Routledge
Taylor & Francis Group

Accounting and Business Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rabr20

Subjective performance evaluation and managerial work outcomes

Iryna Alves & Sofia M. Lourenço

To cite this article: Iryna Alves & Sofia M. Lourenço (2023) Subjective performance evaluation and managerial work outcomes, Accounting and Business Research, 53:2, 127-157, DOI: 10.1080/00014788.2021.1959292

To link to this article: https://doi.org/10.1080/00014788.2021.1959292

Published online: 17 Sep 2021.

Accounting and Business Research, 2023
Vol. 53, No. 2, 127–157, https://doi.org/10.1080/00014788.2021.1959292

Subjective performance evaluation and managerial work outcomes

IRYNA ALVES [a*] and SOFIA M. LOURENÇO [b]

[a]*ISCAL Lisbon Accounting and Business School, Lisbon Polytechnic Institute and ADVANCE, CSG Research Centre, Lisbon, Portugal;* [b]*ISEG Lisbon School of Economics & Management, Universidade de Lisboa and ADVANCE, CSG Research Centre, Lisbon, Portugal*

Organisations design performance evaluation systems to obtain desired work outcomes. This study analyses how subjective performance evaluation (SPE), a specific type of performance evaluation, is related to managerial work outcomes—turnover intention, organisational identification, and performance. To this end, we consider two possible mechanisms: feedback quality and trust in the supervisor. Moreover, we also consider whether adding objective performance measures to SPE alters these relationships. Based on questionnaire responses from 751 top executives and middle managers in small and medium enterprises, we find that SPE is negatively related to feedback quality, but this effect is mitigated when SPE is used jointly with objective performance measures. SPE is not directly related to trust in the supervisor when feedback quality is also considered in the analysis because the two mechanisms are inter-related—we find a positive relationship between feedback quality and trust in the supervisor. Both mechanisms are negatively related to turnover intention, but only trust in the supervisor is positively related to organisational identification. Finally, both turnover intention and organisational identification are positively related to performance. Our findings suggest that companies using SPE can improve work outcomes by adding objective performance measures to their performance evaluation system.

Keywords: subjective performance evaluation; feedback quality; trust; performance

1. Introduction

Performance evaluation is an important element of management control systems as it is generally linked to compensation and rewards and, thus, has the ability to influence individual agent effort (Baker et al. 1988, Lau and Buckland 2001). Specifically, performance evaluation has a critical role in motivating agents to achieve established organisational goals (Baker et al. 1988). Subjective performance evaluation (SPE) is a widely used form of performance evaluation in modern enterprises (Bol and Smith 2011). SPE may appear as either (i) subjective judgments about managerial performance (Hartmann et al. 2010, Prendergast and Topel 1996), (ii) the application of

*Corresponding author. Email: ibalves@iscal.ipl.pt
Accepted by Frank Hartmann.

some flexibility in the weighting of objective performance measures, or (iii) the assessment of factors not initially included in the bonus plan at sole discretion of the evaluator (Gibbs et al. 2004, Höppe and Moers 2011). Thus, organisations use SPE to obtain more complete evaluations of their employees, exploit relevant information not captured by objective measures (Baiman and Rajan 1995, Baker et al. 1994, Prendergast and Topel 1993), improve managerial effort directed towards firm-value creation, and enhance incentive contracting (Gibbs et al. 2004, Lillis et al. 2017, Merchant 1989). Nevertheless, due to its discretionary nature, SPE has many biases and shortcomings, such as central tendency, favouritism, leniency, and spill-over biases (Bol et al. 2016, Bol and Smith 2011, Prendergast and Topel 1996), which may cause dysfunctional employee behaviours and attitudes (Ahn et al. 2010, Golman and Bhatia 2012, Van Rinsum and Verbeeten 2012).

Our study adds to the SPE literature by investigating the relationships between the use of SPE and managerial work outcomes— the intention to leave the organisation (turnover intention), the perception of belonging to the organisation (organisational identification), and job performance (individual performance). Instead of considering direct relationships, we rely on prior research (e.g. Hartmann and Slapnicar 2009, Van Rinsum and Verbeeten 2012) and unravel the mechanisms by which SPE may influence work outcomes. This more thorough analysis may help to explain the contradictory evidence regarding the behavioural effects of SPE. In our study, we consider feedback quality and trust in the supervisor as possible mechanisms that explain the relationship between SPE and work outcomes. Furthermore, and because prior literature has established the importance of using both objective and subjective performance measures (e.g. Anderson et al. 2020, Merchant et al. 2010), we also investigate whether the simultaneous use of SPE and objective performance measures plays a role in these relationships.

To develop and empirically test our predictions, we survey 751 top executives and middle managers in small and medium enterprises (SMEs). Even though SMEs are vital to modern economies (López and Hiebl 2015, Wapshott and Mallett 2015), prior management accounting research has mainly focused on large companies. SMEs often have simple organisational structures and unsophisticated human resource systems (Cardon and Stevens 2004, Wapshott and Mallett 2015). They are also characterised as having working relationships that are more loose and informal, and performance evaluation processes that are often not standardised (Lai et al. 2017). This informality and lack of sophisticated performance measurement systems may lead SMEs to use SPE to a greater extent than large companies. Additionally, SME owners tend to adapt management accounting and control systems according to their prior experiences, which can lead to imprecise management accounting knowledge (Perren and Grant 2000) and may suggest a decrease in the use of objective information. An informal atmosphere may also result in detrimental effects, such as relationships between supervisors and subordinates that are too close (Cassell et al. 2002, Lai et al. 2017), which can lead to evaluation biases in SPE such as assigning a higher rating to preferred subordinates (Prendergast and Topel 1996).

Subordinates may perceive that use of SPE in companies that have unsophisticated performance measurement systems and lack human and technological resources does not provide useful developmental information for them (e.g. Cassell et al. 2002), and thus consider that feedback quality is reduced. However, when SPE is complemented with objective performance measures, its negative effect is expected to be mitigated as subordinates will likely benefit from the advantages of both types of performance evaluation. On one hand, they benefit from the flexibility of SPE that allows accounting for unpredictable events and information not captured by objective performance measures (e.g. Baiman and Rajan 1995, Baker et al. 1994, Prendergast and Topel 1993). On the other hand, they benefit from objective performance measures that provide more impartial and quantifiable information (Ittner et al. 2003) and enhance the perception of

fairness, trust, and procedural justice (e.g. Saridakis et al. 2013). Thus, we consider SMEs as an attractive setting in which to explore SPE, its relationships with feedback quality and trust in the supervisor, the moderation role of objective performance measures, and ultimately the effects of those relationships on different work outcomes.

We find that although SPE is negatively related to feedback quality, this effect is mitigated when SPE is used jointly with objective performance measures. SPE is not directly related to trust in the supervisor when feedback quality is also considered in the analysis because the two mechanisms are inter-related—we find a positive relationship between feedback quality and trust in the supervisor. Both mechanisms are negatively related to turnover intention, but only trust in the supervisor is positively related to organisational identification. Finally, both turnover intention and organisational identification are positively related to performance. Our findings suggest that companies using SPE can improve work outcomes by adding objective performance measures to their performance evaluation system.

We address the call in the literature for further research on how performance measurement systems influence managerial behaviours and attitudes (Van Rinsum and Verbeeten 2012). Our study adds to empirical field studies on the individual effects of using subjectivity in compensation systems (e.g. Bicudo de Castro 2017, Van Rinsum and Verbeeten 2012) by considering a broader set of behavioural outcomes and their interplay. We also add to the relatively limited number of management accounting studies regarding the use of SPE from the subordinate's viewpoint (Bellavance et al. 2013, Bicudo de Castro 2017, Gibbs et al. 2004) rather than the supervisor's perspective (e.g. Ahn et al. 2010, Dai et al. 2018, Fehrenbacher et al. 2018), and provide new insights into the complex inter-relationships between SPE and work outcomes. Specifically, our study shows that the use of SPE is related to managerial turnover intention, organisational identification, and performance via feedback quality and trust in the supervisor. Finally, we contribute to the SPE literature by suggesting that certain negative effects of SPE can be mitigated by complementing it with objective performance measures (see the call by Franco-Santos et al. 2012). Therefore, our study helps reconcile the seemingly contradictory evidence of SPE effects on managerial behaviours and attitudes.

In the next section, we review current literature and develop our research hypotheses. We then describe our research settings, data collection, and survey instruments. Next, we present our empirical results, discuss their implications, and provide robustness tests. Finally, we provide our conclusions.

2. Literature review and hypotheses development

2.1. *Subjective performance evaluation and work outcomes*

One of the main objectives of performance evaluation is to motivate subordinates to increase their effort and, thus, improve their performance (Baker et al. 1988, Lau and Buckland 2001). Other objectives of performance evaluation include (i) focusing individual decisions and actions on strategic goals, (ii) facilitating the development, implementation, and review of business strategies, (iii) promoting communication processes that influence organisational behaviours, and (iv) selecting and retaining key personnel (e.g. Franco-Santos et al. 2012, Hopwood 1972).

