



# Designing Ethical Management Control: Overcoming the Harmful Effect of Management Control Systems on Job-Related Stress

Stefan Linder<sup>1</sup> · Bernard Leca<sup>1</sup> · Adrián Zicari<sup>1</sup> · Veronica Casarin<sup>1</sup>

Received: 29 October 2018 / Accepted: 23 March 2020 / Published online: 13 April 2020  
© Springer Nature B.V. 2020

## Abstract

Ethical aspects of management control systems (MCS) are attracting increasing attention among scholars and practitioners. Much of the work centers on their aims. We complement this scholarship by applying the ethical principle of “no harm,” i.e., non-maleficence, to examine how those aims are achieved. We illustrate this approach by exploring the effects of four MCS designs on job-related stress drawing on the differentiation of stress into two dimensions: a challenge (i.e., unproblematic and even desirable) and a threat dimension (i.e., dangerous; causing psychological strain). Results from a lagged field-survey with 471 managers and employees from the UK and the U.S. support key predictions and offer first insights into designing MCS based on a “no harm” ethics. Our study highlights the benefits of interdisciplinary research in business ethics and hopefully encourages more work on MCS from a perspective based on the non-maleficence principle.

**Keywords** Management control · Non-maleficence · Job-related stress · Challenge · Threat

## Introduction

Ethical aspects of management control systems (MCS) increasingly attract attention (e.g., Chabrak et al. 2016; Cugueró-Escofet and Rosanas 2017; Merchant and White 2017), including how MCS can lead (good) people to do bad things (e.g., Jaworski and Young 1992; Schweitzer et al. 2004), or conversely, how MCS should encourage employees’ development (e.g., Rosanas and Velilla 2005). Much of this interest centers on the *aims* of MCS. We suggest complementing this work by considering the effect of MCS design choices on employees’ well-being based on the non-maleficence principle, that is: an ethical imperative focusing on avoiding harm.

Extending and complementing this debate with such an approach is important for two reasons. First, from a theoretical perspective, Chabrak et al. (2016) have established that MCS can result in anomie, stress, futility and isolation among employees and even lead some to take their lives. Yet, so far we know little about how alternative MCS designs found in practice or discussed in literature affect well-being. Second, from a managerial perspective, MCS should foster organizational performance in compliance with the ethical imperative to protect employees’ well-being.

We therefore believe that an approach based on the non-maleficence principle (Beauchamp and Childress 2013) can help advance literature and practice. Hence, we aim to illustrate how research following such an approach can look like and hope to thereby elicit more scholarly work on the role of MCS design choices for employees’ well-being and heighten awareness of ethical design of MCS in business practice. Moreover, from a broader theoretical perspective, we illustrate the value of interdisciplinary research in business ethics by combining moral philosophy with psychology to inform the design of non-maleficent MCS.

The implications of four MCS design choices for job-related stress serve us to illustrate such an approach. Stress has been associated with coronary heart disease, immunosuppression, gastrointestinal disorders, headaches, tremors, and sleep disturbance (e.g., George and Jones 2005;

✉ Stefan Linder  
linder@essec.fr

Bernard Leca  
leca@essec.fr

Adrián Zicari  
zicari@essec.fr

Veronica Casarin  
casarin@essec.fr

<sup>1</sup> ESSEC Business School, 3 Avenue Bernard Hirsch, BP 50105, 95021 Cergy-Pontoise, France

Stephoe and Kivimäki 2012). Few studies have examined the role of MCS in job-related stress and thus for many MCS design choices we currently do not know how they impact job-related stress. Yet, the limited evidence suggests that MCS affect the stress experienced by managers and employees (e.g., Chabrak et al. 2016; Kenis 1979; Shields et al. 2000). The role of MCS in job-related stress thus seems a salient case for illustrating how taking a well-being perspective based on non-maleficence can look like. Recent advances in psychology that so far have not seen application in MCS research, further enhance the value of using job-related stress for illustrating the approach that we propose. These advances suggest that stress is better understood as two dimensional than unidimensional: as a challenge (i.e., as unproblematic and even desirable) and a threat dimension (i.e., dangerous; causing psychological strain). Thus, advancing our understanding of the effects of alternative MCS designs, both those already studied in the past and those that have not attracted attention yet, when distinguishing threat- from challenge stress can facilitate the design of a more ethical MCS.

We first review literature on the ethical aspects of MCS and explain the principle of non-maleficence (Beauchamp and Childress 2013). Then we illustrate how research applying the principle of non-maleficence could look like by drawing on the case of job-related stress as a multi-dimensional phenomenon and developing predictions for four exemplary MCS design choices. Subsequently, we test our predictions using lagged survey evidence from 471 U.S. and UK employees and managers. We then present some questions and potential dilemmas about how to design more ethical MCS from a non-maleficence perspective that we believe can help guide future research into MCS design.

## Theoretical Background

### The Ethical Aspects of MCS

The growing interest in ethical aspects of MCS (e.g., Chabrak et al. 2016; Cugueró-Escofet and Rosanas 2013, 2017; Jaworski and Young 1992; Schweitzer et al. 2004) has led to the emergence of two streams of research. The first one focuses on the role of management accountants and managers. It views MCS as technical tools whose consequences depend on how they are used. Consequently, it highlights the moral development of management accountants and managers as well as their reactions to moral dilemmas and issues (Ponemon and Gabhart 1993; Etherington and Hill 1998; Fisher and Lovell 2000).

The second approach considers the consequences of design choices in MCS for the behavior of those subjected to the MCS. Langevin and Mendoza (2013) suggest that the

more employees perceive the control system as fair, the less inclined they will be to act unethically. Rosanas and colleagues (Rosanas and Velilla 2005; Cugueró-Escofet and Rosanas 2013, 2017) highlight the importance of justice, and the need for congruence between individual and organizational goals in the design of the MCS and in the use of MCS by managers.

Both streams have greatly enhanced our understanding of ethical aspects of MCS. At the same time, even though MCS design has emerged as an important topic, it so far has mainly been considered a way to encourage ethical behavior. The ethical problem that certain MCS designs may be harmful to managers and/or employees, in contrast, has received little attention so far (a notable exception is Chabrak et al. 2016).

Thus, both theory and practice stand to benefit from better understanding what effects alternative MCS designs have for the well-being of managers and employees. We believe that principlism, namely, the ethical principle of non-maleficence, is a promising starting point for such work.

### Bringing Non-maleficence into the Ethical Discussion on MCS

Non-maleficence obligates one “to abstain from causing harm to others” (Beauchamp and Childress 2013, p. 150). From the principlist perspective, non-maleficence is among the *prima facie* duties in making ethical decisions (Beauchamp and Childress 2013), that is, duties that are morally binding (Ross 1930). The principle of non-maleficence can be found in many types of ethical theory and is sometimes combined with the principle of beneficence into a single principle (Beauchamp and Childress 2013). However, keeping the two separate is helpful: whereas non-maleficence focuses on abstaining from harming someone, beneficence implies “to help others further their important and legitimate interests” (Beauchamp 1999, p. 20) and thus demands to “take positive steps to help others, not merely [to] refrain from harmful acts” (Beauchamp and Childress 2013, p. 202). Non-maleficence and beneficence are two of the four principles of ethical conduct identified by Beauchamp and Childress (2013), along with respect for autonomy and justice.

These principles were first used in bioethics (Beauchamp 1995) and have since been introduced into business ethics (Fisher 2001; Lurie and Albin 2007). More recently, the principles of beneficence and non-maleficence have even been placed at the core of the discipline of well-being marketing (e.g., Sirgy and Lee 2008). Reflecting on what non-maleficence means in a business context, Fisher (2001, p. 21) cites “the duty employers have to provide a safe working environment directly connects to the injunctions not to inflict harm, to prevent harm and to remove harm.”

Even though non-maleficence is among the *prima facie* factors in making ethical decisions and though Beauchamp and Childress (2013) qualify non-maleficence as a basic ethical requirement, we know little about whether or to what extent different MCS design choices comply with it. A few works, however, suggest that MCS do not always fulfill this basic principle, particularly when, as Rosanas and Velilla (2005) state, they are designed to address technical issues or, as reported in Chabrak et al. (2016), they harm employees.

We thus propose to complement extant research on the ethical aspects of MCS by building on the non-maleficence principle and by studying how well MCS design choices comply with this principle. Such a perspective can complement an organizational justice perspective, which considers the controllability or objectivity of MCS (e.g., Bourguignon and Chiapello 2005; Cugueró-Escofet and Rosanas 2017; Rosanas and Velilla 2005). Moreover, it resonates with the growing scholarly interest in an “ethics of care” perspective that calls for more “sensitivity and responsiveness to the feelings, concerns, and particular circumstances of individuals” (Simola 2003, p. 354).

Various forms of harm are conceivable within an organizational context and offer applications for studying how consistent MCS design choices are with the principle of non-maleficence. Chabrak et al. (2016) pointed to anomie, shame, burnout and isolation among employees. Stress merits special attention. The European Union has identified job-related stress as one of the biggest health and safety challenges in Europe (OSHA 2009). We will therefore use job-related stress as our example.

