impact of target setting on employee effort rely on experiments or cross-sectional single firm case studies. Generalizable evidence about how targets influence effort *on average* using large and representative data is missing. Indeed, Sholihin et al. (2011) call for such evidence based on “larger samples from various organizations determined randomly” (p. 145). We address this gap in the literature by investigating the impact of performance management, focusing on appraisal interviews and written target agreements, on individual effort. We do so by making use of four waves of the Linked Personnel Panel (LPP), a large and representative German linked employer-employee data set. As compared to prior studies, these data enable us to both make more causal statements by employing various fixed effects and to test the external validity using a representative data set.

Our results show a robust positive and statistically significant effect of the presence of a performance management and evaluation process (PMEP) on employee effort. When dividing the PMEP into its single components, we find a positive and statistically significant effect of AIs on work engagement. Furthermore, our results show a positive *additional* impact of TAs on work engagement. This effect is statistically significant in all but our lagged dependent variable and individual fixed effects estimations. We explicitly discuss the challenge of analyzing isolated effects of performance management practices using firm data, namely the lag of within-variation: there are too few individuals in our data switching from having no (a) TA to having one (none) without jointly making the same switch in the AI variable. This makes it rather difficult to draw a causal statement about the additional effect of written target agreements on individual effort. In a next step, we present the results of a mediation analysis to learn more about the potential channels underlying our core results. We find that the direct effect of PMEP on work engagement is partially mediated by procedural fairness and goal clarity.

Of course, this study is not without weaknesses, the most important one being causality. Even though we certainly move closer to causality than previous literature using firm data

does, we cannot be sure whether the effects we find are really causal, as there is no random assignment of employees into performance management practices. We try to take this possibility into account by running further robustness checks, which confirm our baseline results. As described above, the main advantage of our data is that, due to the large sample size, the representativeness, and the panel structure of the data, we are able to make more causal and generalizable statements. As a consequence, however, the disadvantage is that we do not have information about the exact content of the AIs and the respective targets. In order to get a more complete picture of the effects of performance management on employee effort, we therefore regard our study as a good complement to prior single-firm econometric case studies.

References

- Aguinis, H., R. K. Gottfredson, and H. Joo (2013). Avoiding a “me” versus “we” dilemma: Using performance management to turn teams into a source of competitive advantage. *Business Horizons* 56(4), 503–512.
- Anderson, D. M. and J. M. Stritch (2015). Goal clarity, task significance, and performance: evidence from a laboratory experiment. *Journal of Public Administration Research and Theory* 26(2), 211–225.
- Bakker, A. B. (2017). Strategic and proactive approaches to work engagement. *Organizational Dynamics* 46(2), 67–75.
- Baron, R. M. and D. A. Kenny (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51(6), 1173.
- Barsky, A. (2008). Understanding the ethical cost of organizational goal-setting: A review and theory development. *Journal of Business Ethics* 81(1), 63–81.
- Baum, C. F. and M. E. Schaffer (2012, November). IVREG2H: Stata module to perform instrumental variables estimation using heteroskedasticity-based instruments. Statistical Software Components, Boston College Department of Economics.
- Bayo-Moriones, A., J. E. Galdon-Sanchez, and S. Martinez-de Morentin (2019). Performance appraisal: dimensions and determinants. *The International Journal of Human Resource Management*, 1–32.
- Bender, S., N. Bloom, D. Card, J. Van Reenen, and S. Wolter (2018). Management practices, workforce selection, and productivity. *Journal of Labor Economics* 36(1), 371–409.
- Boswell, W. R. and J. W. Boudreau (2002). Separating the developmental and evaluative performance appraisal uses. *Journal of Business and Psychology* 16(3), 391–412.
- Cederblom, D. (1982). The performance appraisal interview: A review, implications, and suggestions. *Academy of Management Review* 7(2), 219–227.
- Christ, M. H., K. L. Sedatole, and K. L. Towry (2012). Sticks and carrots: The effect of contract frame on effort in incomplete contracts. *The Accounting Review* 87(6), 1913–1938.