The way in which employees' performances are evaluated (i.e. objectively and/or subjectively) influences their subsequent behaviours and work-related attitudes (Hopwood 1972). In this study we consider objective performance evaluation as an assessment that uses objective measures, which refer to information that is directly quantifiable by using existing data records not influenced by the evaluator. In contrast, SPE refers to an assessment that does

not come from third-party quantifiable data records and is based on the evaluator's perceptions or judgements (e.g. Dai et al. 2018, Gibbs et al. 2004, Woods 2012). SPE may therefore appear as (i) subjective judgments of performance (Bol and Smith 2011, Hartmann et al. 2010, Prendergast and Topel 1996), (ii) the application of some flexibility in the weighting of objective performance measures, or (iii) the assessment of factors not initially included in the bonus plan at sole discretion of the evaluator (Gibbs et al. 2004, Höppe and Moers 2011).

The extant literature provides mixed evidence on the effects of SPE. Specifically, SPE can reduce motivation, increase staff turnover, and decrease performance due to the presence of biases (Ahn et al. 2010, Golman and Bhatia 2012, Ittner et al. 2003, Van Rinsum and Verbeeten 2012).[1] However, when SPE is used with the goal of obtaining non-contractible information to assess actions and efforts that objective measures are unable to capture, it can create a more complete depiction of subordinates' performance (e.g. Bol 2008). Moreover, when targets that are defined through objective measures become *ex post* difficult to achieve, subjectivity appears to be an effective method to keep subordinates more motivated (Gibbs et al. 2004) and, consequently, enhance their work-related outcomes.

To reconcile these different findings, we argue that the behavioural effects of SPE can occur due to two different mechanisms—perceived feedback quality and trust in the supervisor. Perceived feedback quality, which refers to the subordinates' perception about the value of the performance feedback provided by supervisors (Hartmann and Slapnicar 2009), is an important mechanism for work outcomes because it translates the subordinates' view about the usefulness of this information, namely about how they can improve in the future (Hall 2008, Ilgen et al. 1979, Wright 2004). This is particularly relevant in subjective evaluations where subordinates may not know *ex ante* the dimensions/tasks in which they are being evaluated and may suspect of supervisors' biases in this type of evaluation (e.g. Bol 2008, 2011, Kunz 2015, Prendergast and Topel 1993, 1996). Trust in the supervisor, which refers to the belief that subordinates hold about the support of their supervisors (Hartmann and Slapnicar 2009), is also a key driver of behavioural outcomes because subordinates tend to reciprocate supervisors' care and concern (Mayer and Davis 1999), which generate positive attitudes towards the organisation (Lau et al. 2008). Again, in subjective evaluations subordinates may be concerned with supervisors' biases and limitations, which may give rise to perceptions of managerial discrimination and unfairness (Prendergast and Topel 1993). Therefore, both feedback quality and trust in the supervisor are promising mechanisms to understand the behavioural effects of SPE.

Furthermore, the extant literature suggests that (i) the use of different types of measures, in isolation, can lead to suboptimal performance evaluation system, and (ii) the use of objective performance measures may impact the effectiveness of SPE (e.g. Bol and Smith 2011, Dai et al. 2018, Kunz 2015, Manthei and Sliwka 2019). Therefore, the use of SPE in conjunction with objective performance measures is key for a comprehensive evaluation of SPE outcomes. Hence, we propose that the direct relationships between (i) SPE and perceived feedback quality and (ii) SPE and trust in the supervisor are contingent on the simultaneous use of SPE and objective performance measures. In the next sections, we provide the theoretical justification for these relationships.

[1]Despite predictions of agency theory regarding the negative effect of SPE biases on employees' behaviours, the literature provides evidence of a positive effect of leniency bias on performance improvement (e.g. Bol 2011, Zábojník 2014). Bol (2011) explains these results as higher organisational interest in increasing individual performance than in the accuracy of the performance ratings.

2.2. *Subjective performance evaluation, feedback quality, and work outcomes*

2.2.1. *Subjective performance evaluation and feedback quality*

SPE gives a supervisor the opportunity to use relevant but non-contractible information, to consider more dimensions of performance, to apply his/her specific knowledge of the setting in which the subordinate operates, and to adjust the performance evaluation to unexpected events (Baiman and Rajan 1995, Baker et al. 1994, Rajan and Reichelstein 2009). All these factors suggest the possibility that the performance feedback provided when using SPE will be better than using objective performance measures (e.g. Keasey et al. 2000). However, such a possibility may not materialise due to the limitations of SPE. Specifically, due to the subjectivity in the personal judgment of the supervisor and the potential evaluation biases in SPE, subordinates may perceive supervisors' feedback as vague, uninformative, discretional, and ambiguous (Ahn et al. 2010, Fulk et al. 1985, Ittner et al. 2003, Prendergast and Topel 1993).[2] Moreover, SPE can lead subordinates to feel confused as to what constitutes good performance, question whether compensation depends upon their performance, and suspect that inconsistencies occur within the organisation (Ittner et al. 2003), all of which deteriorate the subordinates' perceived quality of performance feedback.

We argue that the negative effects of SPE dominate its potential benefits in perceived feedback quality because subordinates may be focused on supervisors' bias and limitations (e.g. Bol 2008, Bol 2011, Kunz 2015, Prendergast and Topel 1993, 1996) when they receive feedback that is not grounded in objective and verifiable information. Additionally, even supervisors tend to undervalue subjective elements in the evaluation process compared to objective elements because they perceive the former as less scientific than the latter for performance evaluation decisions (Dai et al. 2018). Therefore, we predict that the use of SPE is negatively related to the subordinates' perceived feedback quality and state our first hypothesis as follows:

H1a: There is a negative relationship between SPE and perceived feedback quality.

When SPE is used jointly with objective measures, supervisors have the opportunity to consider a more complete picture of subordinates' performance (Van Rinsum and Verbeeten 2012) by obtaining information from both subjective and objective elements (e.g. Baiman and Rajan 1995, Baker et al. 1994, Woods 2012), which will likely enhance perceived feedback quality. Prior research supports this argument by suggesting that SPE and objective evaluation should be used as complements (Baiman and Rajan 1995, Baker et al. 1994, Ittner et al. 2003, Moers 2005, Rajan and Reichelstein 2009, Van der Stede et al. 2006), which allows supervisors to provide more differentiated and comprehensive feedback to their subordinates. Therefore, we expect perceived feedback quality to be enhanced when SPE is used in conjunction with objective measures by promoting the mutual advantages and reducing the drawbacks of each one (Merchant et al. 2010).

On one hand, by minimising distortions caused by objective measures (Baiman and Rajan 1995, Baker et al. 1988, Baker et al. 1994, Bol 2008, Gibbs et al. 2004, Moers 2005, Rajan

[2]The existing literature recognizes SPE biases such as central tendency, leniency, and spill-over (Berger et al. 2013, Bol 2011, Bol and Smith 2011, Moers 2005). Central tendency bias occurs when a supervisor fails to sufficiently distinguish between subordinates in the evaluation of their performances (Ahn et al. 2010, Golman and Bhatia 2012). Leniency bias occurs when most employees get exaggerated performance evaluations, comparative to their true or average performance (Golman and Bhatia 2012). Spill-over effect occurs when supervisors bias their evaluation of a subjective measure consistent with the level of performance of an unrelated objective measure (Bol and Smith 2011, Dai et al. 2018).

and Reichelstein 2009, Van der Stede et al. 2006, Woods 2012), SPE mitigates performance measurement problems and aligns the interests of the principal with those of the agent (Baker et al. 1988, Bol 2008, Gibbs et al. 2004). Thus, the introduction of subjectivity into performance evaluation practices to correct and augment incomplete objective performance measures (Baker et al. 1994) (i) reduces the noise and the manipulation potential of quantitative performance measures (Gibbs et al. 2004); (ii) improves incentive alignment (Höppe and Moers 2011, Lillis et al. 2017, Merchant 1989); and (iii) reduces risk for employees (Gibbs et al. 2004).

On the other hand, when objective performance measures are added to SPE, they may mitigate supervisors' biases associated with the sole use of SPE (Ahn et al. 2010, Moers 2005). Particularly, the use of objective measures may anchor supervisors' personal judgment on objective and quantifiable information about subordinates' performance, provide a more direct comparison among peers, and give supervisors more specific and verifiable information to share with subordinates during the performance evaluation process (Dai et al. 2018). As a consequence, subordinates may consider feedback from SPE, when objective performance measures are also used, as more informative (Baker et al. 1994, Wright 2004) and more fair (Bol and Smith 2011) than when only SPE is used. Therefore, we expect that the use of objective performance measures in addition to SPE will mitigate the negative relationship between SPE and perceived feedback quality. Accordingly, the following hypothesis is proposed:

H1b: The relationship between SPE and perceived feedback quality is weaker (i.e. less negative) when SPE is used in conjunction with objective performance measures.