### Job-Related Stress

Stress is a psychological state that arises from a mismatch between perceived demands and one’s ability to meet those demands given available resources (Cox 1993; George and Jones 2005; Lazarus and Folkman 1984). This view underpins the job-demands-resources (JD-R) model (Beehr and Bhagat 1985; Van Harrison 1985), and its latest iteration, the differentiated job-demands-resources (dJD-R) model (Crawford et al. 2010; Van den Broeck et al. 2010). The dJD-R model enhances the basic model with a view of stress as multi-dimensional: as a challenge that fosters work engagement and performance, and as a threat (Crawford et al. 2010).

The threat dimension speaks to the potential for harm; the challenge dimension relates to the potential for growth and mastery. The challenge-related dimension therefore has also been associated with increased motivation and work engagement (Bakker and Schaufeli 2008; Behnke and Kaczmarek 2018). This conceptualization corresponds to recent empirical findings of the two-dimensional nature of stress (e.g., Crawford et al., 2010; Ferguson et al. 1999; Rosen et al.

2010; Van den Broeck et al. 2010) and is consistent with the challenge and threat paradigm, which has become a leading theoretical framework for physiological responses during a motivated performance (Blascovich et al. 2004; Seery et al. 2009; Turner et al. 2012; Behnke and Kaczmarek 2018).

Extant MCS literature testing relationships among MCS designs and stress has thus far not considered this distinction between challenge and threat stress. Integrating these advances into the analysis and debate of the ethical aspects of MCS may assist in designing MCS that increase the challenge but not the threat, or that reduce threat while keeping challenge constant. A more nuanced understanding of how MCS affect the different dimensions of stress can thus inform the design of MCS, which comply with the principle of non-maleficence.

### Hypotheses

Out of the many MCS design choices, four seem particularly promising for illustrating the implications of taking the ethical approach that we propose.

Extant research suggests that the tightness of performance standards and the breadth of Performance Measurement Systems (PMS) affect stress, yet it has neither differentiated the relation for multiple dimensions of stress nor measured the psychological state of stress, but stressors or strains (e.g., Burney and Widener 2007; Hopwood 1972; Kenis 1979; Shields et al. 2000). Thus, shedding more light on these two design choices promises to contribute to theory building and be informative for business practice. This seems particularly important for PMS breadth. Broader PMS—such as the balanced scorecard (Epstein and Manzoni 1997; Kaplan and Norton 1992) or the *tableau de bord* (Epstein and Manzoni 1997, 1998)—have become prevalent in business practice. This seems to fit the fact that many outcomes of tasks and jobs today are multi-dimensional (Osterloh and Frey 2000). Broad PMS can cover as many of these dimensions as wished; narrow PMS focus on a few or only one (Burney and Widener 2007; Hall 2008; Holmström and Milgrom 1991). Yet, literature seems to disagree about the effects on job-related stress: Hopwood (1972) suggested that PMS that are limited to financial performance imply more job-related tensions and subordinate dysfunctional behaviors, because the PMS can be considered a biased indicator of managerial performance. Likewise, Burney and Widener (2007) suggest that more comprehensive PMS reduce role-related stress by lowering role ambiguity. At the same time, some literature associates stress with information overload (Eppler and Mengis 2004; Zhuang et al. 2011) and some researchers have warned of the dangers of information overload due to the inclusion of non-financial indicators in performance measurement (e.g., Neumann et al. 2012) or concluded that broad performance measurement can be

related to information overload (Clark et al. 2006; Chong 1996). These studies report stress, confusion, anxiety, and low motivation (Eppler and Mengis 2004; Zhuang et al. 2011). Part of the disagreement in predictions in the extant literature might be related to which dimension of the stress construct is under examination.

Like broader PMS, pay-for-performance systems are increasingly common (e.g., Gerhart and Fang 2014). Therefore, knowledge about how well these design choices comply with the principle of non-maleficence points to the value that an ethical perspective can have for business practice.

Finally, according to Hopwood (1972), the manner in which accounting data are used to evaluate managers affects their job-related stress. Research in organizational behavior and psychology studying the links between leadership styles and well-being at work points into a similar direction. Nielsen and Munir (2009) find that visionary and creative leadership inspires employees and improves their affective well-being. In the MCS literature, Simons' (1994) distinction between diagnostic and interactive use has gained widespread attention, yet the relation with job-related stress has not yet been studied. Consequently, studying how interactive vs. diagnostic systems and how different remuneration systems perform with respect to the non-maleficence principle seems useful in advancing ongoing theory-building efforts and business practice in these areas.

We therefore focus on standard tightness, the breadth of the PMS used, the interactive nature of these PMS, and whether performance-contingent incentives or seniority wages are used, to illustrate what research drawing on the principle of non-maleficence can look like. In the following, we consequently develop hypotheses for these four design choices and subsequently test them empirically.

### Standard Tightness

Empirical work on standard or budgetary tightness indicates a positive relation between standard tightness and stress (e.g., Kenis 1979; Shields et al. 2000). Whereas the work made no distinction among types of stress, their findings are consistent with the predictions of both the JD-R and its refined version, the dJD-R model. Tighter standards imply that, all else equal, obtaining valued outcomes is more difficult and success more uncertain than under conditions of slack resources. Perceived uncertainty of obtaining outcomes is one key variable driving stress in the JD-R model (Van Harrison 1985). Moreover, since most firms employ annual budgets, tight budgets and employees' resulting uncertainty about attaining valued outcomes are not a short-term situation. This extended duration of the misfit further heightens the stress experienced (e.g., Beehr and Bhagat 1985; Van Harrison 1985). Consequently, based on the JD-R model, tight financial standards should be experienced as

stressful. Workload and time urgency have been identified as two core aspects of a job leading to challenge stress, whereas resource inadequacies and role overload relate to the threat dimension (Crawford et al. 2010). Since all can be considered characteristic of work situations with tight standards, stricter standards should be perceived as increasing both the challenge and the threat dimensions.

**Hypothesis 1** Standard tightness increases the challenge dimension of job-related stress.

**Hypothesis 2** Standard tightness increases the threat dimension of job-related stress.

### Breadth of Performance Measurement

Based on the dJD-R model and the two dimensions of job-related stress, broadening the assessment of performance to comprise both financial and non-financial performance should increase challenge- but reduce threat-related stress. Extending the measures of performance from financial to non-financial implies that more information about an employee's performance is available. A broader PMS thus offers more information for evaluation thereby facilitating an employee's learning and mastery. Given the greater possibility for personal growth implied by a broader PMS and since it signals the importance of performing well on multiple dimensions, it is consistent with fostering a growth mindset (Chao et al. 2017) and should strengthen the challenge dimension. At the same time, a broader PMS demands more information processing capacities to exploit the opportunity that it creates (Clark et al. 2006; Chong 1996; Neumann et al. 2012). In the JD-R/dJD-R model this corresponds to higher challenge stress.

A broader PMS also increases challenge because being evaluated on multiple dimensions makes "taking shortcuts" more difficult (e.g., Holmström and Milgrom 1991). Thus, whereas financial PMS allow an individual to cut corners on quality or customer satisfaction to achieve good financial performance, this is not possible with the inclusion of non-financial indicators (Burney and Widener 2007). A broad PMS thus may require identifying new and better ways of doing one's job to achieve good performance on multiple dimensions. It thus requires personal learning and growth. Within the dJD-R model, this is challenge-related stress.

However, employees tend to perceive performance evaluation as a threat to their self-esteem (Mallinger and Greiner 1981). Non-financial indicators precede (in terms of time) financial ones (Kaplan and Norton 1992). Consequently, a performance assessment that considers such leading indicators should reduce employees' perceived uncertainty about attaining desired outcomes. For example, although sales revenue might indicate that today's performance is mediocre,

customer satisfaction (a non-financial indicator leading sales) might “forecast” an improvement in sales revenue. If an organization takes comprehensive performance measurement seriously, this situation should entail a more favorable evaluation of an employee’s or team’s performance. Thus, a PMS that incorporates both non-financial and financial indicators—and hence, is a less biased assessment of performance (Hopwood 1972)—should lead to less perceived threat. The breadth of the PMS might therefore enhance, to stay in the terminology of the dJD-R model, the attainability of the demands—and decrease perceived threat-related stress.

Moreover, failure on one dimension does not mean failure in all of them. A broad PMS can generate more relevant information and clarify the organization’s expectations than a narrow one can (Burney and Widener 2007). Role ambiguity is associated with threat-related stress (Parker and DeCotiis 1983). Reduced ambiguity about one’s role as well as the organization’s expectations corresponds to an increase in resources in the terminology of the dJD-R model—and should reduce perceived stress. Therefore, and consistent with the dJD-R model, a broader PMS implies less perceived threat-related stress.

Thus, we propose that breadth of the PMS will be positively linked to the challenge dimension of perceived stress but negatively to the threat dimension.

**Hypothesis 3** Breadth of performance measurement systems heightens the challenge dimension of job-related stress.

**Hypothesis 4** Breadth of performance measurement systems reduces the threat dimension of job-related stress.

#### Interactive Use of PMS

According to Simons (1994), an interactive control implies that the information generated receives recurring and close attention from upper management and operational managers at all levels of the organization, that it should be interpreted and discussed face-to-face, and that “the process relies on the continual challenge and debate of underlying data, assumptions, and action plans” (Simons 1987, p. 351n; Simons, 1994). In contrast, a diagnostic PMS would evaluate managers’ performance against predetermined targets (Simons 1994).