- Colquitt, J. A. (2001). On the dimensionality of organizational justice: a construct validation of a measure. *Journal of Applied Psychology* 86(3), 386.
- Cropanzano, R., Z. S. Byrne, D. R. Bobocel, and D. E. Rupp (2001). Moral virtues, fairness heuristics, social entities, and other denizens of organizational justice. *Journal of Vocational Behavior* 58(2), 164–209.
- Derfuss, K. (2009). The relationship of budgetary participation and reliance on accounting performance measures with individual-level consequent variables: a meta-analysis. *European Accounting Review* 18(2), 203–239.
- Dossett, D. L., G. P. Latham, and T. R. Mitchell (1979). Effects of assigned versus participatively set goals, knowledge of results, and individual differences on employee behavior when goal difficulty is held constant. *Journal of Applied Psychology* 64(3), 291.
- Eyring, H. and V. Narayanan (2018). Performance effects of setting a high reference point for peer-performance comparison. *Journal of Accounting Research* 56(2), 581–615.
- Fang, E., K. R. Evans, and S. Zou (2005). The moderating effect of goal-setting characteristics on the sales control systems–job performance relationship. *Journal of Business Research* 58(9), 1214–1222.
- Ferreira, A. and D. Otley (2009). The design and use of performance management systems: An extended framework for analysis. *Management Accounting Research* 20(4), 263–282.
- Fisher, J. G., J. R. Frederickson, and S. A. Peffer (2000). Budgeting: An experimental investigation of the effects of negotiation. *The Accounting Review* 75(1), 93–114.
- Franco-Santos, M., L. Lucianetti, and M. Bourne (2012). Contemporary performance measurement systems: A review of their consequences and a framework for research. *Management Accounting Research* 23(2), 79–119.
- Gordon, M. E. and L. P. Stewart (2009). Conversing about performance: Discursive resources for the appraisal interview. *Management Communication Quarterly* 22(3), 473–501.
- Hartmann, F. and S. Slapničar (2009). How formal performance evaluation affects trust between superior and subordinate managers. *Accounting, Organizations and Society* 34(6–7), 722–737.

- Haylock, M. and P. Kampkötter (2019). Dataset: The linked personnel panel (lpp). *Data in Brief* 27, 104824.
- Holzhaecker, M., S. Kramer, M. Matějka, and N. Hoffmeister (2019). Relative target setting and cooperation. *Journal of Accounting Research* 57(1), 211–239.
- Jensen, M. C. and K. J. Murphy (1990). Performance pay and top-management incentives. *Journal of Political Economy* 98(2), 225–264.
- Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of management journal* 33(4), 692–724.
- Kampkötter, P. (2017). Performance appraisals and job satisfaction. *The International Journal of Human Resource Management* 28(5), 750–774.
- Kampkötter, P., K. Marggraf, and J.-H. Zimmermann (2017). Determinants and effects of formal target agreements: an empirical investigation of german firms. *Review of Managerial Science* 11(1), 1–18.
- Kampkötter, P., J. Mohrenweiser, D. Sliwka, S. Steffes, and S. Wolter (2016). Measuring the use of human resources practices and employee attitudes: the linked personnel panel. In *Evidence-based HRM: a Global Forum for Empirical Scholarship*, Volume 4, pp. 94–115.
- Kim, T.-Y. and K. Leung (2007). Forming and reacting to overall fairness: A cross-cultural comparison. *Organizational Behavior and Human Decision Processes* 104(1), 83–95.
- Korsgaard, M. A., D. M. Schweiger, and H. J. Sapienza (1995). Building commitment, attachment, and trust in strategic decision-making teams: The role of procedural justice. *Academy of Management Journal* 38(1), 60–84.
- Krekel, C., G. Ward, and J.-E. De Neve (2019). Employee wellbeing, productivity, and firm performance. *Saïd Business School WP* 4.
- Lau, C. M. and C. Buckland (2001). Budgeting - the role of trust and participation: A research note. *Abacus* 37(3), 369–388.
- Leventhal, G. S. (1980). What should be done with equity theory? In *Social exchange*, pp. 27–55. Springer.

- Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics* 30(1), 67–80.
- Li, S. X. and T. Sandino (2018). Effects of an information sharing system on employee creativity, engagement, and performance. *Journal of Accounting Research* 56(2), 713–747.
- Libby, T. (1999). The influence of voice and explanation on performance in a participative budgeting setting. *Accounting, Organizations and Society* 24(2), 125–137.
- Libby, T. (2001). Referent cognitions and budgetary fairness: A research note. *Journal of Management Accounting Research* 13(1), 91–105.
- Lind, E. and T. Tyler (1988). Critical issues in social justice. the social psychology of procedural justice. new york, ny, us.
- Lind, E. A. (2001). Fairness heuristic theory: Justice judgments as pivotal cognitions in organizational relations. *Advances in Organizational Justice* 56(8).
- Liu, X. and Y. Zhang (2015). Effects of target timing and contract frame on individual performance. *European Accounting Review* 24(2), 329–345.
- Locke, E. A. and G. P. Latham (1990). *A theory of goal setting & task performance*. Prentice-Hall, Inc.
- Locke, E. A. and G. P. Latham (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist* 57(9), 705.
- Locke, E. A. and G. P. Latham (2006). New directions in goal-setting theory. *Current Directions in Psychological Science* 15(5), 265–268.
- Marginson, D. and S. Ogden (2005). Coping with ambiguity through the budget: the positive effects of budgetary targets on managers’ budgeting behaviours. *Accounting, Organizations and Society* 30(5), 435–456.
- Ordóñez, L. D., M. E. Schweitzer, A. D. Galinsky, and M. H. Bazerman (2009). Goals gone wild: The systematic side effects of overprescribing goal setting. *Academy of Management Perspectives* 23(1), 6–16.
- Otley, D. (1999). Performance management: a framework for management control systems research. *Management Accounting Research* 10(4), 363–382.