2.2.2. *Feedback quality and work outcomes*

Feedback is considered to be an essential function of any performance evaluation process because it can help subordinates to identify (i) their strengths and weaknesses; (ii) what needs to be done to improve performance; and (iii) a direction for their future effort (Hall 2008, Zábojník 2014). Ilgen et al. (1979) argue that the effectiveness of feedback depends on its informational value and how it is perceived and accepted by the subordinate. Thus, feedback that is perceived by subordinates as high quality should lead to more favourable work outcomes, such as high performance and organisational identification, and decreased turnover intention. Performance, in this study, is considered as individual job performance characterised by specific functions and competencies (Mahoney et al. 1965). Organisational identification is defined as an individual's perception of oneness with or belongingness to an organisation (Ashforth and Mael 1989). Turnover intention is defined as a conscious and deliberate willingness to leave the organisation (Tett and Meyer 1993).

High-quality feedback should provide detailed information about task requirements and suggestions for improving performance (Hall 2008, Ilgen et al. 1979) that will allow subordinates to improve future performance. High-quality feedback should also clarify organisational goals and the expectations of the organisation relative to the subordinate (Ilgen et al. 1979, Wright 2004). This improved understanding by the subordinate is expected to increase organisational identification. Finally, by providing actionable developmental information and clarifying organisational and individual goals (Hall 2008, Ilgen et al. 1979, Wright 2004), high-quality feedback will foster subordinates' long-term perspective of the organisation because they will expect favourable future evaluations and promotion opportunities, which can decrease their turnover intention.

Extant evidence on the behavioural effects of high-quality feedback is consistent with the expected positive work outcomes as described above. For example, in an experimental study, Tziner and Latham (1989) show that performance feedback combined with goals can increase

work satisfaction and organisational commitment/identification. Wright (2004) shows that quality feedback clarifies organisational goals, which enhances employees' work motivation. Based on this evidence, he suggests that quality feedback can increase productivity. Joo et al. (2015) show that developmental feedback (i.e. helpful and useful information provided by the supervisor that facilitates employees' learning and development on the job) can decrease employees' turnover intention.

We extend this evidence to our model, and test the following hypotheses:

H2a: There is a negative relationship between perceived feedback quality and turnover intention.
H2b: There is a positive relationship between perceived feedback quality and individual performance.
H2c: There is a positive relationship between perceived feedback quality and organisational identification.

2.3. *Subjective performance evaluation, trust, and work outcomes*

2.3.1. *Subjective performance evaluation and trust*

Prior literature suggests that the presence of biases in SPE may give rise to a negative relationship between SPE and trust (e.g. Hopwood 1972, Lau and Buckland 2001, Van Rinsum and Verbeeten 2012). When using SPE, supervisors can (i) assign exaggerated performance evaluations to some employees, comparative to their true or average performance (Golman and Bhatia 2012); (ii) decrease dispersion in performance ratings (Bol 2011, Golman and Bhatia 2012, Moers 2005); and (iii) assign higher grades of performance evaluation to preferred subordinates (Prendergast and Topel 1996). These biases may prompt a subordinate's sense of managerial discrimination and unfairness (Prendergast and Topel 1993), which can then reduce trust in the supervisor (Fulk et al. 1985, Prendergast and Topel 1993). Therefore, the use of subjectivity may decrease (rather than increase) the incentive role of performance evaluation (Ahn et al. 2010). Prior empirical research supports these claims by showing that flexible, non-accounting-based performance evaluation is likely to be associated with greater mistrust in supervisors than inflexible, accounting-based evaluation (Hartmann and Slapnicar 2009, Hopwood 1972, Lau and Buckland 2001, Lau and Shohilin 2005, Van Rinsum and Verbeeten 2012). We posit that the use of SPE may undermine the subordinate's trust in the supervisor. Our hypothesis is as follows:

H3a: There is a negative relationship between SPE and trust in the supervisor.

Notwithstanding, the combined use of SPE and objective measures may highlight the advantages of each type of evaluation while minimising their disadvantages (Merchant et al. 2010). On one hand, SPE permits the supervisor to consider events that subordinates cannot control but that influence their performance. Managers who use SPE can also disregard evaluations based only on explicit measures that focus on a few tasks and, therefore, consider a more complete picture of a subordinate's performance (Van Rinsum and Verbeeten 2012). On the other hand, objective performance measures are often argued to be less susceptible to supervisor's biases (Ahn et al. 2010, Moers 2005) and provide more discriminability in performance assessment (Ahn et al. 2010). These characteristics may mitigate the perception of a discretionary evaluation by the supervisor when only SPE is used. Additionally, the use of objective measures in performance evaluation may signal the supervisor's intention to uphold norms of honesty and clarity (e.g. Lau and Buckland 2001, Lau and Shohilin 2005), which may enhance a subordinate's trust in the supervisor.

Objective measures provide an anchor for supervisors' evaluations as they reflect objective, verifiable, and quantifiable information (Ittner et al. 2003) and they are weighted more than subjective elements by supervisors in the evaluation process (Dai et al. 2018). Objective measures also permit a more direct comparison among peers, which may reduce a supervisor's favouritism for a given subordinate. These mechanisms will reduce the subordinates' perceptions of discrimination and unfairness and, therefore, mitigate their mistrust in the supervisor due to the sole use of SPE. Therefore, we expect that adding objective performance measures to SPE will mitigate the negative relationship between SPE and a subordinate's trust in the supervisor. Accordingly, we test the following hypothesis:

H3b: The relationship between SPE and trust in the supervisor is weaker (i.e. less negative) when SPE is used in conjunction with objective performance measures.

2.3.2. *Trust and work outcomes*

Trust plays a critical role in the development of behavioural norms among individuals in organisations (Lau et al. 2008). Trust in the supervisor is especially important as it can improve subordinates' cooperation (e.g. Jones 1995) and desire to reciprocate (Mayer and Davis 1999), enhance information exchange, and reduce opportunistic behaviour (e.g. Fisher et al. 2005). Hence, greater trust in the supervisor is likely to increase the effectiveness of incentive contracts (e.g. Gibbs et al. 2004) and contribute to an open and supportive atmosphere in the firm (Klein and Kim 1998). In sum, a 'high level of trust in the supervisors is likely to be translated into a favourable attitude towards the organization' (Lau et al. 2008, p. 126), which can increase performance and organisational identification and decrease turnover intention. Specifically, trust in the supervisor generates a desire in the subordinate to cooperate and reciprocate (e.g. Jones 1995, Mayer and Davis 1999), which will likely lead to an increase in effort and, subsequently, performance. Trust in the supervisor also promotes an open and supportive atmosphere in the firm (Klein and Kim 1998), which can increase organisational identification. Finally, trust enhances information exchange and reduces opportunistic behaviour (e.g. Fisher et al. 2005), which can reduce turnover intentions.

Prior empirical research on the behavioural effects of trust generally supports these claims. For example, Van Rinsum and Verbeeten (2012) show that a higher degree of trust in the supervisor is associated with higher effort motivation, which can be translated into higher individual performance as other studies have shown (Gibbs et al. 2004). Sholihin and Pike (2009) show that a higher level of trust in the supervisor is associated with higher organisational commitment and, therefore, organisational identification. In contrast, mistrust by subordinates creates environments that are characterised by a lack of transparency, which lead to high levels of anxiety, stress, and frustration (Lau and Buckland 2001), thus discouraging an individual's perception of belongingness to the organisation. Finally, Costigan et al. (2011) show that trust can decrease turnover intention. Conversely, a person who does not trust his/her supervisor tends to either believe that his/her performance is undervalued or considers the results of his/her evaluation to be unfair, which can lead to lower effort and higher propensity to quit (Prendergast and Topel 1993).

We extend these arguments and empirical evidence to our framework, and test the following hypotheses:

H4a: There is a negative relationship between trust and turnover intention.

H4b: There is a positive relationship between trust and individual performance.

H4c: There is a positive relationship between trust and organisational identification.

2.4. *Relationships among outcome variables*

2.4.1. *Turnover intention and individual performance*

Prior literature suggests that employees may demonstrate diverse withdrawal behaviours and reduce job input as a result of their turnover intentions (e.g. Griffeth et al. 2000). Consistent with this proposition, Bowen (1982, p. 12) argues that 'if employees intending to quit are performing tasks where they have a great deal of control over their productivity, then I/Q (i.e. turnover intention) may lead to lower job performance'. Similarly, using an investment perspective, Hui et al. (2007) suggest that it is unlikely that employees who intend to leave the organisation will make a substantial investment in it (i.e. perform well). Instead, employees with high turnover intention develop low performance expectations that, in turn, undermine their current performance (Curry et al. 1998). The organisational literature corroborates this argument and empirically shows the detrimental effect of turnover intention on individual employee behaviours such as commitment (Haque et al. 2017), job performance (Hui et al. 2007), and overall deviant behaviours (e.g. Christian and Ellis 2014).[3] Based on the arguments presented above, our hypothesis is as follows:

H5a: Turnover intention is negatively related to individual performance.