Based on the dJD-R model one can expect that an interactively used PMS will tend to lower both dimensions of job-related stress. The involvement of top management in the day-to-day use of performance measures signals its commitment to understand the organization’s situation. Using PMS interactively fosters a shared understanding. Thus, interactive use should reduce the risk of misunderstandings, conflicts among action plans, or insufficient resource allocation.

All of these risks make it less likely that an employee’s performance will satisfy top managers. The reduction of uncertainty about one’s role and the organization’s expectations corresponds to a reduction in hindrance demands in the terminology of the dJD-R model—and lower perceived threat-related stress.

Moreover, an organization that interactively uses its PMS invites its members to share competence and knowledge via the exchange among each other. Individuals thus stand to gain new insights—and can do their job better—which within the dJD-R corresponds to an increase in resources. Therefore, we expect that interactive use of a PMS that responds to environmental uncertainty, builds on face-to-face interaction, and that is non-intrusive will lower both the perceived challenge and threat dimensions.

**Hypothesis 5** Interactive use of performance measurement systems reduces the challenge dimension of job-related stress.

**Hypothesis 6** Interactive use of performance measurement systems reduces the threat dimension of job-related stress.

#### Remuneration System

For employees who value money, a financial incentive linked to performance should increase the perceived importance of “good” performance. Higher importance of valued behavioral outcomes leads to greater perceived stress (e.g., Beehr and Bhagat 1985; Van Harrison 1985). Therefore, the choice between a remuneration system where payments are flat (non-performance contingent) and a system with performance-contingent incentives probably affects the stress that individuals experience. From the perspective of non-maleficence, one might ask whether this difference applies only to one or both dimensions of stress.

Within the dJD-R model, performance-contingent remuneration schemes should heighten the perceived challenge. Such a remuneration scheme offers the opportunity for gain, but it can also be a source of information for personal growth. They signal the importance of improving one’s work via learning and mastery—and thus performance-contingent remuneration is consistent with a growth mindset (Chao et al. 2017). Notably, it offers the individual information for better judging her actions and achievements that she would not necessarily have in fixed salary setting. For example, pay-for-performance schemes require a definition of the reward basis (e.g., sales achieved, cost savings attained), a formula to calculate the size of the payment, and a yardstick to compare actual performance. This information can be helpful in evaluating one’s actions, putting the performance achieved in perspective and fostering a growth perspective (Ashford 1986; Shields et al. 2000). Performance-contingent

remuneration thus speaks to the potential for growth and the potential for gain (Ferguson et al. 1999). Consequently, as George and Jones (2005) noted, offering bonuses for high performance pushes employees to benefit from the opportunity.

Besides heightening the challenge dimension of stress, performance-related remuneration systems may affect threat-related stress. As Locke and Taylor (1991) note, a career can allow the pursuit of material, achievement-related, and social values; potential obstacles to these values are perceived as a threat to one's physical well-being or self-esteem, and work becomes stressful. Especially if an employee's income is performance-contingent, his or her material well-being depends on attaining and maintaining good performance. In such a case, not attaining good performance and, thus not earning more money, implies a risk of monetary loss—and subsequently a potential for material harm. A performance-contingent remuneration system may not only be perceived as an opportunity for gain in such a setting, but might be seen as complementing the base salary that is necessary to maintain one's lifestyle. Not getting the same salary then translates into a decline in lifestyle and harm to one's material well-being.<sup>1</sup>

Thus, based on the dJD-R model, the choice of remuneration system should affect both the perceived challenge and the threat-related stress.

**Hypothesis 7** Pay-for-performance schemes rewarding good performance with a bonus heighten the challenge of job-related stress.

**Hypothesis 8** Pay-for-performance schemes rewarding good performance with a bonus increase the threat dimension of job-related stress.

## Methods

### Design and Data

For testing the hypothesized relationships, we draw on data from U.S. and UK managers and employees recruited with the assistance of Prolific Academic, a startup in the Incubator of the University of Oxford. Prolific Academic provides high-quality professional services for collecting data via an online labor market. It is particularly suited for academic social and economic research (e.g., Marreiros et al. 2017) as it meets the standards of good online surveys (Plan and Schitter 2018) by allowing researchers to contact

respondents multiple times to conduct lagged or longitudinal analyses and by offering pre-screening options to define a sample (Peer et al. 2017). This allowed us to restrict the sample to participants living in the UK or the U.S. and working in for-profit firms. Such a sample facilitates the use of extant measurement instruments for our constructs, which is recommended in literature (Podsakoff et al. 2003).

To minimize the risk of common method bias tainting our findings, we used procedural and statistical means (e.g., Podsakoff et al. 2003). Procedural remedies comprised collecting the dependent variable six weeks after the independent, the use of a cover story that avoided a link to job-related stress, relying on existing scales, ensuring respondents' anonymity, and including the scales in a larger set of scales on diverse topics. As a statistical means, we rely on Harman's one-factor test on all items of the independent and dependent variables. Factor analysis identified seven factors with eigenvalues above 1, each explaining between 5 and 19% of total variance. Common method variance should not taint our results (Podsakoff et al. 2003).

To identify and remove careless responses, we tracked the time respondents took to complete the survey and placed instructed response items at some points (Meade and Craig 2012). A total of 630 (out of eligible 1200) individuals participated in the initial data collection to gain information about our independent variables (firms' MCS) and participants' characteristics. After removing careless and incomplete responses, six weeks later we contacted 596 respondents who had provided usable responses to collect data on our dependent variable (job-related stress). Of these, 514 responded. Yet, one person responded by email that she had been hospitalized and currently is not be able to fill-in the survey, ten did not complete the survey, and nine failed to meet the response quality criteria. Moreover, out of the 494 remaining, 23 individuals had changed jobs or failed to answer that question so their data had to be excluded. This left us with 471 observations.

To test whether sample attrition might bias the results, we followed Goodman and Blum (1996) and set a binary variable equal to 1 if participants responded to the second data collection (stayers) and equal to 0 if participants only responded at time 1 (leavers). A logistic regression of this binary variable on respondents' gender, age, tenure, whether the individual is a manager or not, U.S. vs. UK residency, and firm size suggested that none of these variables significantly predicts whether participants stayed in or dropped from the study at time 2. Consequently, attrition should not taint our results.

<sup>1</sup> We thank one of the reviewers for pointing us to variable compensation systems offering potential for gain and for harm.

## Measures

To allow for easy interpretation when testing for potential interaction effects, all MCS measures were mean-centered and standardized.

### Standard Tightness

Drawing on Shields et al. (2000), we asked respondents for each one of four types of resources to indicate the amount available to them and the amount of these resources needed as perceived by them (see "Appendix 1"). As in Shields et al. (2000), responses for each of the four types of resources available were subtracted from the corresponding resources needed as perceived by the respondent, and the differences then summed to form a single index. Higher positive values denote tighter standards.

### Breadth of PMS

Assessment of the breadth of the PMS relied on a multi-item scale by Homburg et al. (2012), which asks respondents to state their agreement with four statements about the characteristics of their firm's PMS (see "Appendix 1"). Cronbach's alpha of the 4-item scale in our sample was 0.78. The items were aggregated into a summated rating scale, with positive ratings corresponding to higher than average total PMS breadth.

### Interactive use of PMS

One of the most often used instruments to assess interactive use of an MCS is Abernethy and Brownell's (1999) classification task, designed for budgets. This one-item scale asks respondents to indicate which of two alternative uses of budgeting better represents the situation at their organization. Even though developed to measure different uses of budgets, the scale adapts itself well to capturing differences in use of PMS. In this modified form (*PMS Inter. use (AB)*) respondents are asked to rate the extent to which their firms use their PMS in an interactive manner in a scale ranging from 1 (interactive use) to 6 (diagnostic use) as shown in "Appendix 1".

### Remuneration Scheme

For assessing whether firms rely on fixed salaries or a remuneration scheme that comprises a performance-based component, we drew on Abernethy et al. (2015). Their one-item scale asks respondents to indicate the approximate percentage of total compensation that is performance based (see "Appendix 1").

## Stress

Measurement of challenge- and threat-related stress relied on scales by Drach-Zahavy and Erez (2002). Unlike many others, their scales answer Cox's (1985) call to measure stress as a self-reported state and avoid confounding antecedents, psychological state of being stressed, and consequences. Moreover, in line with Lazarus' (1998, pp. 185–212) distinction and conceptualization of stress processes and the dJD-R model, Drach-Zahavy and Erez's (2002) scale captures perceived challenge as a first dimension of stress and perceived threat as a second. Several scholars have used it in recent years (e.g., Espedido and Searle 2018) and factor analysis supports the two-dimensionality of the measurement instrument in our data. Cronbach's alphas of the two subscales were 0.74 (challenge) and 0.77 (threat) (see "Appendix 1")

## Controls

Extraversion is associated with seeing more opportunities for growth in an environment and thus a more challenge-related appraisal of a situation (Ferguson et al. 1999). In contrast, risk-aversion heightens stress (Cadsby et al. 2016). To control for these factors, we collected participants' risk-aversion based on Dohmen et al.'s (2011) scale (reverse coded to give risk-aversion) and the 10-item TIPI scale by Gosling et al. (2003). The TIPI is a brief measure of the Big-Five personality dimensions that shows adequate convergence with longer measurement instruments and has been proposed for research when personality is not the primary topic of interest or when very short instruments are necessary due to scarcity of respondent time (Gosling et al. 2003).