- Patterson, M. G., M. A. West, V. J. Shackleton, J. F. Dawson, R. Lawthom, S. Maitlis, D. L. Robinson, and A. M. Wallace (2005). Validating the organizational climate measure: links to managerial practices, productivity and innovation. *Journal of Organizational Behavior* 26(4), 379–408.
- Podsakoff, P. M. and J.-L. Farh (1989). Effects of feedback sign and credibility on goal setting and task performance. *Organizational Behavior and Human Decision Processes* 44(1), 45–67.
- Renn, R. W., C. Danehower, P. M. Swiercz, and M. L. Icenogle (1999). Further examination of the measurement properties of leifer & mcgannon’s (1986) goal acceptance and goal commitment scales. *Journal of Occupational and Organizational Psychology* 72(1), 107–113.
- Schaufeli, W. B. and A. B. Bakker (2003). Utrecht work engagement scale: Preliminary manual. *Occupational Health Psychology Unit, Utrecht University, Utrecht* 26.
- Schaufeli, W. B. and A. B. Bakker (2004). Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior* 25(3), 293–315.
- Seijts, G. H. and G. P. Latham (2001). The effect of distal learning, outcome, and proximal goals on a moderately complex task. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior* 22(3), 291–307.
- Sholihin, M. and R. Pike (2009). Fairness in performance evaluation and its behavioural consequences. *Accounting and Business Research* 39(4), 397–413.
- Sholihin, M. and R. Pike (2013). Investigating the determinants of perceived procedural fairness in performance evaluation. *Journal of Applied Management Accounting Research* 11(1), 29.
- Sholihin, M., R. Pike, M. Mangena, and J. Li (2011). Goal-setting participation and goal commitment: Examining the mediating roles of procedural fairness and interpersonal trust in a uk financial services organisation. *The British Accounting Review* 43(2), 135–146.

- van den Bos, K., E. A. Lind, and H. A. Wilke (2001). The psychology of procedural and distributive justice viewed from the perspective of fairness heuristic theory.
- Voußem, L., S. Kramer, and U. Schäffer (2016). Fairness perceptions of annual bonus payments: The effects of subjective performance measures and the achievement of bonus targets. *Management Accounting Research* 30, 32–46.
- Wentzel, K. (2002). The influence of fairness perceptions and goal commitment on managers' performance in a budget setting. *Behavioral Research in Accounting* 14(1), 247–271.
- Zapata-Phelan, C. P., J. A. Colquitt, B. A. Scott, and B. Livingston (2009). Procedural justice, interactional justice, and task performance: The mediating role of intrinsic motivation. *Organizational Behavior and Human Decision Processes* 108(1), 93–105.

6 Appendix

6.1 Work Engagement - List of Items

Each of the following items should be answered on a five-point Likert scale from 1 (daily) to 5 (never):

1. At my work, I feel bursting with energy.
2. At my job, I feel strong and vigorous.
3. I am enthusiastic about my job.
4. My job inspires me.
5. When I get up in the morning, I feel like going to work.
6. I feel happy when I am working intensely.
7. I am proud of the work that I do.
8. I am immersed in my work.
9. I get carried away when I am working.

6.2 Descriptive Statistics

Table A.1: Descriptive Statistics - Main Dependent and Independent Variables

Variables	(1) N	(2) Mean	(3) Std. Dev.	(4) Min	(5) Max
Appraisal Interview	16,506	0.522	0.500	0	1
Target Agreement/PMEP	16,506	0.356	0.479	0	1
Engagement Index (standardized)	16,506	-0.00748	1.003	-3.206	1.588
Goal Clarity Index (standardized)	16,469	-0.00475	1.003	-2.792	4.563
Procedural Fairness (standardized)	16,392	-0.0142	1.011	-2.437	1.611
Engagement Index (non-standardized)	16,506	3.669	0.837	1	5
Engagement - Energy	16,506	3.388	1.006	1	5
Engagement - Strong	16,506	4.019	0.859	1	5
Engagement - Enthusiastic	16,506	3.755	1.027	1	5
Engagement - Inspiring	16,506	3.369	1.262	1	5
Engagement - Feel Like Working	16,506	3.499	1.133	1	5
Engagement - Happy	16,506	3.819	1.083	1	5
Engagement - Proud	16,506	4.088	1.007	1	5
Engagement - Immersed	16,506	3.623	1.155	1	5
Engagement - Carried Away	16,506	3.460	1.164	1	5
Procedural Fairness (non-standardized)	16,392	3.394	0.999	1	5
Goal Clarity - (non-standardized)	16,469	3.653	0.955	1	8
Goal Clarity - Long-Term Plans	16,475	3.558	1.196	1	8
Goal Clarity - Requirements & Objectives	16,500	3.747	1.048	1	8