2.4.2. *Organisational identification and individual performance*

Prior literature suggests that the 'extent to which people identify with a particular social group determines their inclination to behave in terms of their group membership' (Ellemers et al. 1999, p. 372). Individuals who strongly identify with their organisation are more committed, more involved with the group (i.e. organisation), and invest more effort in the group to which they belong (Ellemers et al. 1999, Liu et al. 2011). Specifically, employees who feel a strong sense of belonging to the organisation have more positive beliefs about the organisation, cooperate more with their peers, and exhibit higher performance (e.g. Ashforth and Mael 1989, Van Dick et al. 2006, Walumbwa et al. 2008). Accordingly, our hypothesis 5b is:

H5b: Organisational identification is positively related to individual performance.

2.5. *Feedback quality and trust*

Prior literature suggests that trust in the supervisor is an essential element of performance evaluation processes and is related to individual acceptance and accurate processing of feedback (Ilgen et al. 1979). Specifically, extant management accounting studies show that greater trust in the supervisor depends on the effectiveness of feedback provided in the performance evaluation process (e.g. Coletti et al. 2005, Hartmann and Slapnicar 2009). In fact, Coletti et al. (2005) argue that high-quality feedback positively influences trust because subordinates consider supervisors' attempts to provide good feedback as indications of the supervisors' desire to help subordinates to improve. Hartmann and Slapnicar (2009) show that perceived feedback quality is related to trust in a setting in which objective performance measures are used. We expect this relationship to be present when SPE is used. The final hypothesis is:

H6: There is a positive relationship between perceived feedback quality and trust in the supervisor.

The aforementioned hypotheses are aggregated into our theoretical model, shown in Figure 1.

[3]The examples of deviant behaviours studied by Christian and Ellis (2014) are items such as having 'taken property from work without permission' and 'used an illegal drug or consumed alcohol on the job'.

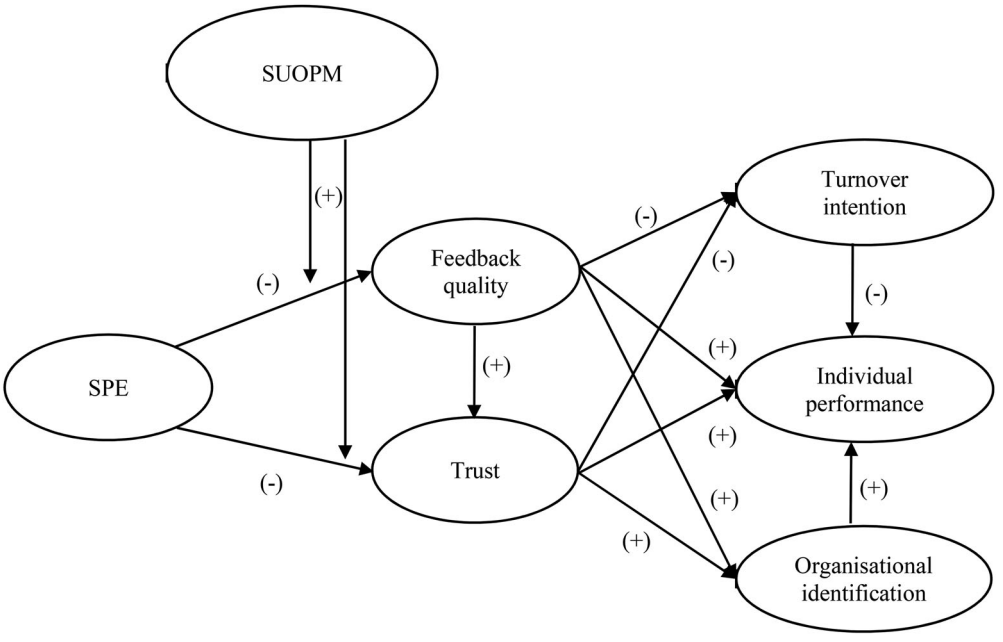

Figure 1. Proposed conceptual model of the relationship between SPE and work outcomes.

3. Research method

3.1. *Sample and data description*

We collect data from two questionnaires completed by SME top executives and middle managers (managers, hereafter) in Portugal.[4] The first questionnaire collects data regarding organisational characteristics and the second regarding the manager.

The first questionnaire was successfully sent to 22,997 SMEs; 4192 usable responses were received. This corresponds to an overall response rate of 18.23%. Using this database, a team of research assistants contacted each firm via telephone to collect the name and e-mail addresses of the managers. This information served to create a second database that was used to send the second questionnaire. Thus, from the initial sample of 4192 SMEs, we obtained 11,748 names with their related job roles and e-mail addresses (personal or organisational). However, due to refusal of participation, inability to contact the firms, or error in the e-mailing process, the number of questionnaires successfully delivered was 7988.[5] We received 1466 responses, but because we have to restrict our sample to respondents who completed the question about SPE, the number of usable observations falls to 1275. This corresponds to a usable response rate of 16%. However, 524 respondents indicated that they do not have a supervisor, which reduces the final sample size to 751.

[4]In our study, SME is defined based on the number of employees, which excludes micro-enterprises. Specifically, SMEs are those with 10–249 employees. This criterion is consistent with the European Commission definition (European Commission 2020). We obtain a list SMEs from Informa D&B, a firm that specialises in corporate information.

[5]The reasons for the inability to contact the firm include: extinct organisation, missing or invalid telephone contact, and no data available on the internet. Errors in the sending-mailing process include: incorrect e-mail address, full mailbox, insufficient capacity of the mailbox, and e-mail considered as spam.

To test for potential (non-) response bias (e.g. Armstrong and Overton 1977), we apply two procedures commonly used in the literature. First, we compare the size, location, and industry representation of the 538 companies at which the 751 respondents work to the 2377 companies of the 7237 non-respondents from the list of 7988 e-mails sent successfully.[6] A chi-square test shows that industry and location categories do not differ between the respondents' sample and the non-respondents' sample. Conversely, an independent sample t-test regarding firm size (measured by the number of employees) shows a difference between respondents ($\overline{X} = 71.06$) and non-respondents ($\overline{X} = 32.70$); the difference is statistically significant ($p < 0.01$ for a two-tailed t-test). This difference in sample means by size is a bias of this study, which implies that the results we find may not generalise to very small companies.

Second, we compare the responses of early and late respondents, as late respondents are more similar to non-responders (Armstrong and Overton 1977). We use the date of the received response to distinguish between early and late respondents. Overall, the results indicate that there are no statistically significant differences between the two groups for the variables examined in our model, except for some measures that were used to build the 'simultaneous use of objective performance measures' variable.[7] This difference is also a bias of this study.

Given that all data are obtained from a single source at a single point in time, we address the concern of common method bias by applying both procedural and statistical remedies (Podsakoff et al. 2003, Speklé and Widener 2018). *Ex ante*, (i) we used simple wording and provided examples, where appropriate, to improve understanding and appropriateness for the sample frame; (ii) we safeguarded the confidentiality of the answers; (iii) we used a distinct response scale for mediator variables as compared to the outcome variables; and (iv) we carefully and thoroughly pre-tested the questionnaire with practitioners (managers not in the sample) and academics (management accounting researchers) to ensure the instrument's clarity, readability, and length. *Ex post,* we apply Harman's one-factor test as well as a confirmatory factor analysis (CFA) to address common method variance concerns regarding the measures (Mossholder et al. 1998, Podsakoff and Organ 1986). The factor solution yields four factors with eigenvalues greater than 1. Together they account for 66% of the total variance. The first factor explains 27.62% of the total variance, which means that it does not account for a majority of the variance. The confirmatory factor analysis shows that the single-factor model does not fit the data well: $X^2(152) = 3908.344$, $p < 0.01$, CMINDF = 25.71, the root mean square error of

[6]Because the total sample of 751 responses may include more than one observation from the same firm, this may be a concern for the tests that require independence of observations. To overcome this problem, we randomly select one observation per firm and eliminate all additional observations from the same firm. The resulting sample of 538 observations, which includes only one observation for each firm, is used to test for (non-) response bias. Additionally, in Table 5, which uses multiple observations from the same firm, we cluster data at the firm level and find that the (untabulated) results are inferentially identical to those reported, with similar significance levels for the hypothesized paths. We also conduct a robustness test in Table 6 by using the sample of 538 observations (which include only one observation for each firm) to generate two groups that have an equal number of observations. We find, in untabulated results, that inferences regarding the moderation effect of the 'simultaneous use of objective performance measures' variable remains robust.

[7]The 'simultaneous use of objective performance measures' variable is based on 22 different objective performance measures. We detect response bias on three of these measures: sales growth ($p < 0.10$), employee satisfaction ($p < 0.10$), and other metrics ($p < 0.01$). Early respondents have a mean of 2.46, 1.90, and 5.07, while late respondents a mean of 0.51, 0.47, and 29.43, respectively. We also compare the demographic variables among early and late respondents, and we do not find statistically significant differences.

approximation (RMSEA) = 0.181, the comprehensive fit index (CFI) = 0.428, Tucker-Lewis Index (TLI) = 0.357.[8] In sum, the results of both procedures indicate that common method bias is unlikely to be a concern for this study.