Additionally, we collected respondents' gender, psychopathy, religiousness, hierarchical position (coded as having or not having a managerial role), country of residence (U.S. residency = 1; UK residency = 0), firm tenure (in years) and firm size (based on number of employees).

Gender seems to matter for how stressful individuals perceive a particular environment (Cox 1993). Individuals scoring high on psychopathic traits are less emotional and less fearful (e.g., Babiak et al. 2010; Boddy 2014), which might imply that they are more stress-resistant. At the same time, psychopathy has been found to be correlated positively with perceived stress (Noser et al. 2014) in a study not differentiating challenge- from threat-related stress. Thus, we cannot judge ex post what role psychopathic traits play with respect to how our respondents experience different MCS designs—and, hence, need to control for it. Measurement of psychopathy relied on the short 4-item psychopathy subscale (Cronbach's alpha 0.70) in Jonason and Webster (2010); religiousness was captured using a single Likert-type question, asking respondents about "How religious would you describe yourself?" and measurement of respondents' need

**Table 1** Descriptive statistics

Variable	Obs	Mean	SD	Min	Max
Challenge	471	4.70	1.15	1.00	7.00
Threat	471	3.13	1.31	1.00	7.00
Std. tightness	471	-0.00	1.00	-3.42	3.27
PMS breadth	471	-0.00	1.00	-3.16	1.94
PMS Inter. use(AB) <sup>a</sup>	471	0.00	1.00	-1.91	1.69
PMS Inter. use(BO)	471	-0.00	1.00	-2.01	1.87
Percentage PfP	471	0.00	1.00	-0.70	3.70
Perf-based incentives	471	-0.00	1.00	-1.59	2.42
Age (categorical)	471	3.70	1.10	2.00	7.00
Female	471	0.58	0.49	0.00	1.00
Managerial position (N/Y)	471	0.44	0.50	0.00	1.00
Tenure	471	6.37	5.53	0.00	31.00
U.S. residence	471	0.25	0.43	0.00	1.00
Firm size (categorical)	471	16.40	9.93	1.00	26.00
Risk-aversion	471	5.27	2.31	1.00	10.00
Religiousness	471	1.69	1.12	1.00	5.00
Need for achievement	471	5.68	1.21	1.00	7.00
Psychopathy	471	2.77	1.25	1.00	7.00
Extraversion	471	3.78	1.54	1.00	7.00
Agreeableness	471	5.07	1.14	1.50	7.00
Conscientiousness	471	5.44	1.10	1.50	7.00
Emotional stability	471	4.63	1.44	1.00	7.00
Openness to experience	471	4.89	1.20	1.00	7.00

<sup>a</sup>PMS Inter. use(AB) refers to the scale based on Abernethy and Brownell (1999), whereas PMS Inter. use(BO) to the scale drawing on items based on Bisbe and Otley (2004)

for achievement relied on Yamaguchi's (2003) four-item measure (Cronbach's alpha 0.87).

Finally, we asked respondents whether they had changed job between the collection of the independent variables (MCS) and dependent variable (stress).

**Table 2** Correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Challenge	1.00							
(2) Threat	-0.09*	1.00						
(3) PMS breadth	0.28***	-0.16***	1.00					
(4) PMS Inter. use (AB) <sup>a</sup>	0.11**	-0.11**	0.22***	1.00				
(5) PMS Inter. use (BO)	0.18***	-0.09**	0.25***	0.11**	1.00			
(6) Std. tightness	-0.09*	0.18***	-0.08*	-0.11**	0.04	1.00		
(7) Percentage PfP	0.22***	-0.05	0.18***	0.06	0.19***	-0.03	1.00	
(8) Perf-based incentives	0.26***	-0.02	0.31***	0.05	0.26***	-0.06	0.47***	1.00

\* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$

<sup>a</sup>PMS Inter. use (AB) refers to the scale based on Abernethy and Brownell (1999); PMS Inter. use (BO) refers to the scale drawing on Bisbe and Otley (2004)

## Results

### Descriptive Results

Our respondents have an average of 6 years of tenure and 44% have a managerial position. About half of them are female. One-quarter are U.S. residents, the others live in the UK. Table 1 provides detailed descriptive statistics.

Pair-wise correlations—as shown in Table 2—suggest that challenge stress is significantly correlated with all four MCS choices. Threat is correlated with PMS breadth, interactive use, and standard tightness—but not the remuneration scheme. Moreover, the coefficient signs differ between challenge and threat stress for PMS breadth, interactive use, and standard tightness.

The correlations thus support our belief that it is worthwhile to study the effect of MCS on stress from a non-maleficence perspective. Some MCS design choices likely increase, others decrease threat-related stress, and others may not affect it.

### Multivariate Results

To test our eight hypothesized relations, we constructed a set of regression models. As models 1 and 3 in Table 3 show, some individual characteristics seem important for explaining challenge and/or threat stress.

Respondents having a managerial position reported higher values on both stress dimensions; extraverted individuals seem to experience more challenge, but less threat-related stress. Religiousness and psychopathy seem to heighten perceived threat- but not challenge-related stress at work. Similarly, conscientiousness and emotional stability seem to affect only the threat dimension. Women seem to perceive less threat-related stress than men; employees in larger firms seem to experience more threat than respondents from smaller firms.

Models 2 and 4 in Table 3 speak to our hypotheses. Tight standards show a statistically significant relationship only



**Table 3** Results

Variables	(1)	(2)	(3)	(4)
	Challenge	Challenge	Threat	Threat
Std. tightness		-0.07 (0.12)		0.13*** (0.00)
PMS breadth		0.20*** (0.00)		-0.12*** (0.00)
PMS Inter. use (AB)		0.03 (0.55)		-0.04 (0.39)
Percentage PfP		0.15*** (0.00)		0.03 (0.53)
Age (categorical)	-0.02 (0.74)	0.00 (0.92)	-0.06 (0.23)	-0.08* (0.09)
Female	-0.02 (0.74)	-0.03 (0.57)	-0.09** (0.05)	-0.09* (0.05)
Managerial position (N/Y)	0.14*** (0.00)	0.14*** (0.00)	0.09** (0.05)	0.08* (0.07)
Tenure	0.01 (0.82)	0.01 (0.79)	-0.06 (0.14)	-0.06 (0.14)
U.S. residence	0.04 (0.41)	0.02 (0.70)	0.00 (0.97)	0.01 (0.84)
Firm size (categorical)	0.06 (0.21)	0.04 (0.39)	0.13*** (0.00)	0.13*** (0.00)
Risk-aversion	-0.03 (0.50)	-0.01 (0.88)	0.09* (0.07)	0.10** (0.04)
Religiosity	0.00 (0.98)	-0.02 (0.60)	0.10** (0.02)	0.10** (0.02)
Need for achievement	0.07 (0.15)	0.09** (0.04)	-0.00 (0.99)	0.00 (0.95)
Psychopathy	0.04 (0.42)	0.01 (0.79)	0.15*** (0.00)	0.16*** (0.00)
Extraversion	0.10** (0.04)	0.09* (0.08)	-0.18*** (0.00)	-0.18*** (0.00)
Agreeableness	0.08 (0.18)	0.04 (0.43)	-0.00 (0.98)	0.01 (0.79)
Conscientiousness	0.03 (0.57)	-0.00 (0.97)	-0.25*** (0.00)	-0.23*** (0.00)
Emotional stability	0.05 (0.41)	0.03 (0.64)	-0.22*** (0.00)	-0.21*** (0.00)
Openness to experience	0.06 (0.31)	0.05 (0.31)	0.02 (0.62)	0.03 (0.53)
Observations	471	471	471	471
$R^2$	0.0978	0.175	0.288	0.324
Adjusted $R^2$	0.0681	0.140	0.265	0.296
Incremental $F$ -test		9.140		5.624
Prob > $F$		0.000		0.000

Multiple regressions with robust standard errors; PMS Inter. use(AB) refers to scale based on Abernethy and Brownell (1999); standardized betas; Robust pval in parentheses

\*\*\* $p < 0.01$ , \*\* $p < 0.1$ , \* $p < 0.05$

with the threat dimension of stress in the expected direction, not with challenge. The evidence thus suggests rejecting hypothesis 1, but is consistent with hypothesis 2, which states that standard tightness drives the threat dimensions

of stress. In contrast, in line with hypotheses 3 and 4, PMS breadth increases challenge-, but lowers threat-related stress. Both hypotheses thus are supported by our data. Using PMS interactively does not seem to affect either dimension,

which suggests rejecting hypotheses 5 and 6. Differences in the remuneration system seem to affect only the challenge dimension. Performance-contingent rewards increase—as per hypothesis 7—the challenge but not the threat dimension of stress. The evidence thus does not support hypothesis 8, yet is consistent with hypotheses 2, 3, 4 and 7. Four of the eight hypotheses are therefore consistent with the empirical results.