Table A.2: Descriptive Statistics - Establishment-Level Control Variables

Variables	(1) N	(2) Mean	(3) Std. Dev.	(4) Min	(5) Max
Establishment-Level Controls:					
Establishment independent (0/1)	16,506	0.693	0.461	0	1
Industry - Manufacturing	16,506	0.299	0.458	0	1
Industry - Metal, Electronics, Automotive	16,506	0.402	0.490	0	1
Industry - Trade, Transportation, News	16,506	0.102	0.303	0	1
Industry - Business-Related Services	16,506	0.125	0.331	0	1
Industry - Information/Communication	16,506	0.0719	0.258	0	1
Region - North	16,506	0.192	0.394	0	1
Region - East	16,506	0.251	0.434	0	1
Region - South	16,506	0.276	0.447	0	1
Region - West	16,506	0.281	0.449	0	1
Size (Number of Employees) - Less Than 100	16,506	0.127	0.333	0	1
Size (Number Employees) - 100 to 249	16,506	0.231	0.422	0	1
Size (Number Employees) - 250 to 499	16,506	0.230	0.421	0	1
Size (Number Employees) - More Than 500	16,506	0.412	0.492	0	1
Principal Owner - Family/Founder	16,506	0.429	0.495	0	1
Principal Owner - Management /Entrepreneurship	16,506	0.151	0.358	0	1
Principal Owner - Financial Investor	16,506	0.0931	0.291	0	1
Principal Owner - Widely Held Stock Capital Market	16,506	0.110	0.312	0	1
Principal Owner - Government/ Public Sector	16,506	0.0233	0.151	0	1
Principal Owner - Other	16,506	0.194	0.395	0	1

Table A.3: Descriptive Statistics - Employee-Level Control Variables

VARIABLES	(1) N	(2) Mean	(3) Std. Dev.	(4) Min	(5) Max
Age Category - under 25 (0/1)	16,506	0.0385	0.192	0	1
Age Category - 25 to 39 (0/1)	16,506	0.233	0.423	0	1
Age Category - 40 to 54 (0/1)	16,506	0.521	0.500	0	1
Age Category - over 55 (0/1)	16,506	0.207	0.405	0	1
Bonuses/Extra Payments (0/1)	16,498	0.594	0.491	0	1
Education - None (0/1)	16,506	0.00418	0.0645	0	1
Education - Lower Secondary School (0/1)	16,506	0.220	0.414	0	1
Education - Intermediate Secondary School (0/1)	16,506	0.424	0.494	0	1
Education - University of Applied Sciences Entrance Qualification (0/1)	16,506	0.110	0.313	0	1
Education - General Higher Education Entrance Qualification (0/1)	16,506	0.235	0.424	0	1
Education - Other (0/1)	16,506	0.00685	0.0825	0	1
Female (0/1)	16,506	0.272	0.445	0	1
Fixed-Term Contract (0/1)	16,506	0.0451	0.208	0	1
Full Time/Part Time (0/1)	16,506	0.127	0.333	0	1
Net Income (in Euros)	16,506	2,418	1,842	1	74,221
Number Members Household	16,506	2.776	1.228	1	14
Serious Relationship (0/1)	16,506	0.841	0.366	0	1
Supervisor (0/1)	16,506	0.290	0.454	0	1
Training Qualification - None (0/1)	16,506	0.0210	0.143	0	1
Training Qualification - Apprenticeship (0/1)	16,506	0.462	0.499	0	1
Training Qualification - Vocational/ Business School (0/1)	16,506	0.0937	0.291	0	1
Training Qualification - Master Craftsman/ Technical College (0/1)	16,506	0.206	0.404	0	1
Training Qualification - University of Applied Sciences (0/1)	16,506	0.0992	0.299	0	1
Training Qualification - University (0/1)	16,506	0.114	0.318	0	1
Training Qualification - Other (0/1)	16,506	0.00436	0.0659	0	1
White Collar (0/1)	16,506	0.625	0.484	0	1