Table 1 reports the descriptive statistics for the demographic variables and organisational characteristics. The respondents are, on average, 44 years old, have 21 years of professional experience, and 14 years of tenure in the company. Most occupy the position of chief financial officer (CFO) or chief executive officer (CEO) (26% and 14%, respectively). Table 1 also shows that respondents are mainly in manufacturing and wholesale and retailing industries (35% and 27%, respectively). Sample firms have an average of 79 employees.

3.2. *Measurement of variables*

To increase the overall quality of the survey instrument, we use, when possible, questions that have been validated in prior studies and follow Bedford and Speklé's (2018a) recommendations.[9] Below, we present more details on the measurement of the variables obtained. Appendix A presents the questions and items used to measure the main constructs. Table 2 presents descriptive statistics for the measurement instruments.

3.2.1. *Subjective performance evaluation (SPE)*

To ensure the clarity and the correct understanding of the notion of SPE, which relies on an internal decision-making process by the supervisor, we explain in the questionnaire that in the performance evaluation of managers there may be objective elements (quantified performance measures) as well as subjective elements (for example, the evaluator's personal judgment that is used in the evaluation of qualitative criteria, in the choice of performance measures, or in the weights assigned to them). By providing this information to respondents, we ensure that the SPE variable captures the different types of subjectivity that are established in the management accounting literature (e.g. Bol and Smith 2011, Gibbs et al. 2004, Hartmann et al. 2010, Höppe and Moers 2011, Ittner et al. 2003), thus providing a comprehensive measure of the SPE variable. Next, we ask respondents whether their supervisor used subjective elements to evaluate their performance during the last evaluation event and, if so, we ask respondents to indicate the weight of the subjective elements in the overall evaluation of their performance. Thus, we measure SPE as the percentage of the subjective elements in the overall assessment of the respondent's performance, which may vary from 0 to 100%.

3.2.2. *Feedback quality*

We assess perceived feedback quality by using an instrument employed in prior accounting literature (e.g. Hartmann and Slapnicar 2009). We ask respondents to indicate on a 5-point Likert scale (1 = completely disagree, 5 = completely agree) the level of agreement with each of 4 items

[8]CFI compares how much better an implied model fits as compared to a null model, TLI contains a penalty for lack of parsimony, and RMSEA adjusts for both sample size and number of degrees of freedom (Kline 2016). CFI and TLI close to 1 (Hair et al. 2014), RMSEA less than 0.08 (Browne and Cudeck 1989), and CMINDF ratio less than 5 (Arbuckle and Wothke 1999) reflect a good model fit.

[9]In some cases, we make slight modifications with the aim to fit the measures to the present research context.

Table 1. Descriptive statistics for demographic variables ($N = 751$).

Panel A: Demographic information of respondents and firms				
Variable	Min	Mean	Max	St Dev
Before logarithmic transformation				
Age (years)	21	43.82	73	8.68
Professional experience (years)	0.5	21.22	52	9.02
Company tenure (years)	0.5	14.43	47	8.68
Firm size (no. of employees)	10	78.54	249	62.02
After logarithmic transformation				
Ln_age	3.04	3.76	4.29	0.20
Ln_professional experience	−0.69	2.95	3.95	0.51
Ln_company tenure	−0.69	2.44	3.85	0.76
Ln_firm size	2.30	4.01	5.52	0.89

Panel B: Job titles	Number	Percentage
CEO	103	13.72
CFO	195	25.97
Commercial/sales/billing manager	87	11.58
Human resources manager	47	6.26
Logistic/marketing manager	46	6.13
Production/quality manager	77	10.25
Other managers	136	18.11
Executive member of the board of directors	60	7.99

Panel C: Breakdown of responses by industry	Number	Percentage
Accommodation and food service activities	36	4.79
Administrative and support service activities	25	3.33
Agriculture, forestry, and fishing	7	0.93
Construction	32	4.26
Financial and insurance activities	37	4.93
Information and communication	45	5.99
Manufacturing	263	35.02
Professional, scientific, and technical activities	41	5.46
Real estate activities	9	1.20
Transport and storage	37	4.93
Water supply, sewerage, waste management and remediation activities	20	2.66
Wholesale and retail trade, repair of motor vehicles and motorcycles	199	26.50

(e.g. 'My supervisor gives me useful feedback about my job performance' and 'I value the feedback I receive from my supervisor').

3.2.3. *Trust in the supervisor*

We assess trust in the supervisor by using an instrument employed in prior accounting literature (e.g. Hartmann and Slapnicar 2009). The instrument for trust captures the extent to which respondents agree to each of 3 items (e.g. 'My supervisor will always act in my favour if he has the chance'), where 1 = completely disagree and 5 = completely agree.

Table 2. Descriptive statistics for the measurement instruments.

Variable	N	Min	Mean	Median	Max	St Dev
SPE	751	0	19.82	0	100	27.50
Feedback quality	704	1	3.69	4.00	5	0.79
Trust in supervisor	700	1	3.55	3.67	5	0.80
Individual performance	672	1	4.86	4.71	7	0.87
Turnover intention	744	1	2.57	2.00	7	1.58
Organisational identification	748	1.5	6.15	6.00	7	0.85
SUOPM (dummy variable)	751	–	0.24	–	–	–

Note: The number of observations on the variables is lower than in the total sample ($N = 751$) due to missing data.

3.2.4. *Individual performance*

To measure individual performance at the managerial level, we use a self-reported survey question developed by Mahoney et al. (1965) and recurrently used in accounting studies (Bedford and Speklé 2018b). The original scale assesses managerial performance along eight dimensions that are related to planning, investigating, coordinating, evaluating, supervising, staffing, negotiating, and representing, and also include an overall assessment of performance. Hall (2008) reduces the scale to 7 items by excluding 2 items (negotiating and representing) because of low factor loadings and not belonging to a unidimensional managerial performance scale. In order to reduce the length of the questionnaire, we use this shortened form and ask respondents to indicate on a 7-point Likert scale (1 = well below average to 7 = well above average) the extent to which their performance was below or above average on each of the 7 remaining items.

3.2.5. *Turnover intention*

To measure managerial turnover intention, we use 3 items from the *Michigan Organizational Assessment Questionnaire* developed by Cammann et al. (1979). The responses for the first and second items ('I often think about quitting' and 'I will probably look for a new job in the next year') are based upon a 7-point Likert-scale with a range from 1 = strongly disagree to 7 = strongly agree. The responses for the third item ('How likely is it that you will actively look for a new job in the next year?') are based on a 7-point Likert-scale with a range from 1 = not at all likely to 7 = extremely likely.

3.2.6. *Organisational identification*

To measure organisational identification, we use one of the theoretical components of Cook and Wall's *Organizational Commitment Instrument* (1980). Respondents indicate their agreement with each of 3 items (e.g. 'I feel myself to be part of the organisation') on a 7-point Likert scale ranging from 1 = strongly disagree to 7 = strongly agree.

3.2.7. *Simultaneous use of objective performance measures (SUOPM)*

First, to capture the use of objective performance measures, we ask respondents to indicate the relative weight placed on different performance measures in the manager's annual compensation (Appendix A). The possible responses are between 0 and 100%. The performance measures listed in this study are consistent with classifications employed in prior management accounting research (e.g. Bellavance et al. 2013, Van der Stede et al. 2006), and correspond to four

categories: financial, non-financial customer-oriented, non-financial employee-oriented, and non-financial internal operations. Second, we create a dummy variable 'SUOPM' that has the value of 1 if the respondent has in his/her performance evaluation both SPE and any objective performance measure, and 0 otherwise.

4. Results and discussion

To investigate our hypothesised model, we employ structural equation modelling (SEM) and multi-group analysis, using Stata 14.1.[10] First, to estimate the entire model (H1a to H5b, with the exception of H1b and H3b), we adopt a generally accepted two-step SEM method (e.g. Hair et al. 2014) that encompasses: (i) evaluation of the measurement model to assure its fit and (ii) examination of the structural model.[11] Multivariate normality of data is a principal assumption of the full information maximum likelihood (FIML) in SEM. However, Kline (2016) notes that multivariate non-normality can usually be identified through univariate procedures. An examination of kurtosis and skewness reveals that all variables indicated in Figure 1 are well below 10.0 and 3.0, respectively, which demonstrate that the data are within tolerable levels of univariate normality (Kline 2016).[12] Second, we conduct multi-group analysis to test for moderating effect of SUOPM (H1b and H3b).

4.1. *Measurement model and assessment of model fit*

In a first step, we evaluate the measurement model with a confirmatory factor analysis (CFA). In the initial approach, we eliminate two items from further analyses. First, we remove one item of the feedback quality scale (fbq2: 'The performance information I get from my supervisor is generally not very meaningful') due to low standardised loading (factor loading = 0.25), which can be due to reversed scaling (e.g. Roszkowski and Soven 2010). Second, we exclude one reverse-score item of the organisational identification scale (ident3: 'I would not recommend a close friend to join our staff') that has the largest measurement error variance, which causes both low composite reliability (CR) and low average extracted variance (AVE).

Furthermore, following extant literature (e.g. Groen et al. 2017), we also allow for covariance of the error terms for two items of the feedback quality measure ('The feedback I receive from my supervisor helps me do my job' and 'I value the feedback I receive from my supervisor') as suggested by Stata and is theoretically justified—these two items are very similar in that they both deal with the value and usefulness of the supervisor's feedback.