### Additional Analyses

To test the robustness of our findings and shed additional light on the role of MCS from a non-maleficence perspective, we ran models that allow for a non-linear relation between MCS and job-related stress by incorporating the squared terms of our four MCS variables in the regressions. Results (available from the authors) do not lead to different conclusions about our eight hypotheses, from the results presented earlier. This additional evidence thus is consistent with Hypotheses 2, 3, 4 and 7.

Subsequently, we replicated the regression models using alternative measurement instruments for capturing differences in the use of the PMS and the design of the remuneration system. Whereas both instruments underlying the results reported in Table 3 are established ones, they draw only on a single item for capturing fairly complex constructs. Therefore, we draw on Shields et al.'s (2000) 3-item instrument (alpha of 0.85) for testing the robustness of our findings regarding the relationship between the remuneration system and job-related stress. For interactive use of the PMS, we rely on Bisbe and Otley's (2004) instrument, adjusted so to focus on the interactive use the PMS (Cronbach's alpha being 0.70). Both instruments were collected at time 1 alongside the instruments based on Abernethy and Brownell (1999) and Abernethy et al. (2015) used in Table 3. Items were aggregated to summated rating scales and both variables were mean-centered and standardized. Results—shown in columns 2 and 5 of Table 4—for our eight hypotheses do not differ when using these alternative measures. The same holds true when using the measure of Abernethy and Brownell (1999) for interactive use in a model with the alternative measure of the remuneration system by Shields et al. (2000) or when relying on the alternative assessment of interactive use based on Bisbe and Otley (2004) in a model with the scale by Abernethy et al. (2015) for the remuneration system (results available from the authors).

This additional evidence is thus consistent with hypotheses 2, 3, 4 and 7, but suggests rejecting hypotheses 1, 5, 6 and 8. Support or rejection of our hypotheses thus does not depend on what instrument is used for capturing interactive use and the remuneration system.

As a further test of the robustness, we ran models allowing for interaction effects of the four MCS and job-related stress.

For these models, Table 4, columns 3 and 6, show the results when assessing interactive use of the PMS by means of the Bisbe and Otley (2004) instrument and the remuneration system based on Shields et al. (2000). All but one interaction effect are statistically non-significant. The only interaction that attains borderline significance ( $p < 0.10$ ) is between PMS breadth and interactive use. It points to a potential increase of challenge- and reduction of threat-related stress if PMS are broad and used in an interactive manner. However, results depend on the use of the measure by Bisbe and Otley (2004) of interactive use; the coefficient fails (by and large) to attain statistical significance when the instrument based on Abernethy and Brownell (1999) is used instead. Moreover, the direct effects remain unchanged when allowing for interactions. Hence, the conclusions for our eight hypotheses do not differ when considering possible interactions.

Additionally, we regressed the four MCS on a measure of emotional exhaustion (burnout) from the Maslach and Jackson (1981) inventory collected at time 2. Emotional exhaustion is deleterious to well-being. Results (available from the authors) show that standard tightness exhibits a statistically significant positive relation (at  $p < 0.01$ ) with emotional exhaustion. This indicates that tight standards violate the principle of non-maleficence. In contrast, PMS breadth is negatively related to emotional exhaustion ( $p < 0.05$ ). In sum, the evidence is consistent with the expectation that PMS breadth, unlike increasing standard tightness, reduces harm irrespective of whether one looks at threat stress or emotional exhaustion.

### Discussion

Four out of eight hypotheses are consistent with our results. Standard tightness seems to increase threat stress; PMS breadth reduces threat-related stress and increases challenge stress. And as predicted, pay-for-performance schemes increase the challenge dimension of stress. In contrast, standard tightness seems not to affect the challenge dimension, and performance-contingent remuneration seems not to heighten the threat dimension. Thus, hypotheses 1 and 8 cannot be supported. Furthermore, using PMS interactively does not show direct effects on either dimension—suggesting to reject hypotheses 5 and 6.

Even though our data do not support all hypotheses, they show that some MCS designs are more likely to fulfill the criterion of non-maleficence than others. The recent emphasis on complementing financial measures with non-financial indicators, for example within a balanced scorecard (Epstein and Manzoni 1997; Kaplan and Norton 1992), is likely to reduce harmful effects of PMS on employees' well-being in terms of threat-related stress and emotional exhaustion. Our findings, hence, lend additional reason from

**Table 4** Robustness

Variables	(1) Challenge	(2) Challenge	(3) Challenge	(4) Threat	(5) Threat	(6) Threat
Std. tightness		-0.08 (0.12)	-0.07 (0.16)		0.14*** (0.00)	0.13*** (0.00)
PMS breadth		0.17*** (0.00)	0.18*** (0.00)		-0.12*** (0.00)	-0.13*** (0.00)
PMS Inter. use (BO)		0.06 (0.19)	0.06 (0.21)		-0.05 (0.26)	-0.05 (0.25)
Perf-based incentives		0.15*** (0.00)	0.15*** (0.00)		0.04 (0.37)	0.04 (0.38)
Std. tightness × Inter. use (BO)			0.05 (0.35)			0.01 (0.91)
Std. tightness × Perf-based incentives			-0.04 (0.46)			0.05 (0.26)
PMS breadth × Inter. use (BO)			0.09* (0.06)			-0.07* (0.10)
PMS breadth × Std. tightness			-0.05 (0.33)			-0.05 (0.33)
PMS breadth × Perf-based incentives			0.01 (0.78)			-0.05 (0.26)
Inter. use(BO) × Perf-based incentives			0.01 (0.92)			0.05 (0.39)
Age (categorical)	-0.02 (0.74)	0.00 (0.97)	-0.00 (0.99)	-0.06 (0.23)	-0.07 (0.12)	-0.06 (0.16)
Female	-0.02 (0.74)	-0.03 (0.48)	-0.04 (0.36)	-0.09** (0.05)	-0.09** (0.04)	-0.08* (0.07)
Managerial position (N/Y)	0.14*** (0.00)	0.12*** (0.01)	0.12*** (0.01)	0.09** (0.05)	0.07 (0.10)	0.08* (0.08)
Tenure	0.01 (0.82)	0.02 (0.66)	0.02 (0.75)	-0.06 (0.14)	-0.07 (0.13)	-0.07* (0.09)
U.S. residence	0.04 (0.41)	0.03 (0.59)	0.03 (0.50)	0.00 (0.97)	0.01 (0.88)	-0.00 (0.95)
Firm Size (categorical)	0.06 (0.21)	0.03 (0.52)	0.02 (0.60)	0.13*** (0.00)	0.14*** (0.00)	0.14*** (0.00)
Risk-aversion	-0.03 (0.50)	-0.01 (0.79)	-0.02 (0.65)	0.09* (0.07)	0.10** (0.03)	0.10** (0.03)
Religiosity	0.00 (0.98)	-0.03 (0.50)	-0.02 (0.56)	0.10** (0.02)	0.10** (0.02)	0.09** (0.03)
Need for achievement	0.07 (0.15)	0.08* (0.07)	0.08* (0.08)	-0.00 (0.99)	0.00 (0.93)	0.01 (0.86)
Psychopathy	0.04 (0.42)	0.00 (0.94)	0.00 (0.98)	0.15*** (0.00)	0.16*** (0.00)	0.16*** (0.00)
Extraversion	0.10** (0.04)	0.08 (0.13)	0.07 (0.16)	-0.18*** (0.00)	-0.17*** (0.00)	-0.16*** (0.00)
Agreeableness	0.08 (0.18)	0.06 (0.26)	0.06 (0.31)	-0.00 (0.98)	0.01 (0.79)	0.02 (0.66)
Conscientiousness	0.03 (0.57)	-0.00 (0.99)	-0.00 (0.94)	-0.25*** (0.00)	-0.23*** (0.00)	-0.22*** (0.00)
Emotional stability	0.05 (0.41)	0.03 (0.53)	0.04 (0.50)	-0.22*** (0.00)	-0.21*** (0.00)	-0.22*** (0.00)
Openness to experience	0.06 (0.31)	0.04 (0.43)	0.04 (0.46)	0.02 (0.62)	0.03 (0.47)	0.03 (0.53)
Observations	471	471	471	471	471	471
R <sup>2</sup>	0.0978	0.180	0.194	0.288	0.325	0.336
Adjusted R <sup>2</sup>	0.0681	0.145	0.149	0.265	0.297	0.299
Incremental F-test		8.345	0.992		5.850	0.720
Prob > F		0.000	0.430		0.000	0.634

Multiple regressions with robust standard errors drawing on Shields et al. (2000) measure of remuneration (Perf-based incentives) and Bisbe and Otley's (2004) measure of interactive use (PMS Inter. use (BO)); test of interactions; standardized betas reported; Robust pval in parentheses

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$



**Table 5** Summary of findings and some tentative propositions for future research

MCS design choices	Impact on stress				Tentative propositions for the design of non-maleficent MCS
	Challenge		Threat		
	Prediction	Result	Prediction	Result	
Standard tightness	H1: increase	No influence	H2: increase	Increase	Proposition 1: ethical MCS that care about employees' well-being would need to de-emphasize standard tightness
Breadth of PMS	H3: increase	Increase	H4: decrease	Decrease	Proposition 2: ethical MCS that care about employees' well-being would need to have broad PMS
Interactive PMS	H5: decrease	No influence	H6: decrease	No influence	Proposition 3: the interactive (versus diagnostic) use of PMS does not pose harm to employees' well-being
Pay-for-performance scheme	H7: increase	Increase	H8: increase	No influence	Proposition 4: pay-for-performance scheme does not pose harm to employees' well-being

a non-maleficence perspective to efforts that aim at complementing financial PMS with non-financial indicators. In contrast, tightening performance standards seems—based on its impact on threat-related stress and on emotional exhaustion—problematic from a non-maleficence perspective.

Chabrak et al. (2016) explain how the stress of financialized MCS in a telecommunications company can drive some employees to suicide. Our results point in a similar direction: they show that *some* MCS designs may harm employees and managers. They thus underscore that the MCS design from a non-maleficence perspective that we suggest merits more attention in literature and practice.

Moreover, our findings raise a number of questions for MCS design from a non-maleficence perspective: (1) would an ethical design of MCS require all MCS to increase the challenge dimension of stress, while reducing the threat dimension? Could a firm select a design in which some MCS increase both dimensions, while others decrease both, because it is less costly to implement? (2) Is there an ethical dilemma when the same MCS influences the challenge and the threat simultaneously? Firms in competitive industries need to maintain employees' high work engagement and effort. Since challenge stress is associated with higher work effort and work engagement, firms may feel pressured to intensify employees' challenge stress. Hence, what does a non-maleficence approach imply if an organization wants to foster the challenge dimension to avoid being wiped out by competitors and prevent job-losses, but the MCS design would increase employees' perceived threat dimension, too? Table 5 summarizes our findings and some preliminary propositions. We believe that they are promising avenues for advancing our knowledge and developing recommendations for practice.

More work corroborating and extending our findings thus seems necessary to develop a sound basis for recommendations for more ethical MCS designs and to help answer these questions. Moreover, several characteristics of our empirical study call for research to refine our conclusions.

We focused on four MCS design choices as our aim with this manuscript is not to offer a comprehensive assessment of MCS with respect to their effects on stress, but merely to demonstrate that research based on the principle of non-maleficence can advance the discussion on more ethical MCS. Yet, reality is more complex than the four design choices. For example, remuneration schemes differ along more than just the dimension of whether or not they include a performance-contingent component. Therefore, future research should study the effects of different designs of remuneration systems. Similarly, other MCS, such as budgets or planning and goal setting processes merit study. For example, the rich work on participative target setting suggests that participation matters for clarifying expectations and that less ambiguity may reduce stress (see Merchant and Van der Stede 2012, pp. 317–318). Our study does not account for this—which may have contributed to the modest variance explained in our models. Work identifying whether participative standard setting affects one or both dimensions of stress, such as the challenge dimension, would paint a more comprehensive picture of MCS design from an ethical perspective based on non-maleficence. More research on other MCS design choices is thus merited, especially as the modest R<sup>2</sup> of our models suggests that additional factors must matter for challenge- and threat-related stress.

Most extant research interested in interactive uses of PMS relies on only a single measurement instrument for capturing this construct. Our study tested the relation of interactive use of PMS with job-related stress drawing on two measurement instruments and we do not find different results for the direct effects on job-related stress when replacing Abernethy and Brownell's (1999) scale with Bisbe and Otley's (2004). Yet, it is rather surprising that the way PMS are used should not matter for the level of challenge and threat-related stress. Moreover, we find a borderline significant interaction effect of PMS breadth and interactive use when the latter is measured based on Bisbe and Otley (2004), but not on Abernethy and Brownell (1999).

This raises the question which of the two measures more accurately captures changes in the way the PMS is used. Future research should rely on other instruments for capturing interactive use (e.g., Henri 2006; Widener 2007; Linder and Torp 2017) and allow for a triangulation. Such work would permit to better ascertain what role interactive use plays for either of the two types of job-related stress. Notably, it would allow assessing whether the significant interaction found when using the measure based on Bisbe and Otley (2004) between PMS breadth and interactive use is just an artifact of the one particular instrument used for assessing interactive use or indicative of a substantive need for research from a non-maleficence perspective to look beyond the effects of individual design choices for well-being and factor in combinations of choices.

We control for whether respondents live in the U.S. or the UK. However, assessing the generalizability of our findings and the existence and importance of the dilemmas described earlier will require that future research considers the role of cultural differences among countries (e.g., Hofstede 1967). Moreover, we collected the dependent variable six weeks after the independent in order to minimize the risk of common method bias. Most MCS design choices seem unlikely to change over the course of a few weeks. Yet, future research may want to control for changes in MCS design during the data collection period.

Our interest is with illustrating that MCS design choices matter from a non-maleficence perspective. Thus, we only control for individual differences in psychopathy and the BIG 5 personality traits with two short instruments. For example, in contrast to the original NEO PI-R scale consisting of 240 items by Costa and McCrae (1992), the TIPI is an extremely brief measure of the Big-Five personality dimensions and has been proposed for research when personality is not the primary topic of interest or when very short instruments are necessary due to scarcity of respondent time (Gosling et al. 2003). It thus seems well-suited for our purpose and a sample of employees and managers, where long surveys might lead to non-response bias or incomplete responses. It is, however, noteworthy that the emotional stability dimension in the TIPI corresponds to the Neuroticism dimension in the NEO PI-R and it is clear that a ten-item scale cannot capture the complexity of personality traits as well as the NEO PI-R can. Therefore, future research capturing these individual differences with more elaborate instruments and focusing more in detail on the role of individual differences and their potential interplay with MCS design choices is warranted.

## Conclusion

Ethical aspects of MCS increasingly are attracting scholarly attention (e.g., Chabrak et al. 2016; Cugueró-Escofet and Rosanas 2017; Merchant and White 2017). As Chabrak

et al. (2016) show, the ethical aspects of MCS merit attention not only because of the harm that MCS can cause by their aims, but also by *how* they pursue those aims. The research and practice of ethical design of MCS, therefore, may benefit from considering the effect of MCS design choices on employees' well-being. Such a perspective may yield valuable insights for MCS design by advancing our knowledge about whether or to what extent different MCS design choices comply with ethical principles such as non-maleficence.

To illustrate and hopefully encourage other scholars to follow such an approach, we drew on the dJD-R model from stress research and studied the effects of four MCS design choices on challenge- and threat-related stress within a sample of managers and employees from the UK and the U.S. Results support the expectation that *some* MCS designs are likely to be harmful to well-being. Notably, tight standards seem to increase threat-related stress and emotional exhaustion. Moreover, some design choices may comply better with the principle of non-maleficence than others. Broader PMS seem to be a case in point by lowering threat-related stress and emotional exhaustion.

Whereas our empirical study looks only at four MCS design choices and is limited to empirical data collected within an Anglo-Saxon context, it shows that more work on MCS design choices from a no-harm perspective is warranted and will hopefully stimulate more such investigations. Moreover, our study resonates with the growing scholarly interest in an "ethics of care" perspective applied to organization research (e.g., Molterer et al. 2019). Additionally, we point to various open questions and dilemmas when considering MCS from the vantage point of the principle of non-maleficence. We thereby contribute to the debate about MCS design in the light of the tension between the need to spur high performance to ensure firm survival and ethical imperatives and considerations.

Finally, our work combines insights from principlism with insights from psychology and demonstrates the benefits of such interdisciplinary approaches for advancing our understanding of ethical aspects of MCS. We hope that this inspires more interdisciplinary research on ethical MCS design, and that, our study thereby contributes—in line with Spence and Rutherford (2003)—to business ethics research initiating and developing interdisciplinary studies and bridging separate research traditions.

## Compliance with Ethical Standards

**Ethical Approval** All procedures performed complied with the ethical standards of the institutional research committee and with the 1964 Helsinki Declaration and its later amendments (World Medical Association 2013).