Table 3 shows the results of CFA (after removing the two problematic items), which support the unidimensionality of the constructs and their convergent validity, since (i) all standardised loadings are significant ($p < 0.01$) and greater than 0.5; (ii) AVE of each construct is equal to or greater than 0.5; and (iii) composite reliability (CR) is greater than 0.7 (Hair et al. 2014).

[10]Our empirical method uses SEM because it accounts for the measurement error in the latent variables and allows for the examination of relationships among multiple dependent variables. Moreover, 'it is still generally true that SEM is a large-sample technique' (Kline 2016, p. 14), which is the case in our study.
[11]We use the full information maximum likelihood (FIML) estimation method for both the measurement model and full structural model. FIML is superior to other imputation techniques as it gives unbiased estimates of means, variances, and other parameters (Hair et al. 2014).
[12]We also perform the examination of kurtosis and skewness for respondents' tenure, professional experience, and age, as well as firm size after their logarithmic transformation. All variables were within tolerable values.

Table 3. Results of confirmatory factor analysis.

Construct indicators	Standardised factor loadings	AVE	CR
Feedback quality		0.60	0.85
fbq1	0.87		
fbq3	0.81		
fbq4	0.62		
Trust in supervisor		0.59	0.82
trust1	0.76		
trust2	0.87		
trust3	0.67		
Individual performance		0.59	0.91
perf1	0.77		
perf2	0.70		
perf3	0.79		
perf4	0.79		
perf5	0.81		
perf6	0.77		
perf7	0.75		
Turnover intention		0.72	0.89
turnov1	0.70		
turnov2	0.94		
turnov3	0.89		
Organisational identification		0.50	0.73
ident1	0.69		
ident2	0.73		

Note: Result of Kaiser–Meyer–Olkin test is 0.88, indicating good sampling adequacy (greater than 0.80) and the Bartlett test of Sphericity is highly significant ($p < 0.01$).

Table 4 provides further support for the discriminant validity of the questionnaire constructs since the Cronbach's alphas (on the diagonal) exceed the bi-variate correlations in all cases (Bagozzi et al. 1991). Overall, all Cronbach's alphas are comfortably above the lower limits of acceptability, generally considered to be 0.50–0.60 (Nunnally 1978). In sum, the measurement model exhibits strong psychometric properties.

Table 4. Multi-trait matrix[a].

Variable	1	2	3	4	5	6	7
1. SPE	n.a.						
2. Feedback quality	−0.15***	**0.83**					
3. Trust in supervisor	−0.12***	0.66***	**0.81**				
4. Individual performance	0.02	0.00	0.04	**0.91**			
5. Turnover intention	0.14***	−0.41***	−0.45***	0.04	**0.88**		
6. Organisational identification	−0.15***	0.42***	0.47***	0.11***	−0.52***	**0.66**	
7. SUOPM	0.37***	0.03	−0.03	0.03	−0.00	0.01	n.a.

[a]The diagonal of the matrix is the Cronbach's alpha for each variable. The remainder of the table reports the bi-variate Pearson's correlation coefficients, except for the correlations between SUOPM (dichotomous variable) and the other six variables. In the latter cases, we use point-biserial correlation (Kline 2016). n.a. is non-applicable.
Note: *** indicates the significance of the p-value at < 0.01, all p-values are two-tailed. The number of observations per correlation varies due to missing data. The pairwise correlation between SPE and SUOPM has the maximum number of observations (751) while the correlation between trust in supervisor and individual performance has the minimum number of observations in this analysis (653).

Furthermore, the confirmatory factor analysis shows that the measurement model does well in all fit indicators: $X^2(137) = 406.142$, $p < 0.01$, CMINDF = 2.96, CFI = 0.959, TLI = 0.949, RMSEA = 0.051. According to the chi-square test, the data fit the model when chi-square is non-significant. However, according to Lawrence et al. (1982), the chi-square test is sensitive to sample size and, therefore, may be misleading. Specifically, large samples inflate the chi-square and decrease the likelihood of achieving a good model fit. Thus, although the chi-square is significant, the CMINDF is less than 5, indicating an acceptable fit of the model (Arbuckle and Wothke 1999). Additionally, according to Browne and Cudeck (1989), the model should be rejected only if RMSEA is above 0.1, which is not the case in our study. Therefore, our measurement model clearly satisfies these basic criteria.

4.2. *Structural model results*

In the second step, we establish the structural model by adding the proposed structural paths to the measurement model (as specified in the hypotheses). Following prior research (e.g. Riketta 2005), we allow the endogenous variables of turnover intention and organisational identification to be correlated.[13]

Table 5 presents the results of the structural model in terms of path coefficients and goodness-to-fit indices. Overall, fit indices indicate that the base model is a good fit for the data. However, the existence of a well-fitting model does not ensure that it is the only appropriate model (Kline 2016). Accordingly, we compare the base model to alternative theory-based models, which are more parsimonious, to rule out alternative model specifications. The first alternative model (Model1) is a non-mediated model, the second alternative model (Model2) considers trust as the only mediator, and the third alternative model (Model3) considers feedback quality as the only mediator.[14] The results presented in Table 5 show that the base model exhibits better fit than all the other models.[15] Significant results of the base model are presented in Figure 2.

4.3. *Multi-group analysis for moderator variable SUOPM*

To test the moderating effect of SUOPM, we apply multi-group analysis to our base model.[16] Using the SUOPM variable, we divide the entire sample into two groups: the first group consists of 181 managers whose compensation contracts include both SPE and objective performance measures (SUOPM group); the second group comprises 570 managers whose compensation contracts lack either SPE or objective performance measures or lack both (non-SUOPM group). However, prior research advocates that unequal sample size of two groups may be problematic for multi-group analysis (i) due to its large effect on the constrained estimates and (ii) because a large sample has the power to demonstrate statistical significance even when the effect is subtle,

[13]A meta-analytic review of behavioural studies reports that organisational identification exhibits strong and negative correlations with turnover intention (e.g. Riketta 2005).

[14]Although the alternative Model1 shows an insignificant association between SPE and performance ($p > 0.10$), prior research indicates that significant indirect effects can be detected, even when the total effect is not statistically significant (Kenny and Judd 2014).

[15]Our results are inferentially identical to those reported in the Table 5 when we cluster data at the firm level to account for the fact that some responses come from managers who work in the same firm.

[16]Even though we only hypothesize different relationships when objective measures are used in conjunction with SPE for the first part of our model, for the sake of completeness and full transparency, we report the multi-group analysis for all the relationships in our model. If we run the multi-group analysis only for the first part our model, our results are inferentially identical to those reported in Table 6.

Table 5. Results of SEM analyses.

Pred. sign	Independent variable	Dependent variable	Base model Coeff.	Model1 Coeff.	Model2 Coeff.	Model3 Coeff.
−	SPE	Feedback quality	−0.181***	–	–	−0.195***
−		Trust in the supervisor	−0.017	–	−0.169***	–
		Turnover intention	–	0.152***	–	–
		Individual performance	–	0.043	–	–
		Organisational identification	–	−0.180***	–	–
+	Feedback quality	Trust in the supervisor	0.813***	–	–	–
−		Turnover intention	−0.160*	–	–	−0.487***
+		Individual performance	−0.103	–	–	−0.088
+		Organisational identification	0.065	–	–	0.571***
−	Trust in the supervisor	Turnover intention	−0.398***	–	−0.532***	–
+		Individual performance	0.039	–	−0.042	–
+		Organisational identification	0.618***	–	0.676***	–
−	Turnover intention	Individual performance	0.241***	0.243***	0.246***	0.238***
+	Organisational identification	Individual performance	0.336***	0.307***	0.334***	0.354***
Model fit		N	751	751	751	751
		X^2	412.578	254.917	324.289	331.746
		Df	140	60	98	97
		p-value	0.000	0.000	0.000	0.000
		CMINDF	2.95	4.25	3.31	3.42
		CFI	0.959	0.954	0.957	0.956
		TLI	0.949	0.940	0.947	0.946
		RMSEA	0.051	0.066	0.055	0.057
Model comparison		X^2 difference		2,255.07	1,328.61	1,201.57
		Df		11	7	6
		p-value		$p < 0.01$	$p < 0.01$	$p < 0.01$

Note: * and *** indicates $p < 0.10$ and $p < 0.01$, respectively, two-tailed test. Standardised coefficients are presented. First column reports predicted sign of the relationships.