## Appendix

Construct	Item	Text	
Job-related stress		How do you perceive your current job?	Likert 1–7 (1 = strongly disagree; 7 = strongly agree)
	Drach-Zahavy and Erez (2002)_1	My work seems like a challenge to me	
	Drach-Zahavy and Erez (2002)_3	My job provides opportunities to overcome obstacles	
	Drach-Zahavy and Erez (2002)_6	My job provides opportunities to exercise reasoning skills	
	Drach-Zahavy and Erez (2002)_8	My job provides opportunities to strengthen my self-esteem	
	Drach-Zahavy and Erez (2002)_2	I'm worried that my work might reveal my weaknesses	
	Drach-Zahavy and Erez (2002)_4	I am worried that my work might threaten my self-esteem	
	Drach-Zahavy and Erez (2002)_5	My work seems like a threat to me	
Standard tightness	Drach-Zahavy and Erez (2002)_7	The work I do seems long and tiresome	Likert 1–7 (1 = No resources; 7 = An extremely large quantity of resources)
		Please indicate the total amounts of each of the following resources you believed were required for you to achieve your performance standards in the past 12 months [minus the response to (for each item) to the following question]:	
		Please indicate the total amounts of each of the following resources that were made available to you in the past 12 months	
	Shields et al. (2000)_time	Your time at work	
	Shields et al. (2000)_technology	Technology (e.g., CAD)	
	Shields et al. (2000)_intern	Assistance from other employees of your company (e.g., time, skills, information)	
Shields et al. (2000)_extern	Assistance from external sources (e.g., consultants, suppliers, customers)		
PMS breadth		The performance measurement system in the company that I work for ...	Likert 1–7 (1 = strongly disagree; 7 = strongly agree)
	Homburg et al. (2012)_1	... provides financial and non-financial measures	
	Homburg et al. (2012)_2	... provides a balanced picture of the function that I work in	
	Homburg et al. (2012)_3	... provides measures of different perspectives (e.g., financial, customer, competitive, innovation, internal processes)	
	Homburg et al. (2012)_4	... provides besides results-oriented measures (e.g., sales, customer satisfaction) also input-oriented indicators (e.g., amount of time spent, duration of a process)	
Interactive use			

Construct	Item	Text	
	PMS Inter. use (AB), i.e. Abernethy and Brownell (1999)	The information generated by the performance measurement system is an important and recurring agenda addressed by the highest level of management. The performance measurement process demands frequent and regular attention from managers at all levels of the organization and the information provided by the system is interpreted and discussed in face-to-face meetings with subordinates and peers. The performance measurement process relies on the continual challenge and debate of underlying data, assumptions and action plans.[vs.] The performance measurement system aims at achieving predetermined outcomes and the information produced by the system is used primarily to inform top managers if actions or outcomes are not in accordance with plans. Staff specialists (i.e., finance departments) play a pivotal role in preparing and interpreting the information produced by the system. Data are reported through formal reporting procedures and top managers tend to be involved in the process infrequently and on an exception basis	Likert 1–6 (1 = left statement more; 6 = right statement more); reverse coded
	PMS Inter. use (BO)	Please select the one statement of each pair which you more strongly believe to be the case as far as you are concerned. Please try to respond to each pair independently when making your choice; do not be influenced by your previous choices	Likert 1–6 (1 = left statement more; 6 = right statement more)
	Bisbe and Otley (2004)_1	Only when there are deviations from planned performance are performance indicators the main subject for face-to-face discussions with my superior.[vs.] Whether there are deviations from planned performance or not, performance indicators are the main subject for face-to-face discussion with my superior	
	Bisbe and Otley (2004)_2	My superior pays periodic or occasional attention to performance indicators.[vs.] My superior pays regular and frequent attention to performance indicators. I use them permanently	
	Bisbe and Otley (2004)_3	For many managers in my company, performance measurement systems require periodic or occasional attention, but not permanent attention.[vs.] In my company, performance measurement requires permanent attention from managers at all levels	
Remuneration system	Abernethy et al. (2015)	Please indicate what percentage of your total compensation is performance-based (e.g., commission, bonus etc.)	Likert 1–9 (1 = 0%; 9 = above 70%)
	Shields et al. (2000)_1	How large are the following aspects in the company that you work for The degree to which valued rewards to you increase with increases in your measured performance	Likert 1–7 (1 = extremely low; 7 = extremely high)
	Shields et al. (2000)_2	The degree to which your rewards are totally determined by measured performance relative to performance standards	
	Shields et al. (2000)_3	The extent to which those employees that are in the top 25% in terms of performance standards receive larger valued rewards than do those whose performance in relation to the standards are not in the top 25%	

## References

- Abernethy, M. A., & Brownell, P. (1999). The role of budgets in organizations facing strategic change: An exploratory study. *Accounting, Organizations and Society*, 24(3), 189–204.
- Abernethy, M. A., Dekker, H. C., & Schulz, A. K.-D. (2015). Are employee selection and incentive contracts complements or substitutes? *Journal of Accounting Research*, 53(4), 633–668.
- Ashford, S. J. (1986). Feedback-seeking in individual adaptation: A resource perspective. *Academy of Management Journal*, 29(3), 465–487.
- Babiak, P., Neumann, C. S., & Hare, R. D. (2010). Corporate psychopathy: Talking the walk. *Behavioral Sciences and the Law*, 28, 174–193.
- Bakker, A. B., & Schaufeli, W. B. (2008). Positive organizational behavior: Engaged employees in flourishing organizations. *Journal of Organizational Behavior*, 29(2), 147–154.
- Beauchamp, T. L. (1995). Principlism and its alleged competitors. *Kennedy Institute of Ethics Journal*, 5(3), 181–198.
- Beauchamp, T. L. (Ed.). (1999). *Enquiry concerning human understanding*. Oxford Philosophical Text. Oxford: Oxford University Press.
- Beauchamp, T. L., & Childress, J. F. (2013). *Principles of biomedical ethics* (7th ed.). New York: Oxford University Press.
- Beehr, T. A., & Bhagat, R. S. (1985). *Human stress and cognition in organizations: An integrated perspective*. New York: Wiley-Interscience.
- Behnke, M., & Kaczmarek, L. D. (2018). Successful performance and cardiovascular markers of challenge and threat: A meta-analysis. *International Journal of Psychophysiology*, 130, 73–79.
- Bisbe, J., & Otley, D. (2004). The effects of the interactive use of management control systems on product innovation. *Accounting, Organizations and Society*, 29(8), 709–737.
- Blascovich, J., Seery, M. D., Mugridge, C. A., Norris, R. K., & Weisbuch, M. (2004). Predicting athletic performance from cardiovascular indexes of challenge and threat. *Journal of Experimental Social Psychology*, 40(5), 683–688.
- Boddy, C. R. (2014). Corporate psychopaths, conflict, employee affective well-being and counterproductive work behavior. *Journal of Business Ethics*, 121, 107–121.
- Bourguignon, A., & Chiapello, E. (2005). The role of criticism in the dynamics of performance evaluation systems. *Critical Perspectives on Accounting*, 16(6), 665–700.
- Burney, L., & Widener, S. K. (2007). Strategic performance measurement systems, job-relevant information, and managerial behavioral responses—Role stress and performance. *Behavioral Research in Accounting*, 19(1), 43–69.
- Cadsby, C. B., Song, F., & Tapon, F. (2016). The Impact of Risk-Aversion and Stress on the Incentive Effect of Performance Pay. In S. J. George & J. R. Hamman (Eds.), *Experiments in organizational economics (Research in Experimental Economics, Vol 19)* (pp. 189–227). Bingley: Emerald Group Publishing Limited.
- Chabrak, N., Craig, R., & Daidj, N. (2016). Financialization and the employee suicide crisis at France Telecom. *Journal of Business Ethics*, 139(3), 501–515.
- Chao, M. M., Visaria, S., Mukhopadhyay, A., & Dehejia, R. (2017). Do rewards reinforce the growth mindset?: Joint effects of the growth mindset and incentive schemes in a field intervention. *Journal of Experimental Psychology*, 146(10), 1402–1419.
- Chong, V. K. (1996). Management accounting systems, task uncertainty and managerial performance: A research note. *Accounting, Organizations and Society*, 21(5), 415–421.
- Clark, B., Abela, A. V., & Ambler, T. (2006). An information processing model of marketing performance measurement. *Journal of Marketing Theory and Practice*, 14(3), 191–208.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources Inc.
- Cox, T. (1985). The nature and measurement of stress. *Ergonomics*, 28(8), 1155–1163.
- Cox, T. (1993). *Stress research and stress management: putting theory to work*. Nottingham: University of Nottingham.
- Crawford, E. R., LePine, J. A., & Rich, B. L. (2010). Linking job demands and resources to employee engagement and burnout: A theoretical extension and meta-analytic test. *Journal of Applied Psychology*, 95(5), 834–848.
- Cugueró-Escofet, N., & Rosanas, J. M. (2013). The just design and use of management control systems as requirements for goal congruence. *Management Accounting Research*, 24(1), 23–40.
- Cugueró-Escofet, N., & Rosanas, J. M. (2017). The ethics of metrics: Overcoming the dysfunctional effects of performance measurements through justice. *Journal of Business Ethics*, 140(4), 615–631.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522–550.
- Drach-Zahavy, A., & Erez, M. (2002). Challenge versus threat effects on the goal-performance relationship. *Organizational Behavior and Human Decision Processes*, 88(2), 667–682.
- Eppler, M., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *Information Society*, 20(5), 325–344.
- Epstein, M., & Manzoni, J. F. (1998). Implementing corporate strategy: From Tableaux de Bord to Balanced Scorecards. *European Management Journal*, 16(2), 190–203.
- Epstein, M. J., & Manzoni, J. F. (1997). The balanced scorecard and tableau de bord: Translating strategy into action. *Management Accounting*, 79(2), 28–36.
- Espedido, A., & Searle, B. (2018). Goal difficulty and creative performance: The mediating role of stress appraisal. *Human Performance*, 31(3), 179–196.
- Etherington, L. D., & Hill, N. T. (1998). Ethical development of CMAs: A focus on non-public accountants in the United States. In L. Ponemon (Ed.), *Research on accounting ethics* (Vol. 4, pp. 225–245). Stamford CT: JAI Press.
- Ferguson, E., Matthews, G., & Cox, T. (1999). The Appraisal of Life Events (ALE) scale: Reliability and validity. *British Journal of Health Psychology*, 4(2), 97–116.
- Fisher, J. (2001). Lessons for business ethics from bioethics. *Journal of Business Ethics*, 34(1), 15–24.
- Fisher, C., & Lovell, A. T. A. (2000). *Accountant's responses to ethical issues at work*. London: CIMA Publishing.
- George, J. M., & Jones, G. R. (2005). *Understanding and managing organizational behavior*. Upper Saddle River, NJ: Prentice-Hall.
- Gerhart, B., & Fang, M. (2014). Pay for (individual) performance: Issues, claims, evidence and the role of sorting effects. *Human Resource Management Review*, 24(1), 41–52.
- Goodman, J. S., & Blum, T. C. (1996). Assessing the non-random sampling effects of subject attrition in longitudinal research. *Journal of Management*, 22(4), 627–652.
- Gosling, D. S., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), 504–528.
- Hall, M. (2008). The effect of comprehensive performance measurement systems on role clarity, psychological empowerment and managerial performance. *Accounting, Organizations and Society*, 33(2–3), 141–163.
- Henri, J. F. (2006). Management control systems and strategy: A resource-based perspective. *Accounting, Organizations and Society*, 31(6), 529–558.