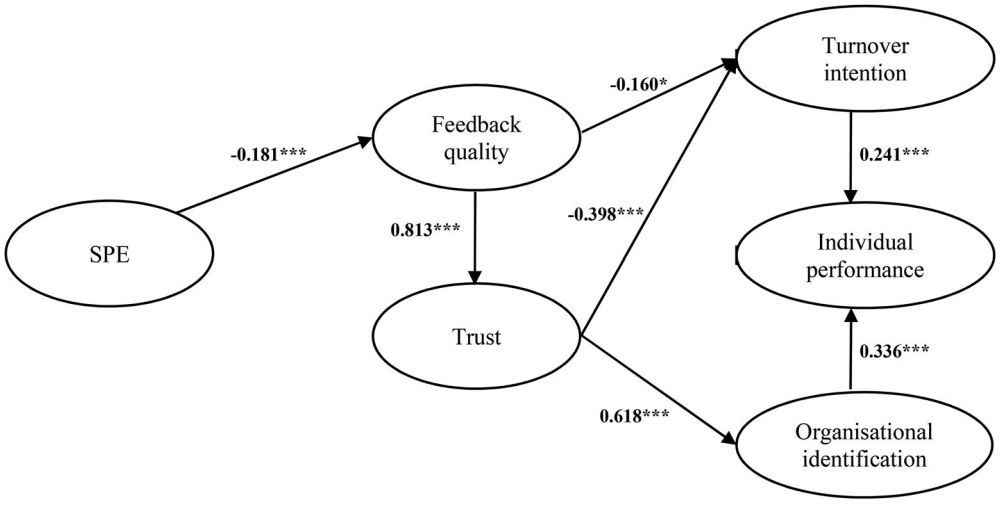

Figure 2. Graphical depiction of significant results for the base model.

while a small sample has little power to demonstrate statistical significance (Acock 2013). To mitigate this concern, we randomly generate a subsample of 181 managers from the non-SUOPM group of 570 managers and conduct multi-group analysis for two groups with equal sample size (181 managers from the SUOPM group and 181 managers from the non-SUOPM group). We use a X^2 difference test to compare an unrestricted model (where structural paths are allowed to differ for the two SUOPM groups) to a restricted model (where structural paths are constrained to be the same for the two SUOPM groups). The results are presented in Table 6.[17] The results show that the models differ significantly ($p < 0.01$), which means that an unrestricted model fits the data better than a restricted model, indicating that the two groups, divided by the moderating variable, are significantly different from each other.

4.4. *Discussion of hypotheses*

As shown in Table 5, we find strong support for H1a because SPE is negatively associated with the perceived quality of feedback ($p < 0.01$). This finding supports the argument that subordinates perceive supervisors' feedback from SPE as vague and uninformative due to the presence of evaluation biases (Fulk et al. 1985, Ittner et al. 2003, Prendergast and Topel 1993).

H1b predicts that SUOPM moderates the relationship between SPE and perceived feedback quality. The individual path coefficients (see Table 6) are significant and negative for both groups and, as expected, stronger for the non-SUOPM group (coeff. for non-SUOPM group = −0.423, $p < 0.01$; coeff. for SUOPM group = −0.199, $p < 0.01$). Furthermore, the two coefficients are significantly different ($p < 0.01$), which supports H1b.[18] These findings support the view that

[17]Our results are inferentially identical to those reported in the Table 6 when we cluster data at the firm level to account for the fact that some responses come from managers who work in the same firm.

[18]We find similar untabulated results that support the moderating effects of the SUOPM on the SPE-feedback quality relationship when we compare the samples with unequal size of the two groups (i.e. 181 managers from SUOPM group and 570 managers from non-SUOPM group). The paths are significant for both groups and stronger for the non-SUOPM group (coeff. = −0.217, $p < 0.01$) than for the SUOPM group (coeff. = −0.127, $p < 0.05$), and between-group coefficients are significantly different ($p < 0.10$).

Table 6. Results of the moderator analysis.

Pred. Sign	Independent variable	Dependent variable	Multi-group analysis for SUOPM: non-SUOPM ($N = 181$)	Multi-group analysis for SUOPM: SUOPM ($N = 181$)	Multi-group analysis for SUOPM: X^2 difference
−	SPE	Feedback quality	−0.423***	−0.199***	12.427***
−		Trust in the supervisor	0.080	−0.023	1.961
+	Feedback quality	Trust in the supervisor	0.833***	0.622***	1.243
−		Turnover intention	−0.314**	−0.099	0.619
+		Individual performance	−0.069	−0.182	0.400
+		Organisational identification	0.221	0.129	0.216
−	Trust in the supervisor	Turnover intention	−0.286*	−0.426***	1.342
+		Individual performance	0.128	0.118	0.000
+		Organisational identification	0.539***	0.591***	0.053
−	Turnover intention	Individual performance	0.141	0.294*	0.300
+	Organisational identification	Individual performance	0.073	0.269	0.396
Model comparison statistics		N			362
		X^2 unrestricted (restricted)			535.192 (778.06)
		Df unrestricted (restricted)			311 (327)
		p-value			0.000
		CMINDF			1.72
		CFI			0.924
		TLI			0.916
		RMSEA			0.063
		X^2 difference			242.87
		Df			16
		p-value			$p<0.01$

Note: *, ** and *** indicates $p < 0.10$, $p < 0.05$ and $p < 0.01$, respectively, two-tailed test. Standardised coefficients are presented. First column reports predicted sign of the relationships.

managers whose evaluations include both SPE and objective measures perceive feedback from SPE as more efficient—that is more informative (Baker et al. 1994, Wright 2004) and more fair (Bol and Smith 2011)—than managers whose compensation contracts do not include SUOPM. This is an important finding because it suggests a mechanism to mitigate the negative consequences of SPE.

Table 5 also shows that, as predicted, feedback quality is negatively associated with turnover intention (coeff. = −0.160, $p < 0.10$), which supports H2a. However, contrary to our expectation, the results do not support H2b and H2c. The relationships between feedback quality and individual performance (H2b), and feedback quality and organisational identification (H2c) are not statistically significant. The lack of a significant association between feedback quality and organisational identification and performance may be justified by the presence of trust in our model. In fact, as Model3 of Table 5 shows, when perceived feedback quality is considered as the only mediator, the relationship between perceived feedback quality and organisational identification is significant and has the predicted sign (coeff. = 0.571, $p < 0.01$) but the relationship between perceived feedback quality and performance is still not significant. Despite the lack of a direct significant relationship between perceived feedback

quality and performance, in the base model the relationship between feedback quality and trust is positive and statistically significant (as expected), and trust is significantly related to turnover intention and organisational identification, which are ultimately related to performance (as we will discuss below).

Interestingly, we find no evidence to support H3a, which predicts a negative relationship between SPE and trust in supervisor. A possible explanation for this is that the relationship between SPE and trust is indirect via feedback quality. Model2 corroborates this speculation. That is, when we consider trust as the only mediator, the relationship between SPE and trust becomes significant and has the predicted sign (coeff. = −0.169, $p < 0.01$). Moreover, untabulated results show that when the base model does not include the path between feedback quality and trust, the relationships between both SPE and feedback quality (coeff. = −0.185, $p < 0.01$) and SPE and trust (coeff. = −0.161, $p < 0.01$) are negative and significant.

Furthermore, we do not find support for H3b, which predicts that SUOPM moderates the relationship between SPE and trust in the supervisor. As shown in Table 6, we do not find a statistically significant difference between the two groups regarding this path. Again, the lack of a significant result in this test may occur because of the indirect association between SPE and trust via feedback quality. Our speculation is corroborated by multi-group analysis for Model 2 (i.e. without feedback quality). Untabulated results show that the individual path coefficients are significant and negative for both groups and, as expected, stronger for non-SUOPM group (coeff. for non-SUOPM group = −0.277, $p < 0.01$; coeff. for SUOPM group = −0.139, $p < 0.05$). Furthermore, the two coefficients are significantly different ($p < 0.10$).

In the hypotheses development, we predict that trust in the supervisor is related to turnover intention, performance, and organisational identification. The results of the structural model are partially consistent with these predictions. Specifically, we observe the hypothesised negative relationship between trust and turnover intention, and the positive relationship between trust and organisational identification, thus supporting H4a and H4c ($p < 0.01$), respectively. These results corroborate the extant management accounting literature that highlights the importance of trust in enhancing organisational commitment and identification (Sholihin and Pike 2009) and reducing the turnover intention (Prendergast and Topel 1993). However, we do not find statistical evidence to support H4b, which predicts a positive association between trust and individual performance ($p = 0.742$). This is in line with Dirks and Ferrin (2001), who argue that the effect of trust on performance is weak and inconsistent.

Regarding the relationships among outcome variables, we do not find support for H5a, which predicts a negative relationship between turnover intention and individual performance. Instead, we observe a positive and significant association between these two variables (coeff. = 0.241, $p < 0.01$). This result is in accordance with evidence that suggests turnover intention may lead to an increase in individual performance if this performance is visible to potential employers (Bowen 1982) or if the employee's performance determines pay and benefits (Christian and Ellis 2014). Moreover, we find strong support for H5b because organisational identification is positively related to individual performance (coeff. = 0.336, $p < 0.01$). This result corroborates the extant literature, which shows a positive effect of organisational identification on individual performance (e.g. Liu et al. 2011, Walumbwa et al. 2008).

We also find support for H6 because the relationship between feedback quality and trust is positive and statistically significant (coeff. = 0.813, $p < 0.01$). This evidence from our SPE model is consistent with prior research that uses other types of performance evaluation (e.g. Coletti et al. 2005, Hartmann and Slapnicar 2009).