- Hofstede, G. (1967). *The game of budget control: How to live with budgetary standards and yet be motivated by them*. Assen: Van Gorcum & Comp.
- Holmström, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization*, 7(Special), 24–52.
- Homburg, C., Artz, M., & Wieseke, J. (2012). Marketing performance measurement systems: Does comprehensiveness really improve performance? *Journal of Marketing*, 76(3), 56–77.
- Hopwood, A. G. (1972). An empirical study of the role of accounting data in performance evaluation. *Journal of Accounting Research*, 10, 156–182.
- Jaworski, B. J., & Young, S. M. (1992). Dysfunctional behavior and management control: An empirical study of marketing managers. *Accounting, Organizations and Society*, 17(1), 17–35.
- Jonason, P. K., & Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment*, 22(2), 420–432.
- Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard—measures that drive performance. *Harvard Business Review*, 70(1), 71–79.
- Kenis, I. (1979). Effects of budgetary goal characteristics on managerial attitudes and performance. *The Accounting Review*, 54(4), 707–721.
- Langevin, P., & Mendoza, C. (2013). How can management control system fairness reduce managers' unethical behaviours? *European Management Journal*, 31(3), 209–222.
- Lazarus, R.S. (1998). *Fifty years of research and theory of R. S. Lazarus*. Mahwah, NJ: Lawrence Erlbaum.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.
- Linder, S., & Torp, S. S. (2017). Middle managers' engagement in autonomous strategic actions: Does it really matter how top managers use budgets? *IEEE Transactions on Engineering Management*, 64, 450–463.
- Locke, E.A. & Taylor, M.S. (1991). Stress, coping, and the meaning of work. In A. Alan Monat & R. S. Lazarus (eds.) *Stress and coping: an anthology*, (pp. 140–157). New York: Columbia University Press.
- Lurie, Y., & Albin, R. (2007). Moral dilemmas in business ethics: From decision procedures to edifying perspectives. *Journal of Business Ethics*, 71(2), 195–207.
- Mallinger, M. A., & Greiner, L. E. (1981). Stress and coping behavior in the appraisal process. In R. D. Lansbury (Ed.), *Performance appraisal*. Macmillan Company of Australia: Melbourne.
- Marreiros, H., Tonin, M., Vlassopoulos, M., & Schraefel, M. (2017). Now that you mention it: A survey experiment on information, inattention and online privacy. *Journal of Economic Behavior & Organization*, 140(C), 1–17.
- Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Occupational Behavior*, 2(2), 99–113.
- Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological methods*, 17(3), 437–455.
- Merchant, K. A., & Van Der Stede, W. A. (2012). *Management control systems: Performance measurement, evaluation and incentives* (3rd ed.). Harlow, England: Pearson Education.
- Merchant, K.A. & White, L.F. (2017). Linking the ethics and management control literatures. *Advances in Management Accounting*, 28(1) 29, 1–29.
- Molterer, K., Hoyer, P., & Steyaert, C. (2019). A practical ethics of care: Tinkering with different 'Goods' in residential nursing homes. *Journal of Business Ethics*. <https://doi.org/10.1007/s10551-018-04099-z>.
- Neumann, B. R., Cauvin, E., & Roberts, M. L. (2012). Management control systems dilemma: Reconciling sustainability with information overload. In M. J. Epstein & J. Y. Lee (Eds.), *Advances in management accounting (Advances in Management Accounting, Volume 20)* (pp. 1–28). Bingley: Emerald Group Publishing Limited.
- Nielsen, K., & Munir, F. (2009). How do transformational leaders influence followers' affective well-being? Exploring the mediating role of self-efficacy. *Work & Stress*, 23(4), 313–329.
- Noser, A. E., Zeigler-Hill, V., & Besser, A. (2014). Stress and affective experiences: The importance of dark personality features. *Journal of Research in Personality*, 53, 158–164.
- OSHA. (2009). *European Risk Observatory Report*. Luxembourg: European Agency for Safety and Health at Work.
- Osterloh, M., & Frey, B. S. (2000). Motivation, knowledge transfer, and organizational forms. *Organization Science*, 11(5), 538–550.
- Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17, 22–27.
- Parker, D. F., & DeCotiis, T. A. (1983). Organizational determinants of job stress. *Organizational Behavior and Human Performance*, 32(2), 160–177.
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153–163.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903.
- Ponemon, L. A., & Gabhart, D. R. L. (1993). *Ethical Reasoning in accounting and auditing*. Vancouver, BC: Canada Research Foundation.
- Rosanas, J. M., & Velilla, M. (2005). The ethics of management control systems: Developing technical and moral values. *Journal of Business Ethics*, 57(1), 83–96.
- Rosen, C. C., Chang, C.-H., Djurdjevic, E., & Eatough, E. (2010). Occupational stressors and job performance: An updated review and recommendations. *Research in Occupational Stress and Well-Being*, 8, 1–60.
- Ross, W. D. (1930). *The right and the good*. Oxford: Oxford University Press.
- Schweitzer, M. E., Ordóñez, L., & Douma, B. (2004). Goal setting as a motivator of unethical behavior. *Academy of Management Journal*, 47(3), 422–432.
- Seery, M. D., Weisbuch, M., & Blascovich, J. (2009). Something to gain, something to lose: The cardiovascular consequences of outcome framing. *International Journal of Psychophysiology*, 73(3), 308–312.
- Shields, M. D., Deng, F. J., & Kato, Y. (2000). The design and effects of control systems: Tests of direct- and indirect-effects models. *Accounting, Organizations and Society*, 25(2), 185–202.
- Simola, S. (2003). Ethics of justice and care in corporate crisis management. *Journal of Business Ethics*, 46(4), 351–361.
- Simons, R. (1987). Planning, control, and uncertainty: A process view. In W. J. Bruns Jr. & R. S. Kaplan (Eds.), *Accounting and management: Field Study Perspectives* (pp. 339–362). Boston, MA: Harvard Business Press.
- Simons, R. (1994). *Levers of control: How managers use innovative control systems to drive strategic renewal*. Boston, MA: Harvard Business Press.
- Sirgy, M. J., & Lee, D. M. (2008). Well-being marketing: An ethical business philosophy for consumer goods firms. *Journal of Business Ethics*, 77(4), 377–403.
- Spence, L., & Rutherford, R. (2003). Small business and empirical perspectives in business ethics: Editorial. *Journal of Business Ethics*, 47(1), 1–5.
- Stephoe, A., & Kivimäki, M. (2012). Stress and cardiovascular disease. *Nature Reviews. Cardiology*, 9(6), 360–370.

- Turner, M. J., Jones, M. V., Sheffield, D., & Cross, S. L. (2012). Cardiovascular indices of challenge and threat states predict competitive performance. *International Journal of Psychophysiology*, 86(1), 48–57.
- Van den Broeck, A., De Cuyper, N., De Witte, H., & Vansteenkiste, M. (2010). Not all job demands are equal: Differentiating job hindrances and job challenges in the Job Demands-Resources model. *European Journal of Work and Organizational Psychology*, 19(6), 735–759.
- Van Harrison, R. (1985). The person-environment fit model and the study of job stress. In T. A. Beehr & R. S. Bhagat (Eds.), *Human stress and cognition in organizations: An integrated perspective* (Vol. 2, pp. 23–55). New York: Wiley-Interscience.
- Widener, S. K. (2007). An empirical analysis of the levers of control framework. *Accounting, Organizations and Society*, 32(7), 757–788.
- World Medical Association. (2013). *Declaration of Helsinki: ethical principles for medical research involving human subjects*. Retrieved from [www.wma.net/en/30publications/10policies/b3/index.html](http://www.wma.net/en/30publications/10policies/b3/index.html)
- Yamaguchi, I. (2003). The relationships among individual differences, needs and equity sensitivity. *Journal of Managerial Psychology*, 18(4), 324–344.
- Zhuang, L., Qiu, Y., & Peng, L. (2011). Is it the more the merrier? An exploratory study into the growing problem of information overload. *Journal of Technology Management in China*, 6(1), 69–83.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.