Additionally, in the non-mediated model (Model1) we find a significant positive relationship between SPE and turnover intention (coeff. = 0.152, $p < 0.01$) and a significant negative

relationship between SPE and organisational identification (coeff. = −0.180, $p < 0.01$).[19] These findings support the argument that various drawbacks of SPE, such as favouritism, ignorance of some measures by evaluators, frequently changed criteria of bonus payment, and the presence of biases in subjective evaluation, may negatively influence employee behaviour (for example, increase turnover and decrease organisational identification among discontented employees) (Ittner et al. 2003, Prendergast and Topel 1993).

4.5. *Robustness tests*

We perform several additional tests to ensure that our findings are robust.[20] First, we estimate the SEM model by adding one control path from family ownership to SPE and find that the statistical inferences are unaffected by this procedure (with the exception of one path between feedback quality and turnover intention that becomes insignificant).[21] The association between family ownership and SPE is not significant, which suggests that ownership structure does not influence the level of subjectivity in SMEs.

Second, we estimate the SEM model by adding one control path from risk attitude to turnover intention.[22] The results show a positive and significant relationship between risk attitude and turnover intention (coeff. = 0.218, $p < 0.01$), which indicates that managers with a greater risk tolerance have a higher turnover intention (e.g. MacCrimmon and Wehrung 1985). More importantly, we find that the results remain qualitatively unchanged because all of the results that were previously (non-) significant are still (non-) significant.

Third, we estimate a series of SEM models that control for manager's age, tenure, professional experience, and position.[23] We implement these controls by modelling paths

[19]We observe similar results when we estimate the base model (not tabulated) with additional direct paths from SPE to the three outcome variables (turnover intention, organisational identification, and individual performance). Our inferences for all hypothesized paths remain unchanged.

[20]To adequately compare model fit of the hypothesized model to the measurement model and to ensure that the results are not driven by the inclusion of control variables (Groen et al. 2017), we do not include control variables in our main analysis (see Table 5). Nevertheless, in order to provide a more comprehensive test, we add control variables in our robustness analyses.

[21]Prior research suggests that family businesses make less use of professional management control practices such as appraisal systems due to limited organisational capabilities (e.g. inadequate management of personnel, limited financial, and human resources) (e.g. Neubauer et al. 2012). We consider family ownership to account for the possibility that family firms exhibit specific characteristics that can affect the level of subjectivity. To measure family ownership, we ask respondents to indicate the percentage of the company's equity held by family members. The response varies from 0 to 100%.

[22]We use the variable *risk attitude* to control for risk-seeking behaviours in turnover intention. Extant research suggests that turnover intention may be an indicator of willingness to engage in risky behaviour (MacCrimmon and Wehrung 1985). Moreover, Allen et al. (2005) find that the relationship between turnover intentions and actual turnover relation is stronger for individuals with low risk aversion. We assess risk attitude by adapting an instrument developed by Pennings and Smidts (2000) for our sample. We ask respondents to indicate on a 7-point Likert scale the level of agreement with each of 3 items related to risk attitude, ranging from 1 = strongly disagree to 7 = strongly agree. Factor analysis of the variable *risk attitude* reveals that the three questions load on a single factor. All standardized loadings are significant ($p < 0.01$) and greater than 0.57, AVE is 0.598 and CR is 0.83.

[23]Since past research shows that demographic characteristics influence individual perceptions, behaviours, and attitudes (e.g. Burdett et al. 2011, Martin and Edwards 2009), we control for their effects. The respondent's tenure, professional experience, and age are self-reported measures that are collected via the questionnaire and transformed into their natural logarithms, which mitigates their high skewness and kurtosis. We also control for position by including dummy variables for CEO, CFO, production and quality manager, logistics and marketing manager, commercial, sales and billing manager, human resource

between each of the control variables and each of the five endogenous variables (feedback quality, trust, organisational identification, turnover intention, and performance). We find that the results are inferentially identical, with similar significance levels for the hypothesised paths.

We observe significant associations between the following control variables and the endogenous variables: (i) age and trust (coeff. = −0.054, $p < 0.10$), age and turnover intention (coeff. = −0.077, $p < 0.05$), age and performance (coeff. = 0.068, $p < 0.10$); (ii) tenure and turnover intention (coeff. = −0.113, $p < 0.01$), tenure and organisational identification (coeff. = 0.130, $p < 0.01$), tenure and performance (coeff. = −0.072, $p < 0.10$); (iii) professional experience and turnover intention (coeff. = −0.063, $p < 0.10$), professional experience and organisational identification (coeff. = 0.087, $p < 0.05$); and (iv) board member and turnover intention (coeff. = −0.088, $p < 0.05$), CEO and performance (coeff. = 0.120, $p < 0.05$), CEO and organisational identification (coeff. = 0.194, $p < 0.01$), CFO and organisational identification (coeff. = 0.099, $p < 0.10$), logistics and marketing manager and organisational identification (coeff. = 0.087, $p < 0.05$), and board member and organisational identification (coeff. = 0.156, $p < 0.01$).

These results indicate that older managers report higher performance and have a lower turnover intention, despite their lower trust in the supervisor. Moreover, more tenured and more experienced managers have a higher organisational identification and a lower turnover intention, and more tenured managers report lower performance.

Regarding the position, the results show that CEOs report higher performance, while board members exhibit lower turnover intention. Furthermore, CEOs, CFOs, logistics and marketing managers, and board members have higher organisational identification.

Fourth, we run the SEM model by including a control path between firm size and each of the six variables of the model.[24] We find that the results are also inferentially identical, with similar significance levels for the hypothesised paths. We also observe a positive and significant relationship between firm size and individual performance (coeff. = 0.087, $p < 0.05$).

Finally, our sample includes board members and CEOs, which may give rise to the criticism that they do not have a supervisor similar to middle managers in the sample. To overcome this criticism, we run our base model (i) without board members, (ii) without CEOs, and (iii) without both board members and CEOs. Untabulated results are inferentially identical to those reported in Table 5, with similar significance levels for the hypothesised paths (with the exception of one path between feedback quality and turnover intention that becomes insignificant when we exclude CEOs from our sample).

5. Conclusions

This study investigates the relationship between SPE and managerial behaviours and attitudes in SMEs, taking into consideration the role of feedback quality and trust in the supervisor, and whether the relationships between SPE and these variables are moderated by the simultaneous use of SPE and objective performance measures. SMEs are an interesting setting in which to explore these relationships due to their simple organisational structures, unsophisticated

manager, and board member; the default is 'other managers', which includes, for example, operations managers, environment managers, and country managers.

[24]Prior literature suggests that firm size influences the quality of employee relations (e.g. Talacchi 1960). We use firm size to account for the possibility that small and medium size companies can exhibit different properties that affect the level of subjectivity, perceived feedback, and trust in supervisor, as well as work-related outcomes. To measure firm size, we use the natural logarithm of the number of employees, which mitigates the high skewness and kurtosis of the original variable.

human resource systems (Cardon and Stevens 2004, Wapshott and Mallett 2015), informal working relationships, and unstandardised performance evaluation processes (Lai et al. 2017).

We find that SPE is negatively associated with feedback quality. This result is consistent with the extant literature that suggests that feedback from SPE may be considered as ambiguous (Ahn et al. 2010), vague, and uninformative (Fulk et al. 1985, Ittner et al. 2003, Prendergast and Topel 1993). We also find that SPE is negatively associated with trust via feedback quality. This result corroborates prior research, which suggests an undermining effect of SPE on trust in the supervisor (Van Rinsum and Verbeeten 2012) and provides further explanation for this effect. Our findings are in line with the argument that feedback quality acts as a second mechanism by which performance evaluation affects trust in the supervisor (e.g. Coletti et al. 2005, Hartmann and Slapnicar 2009, Lau and Buckland 2001). Furthermore, we find that feedback quality is negatively associated with turnover intention both directly and indirectly via trust in the supervisor. The positive association between feedback quality and trust is consistent with prior research, which suggests that effective feedback leads subordinates to have high trust in supervisors (e.g. Coletti et al. 2005, Hartmann and Slapnicar 2009). Additionally, we find that trust is indirectly associated with individual performance via a negative association with turnover intention and a positive association with organisational identification, both of which are positively associated with individual performance. These findings explain previous inconsistent results regarding the effect of trust on performance (Dirks and Ferrin 2001) by showing that there are both positive (via organisational identification) and negative (via turnover intention) paths.

Additionally, our moderator analyses reveal that the relationship between SPE and feedback quality is contingent on the simultaneous use of SPE and objective performance measures. This result corroborates management control literature that suggests a beneficial effect of using both SPE and objective performance measures (e.g. Baiman and Rajan 1995, Baker et al. 1994, Bol 2008, Van der Stede et al. 2006, Woods 2012), and extends it to behavioural effects. Specifically, our study suggests that SMEs may increase managerial work outcomes by implementing a performance measurement system based on the use of both SPE and objective performance measures. The introduction of objective performance measures in addition to SPE in SMEs may be important as they force these companies to ‘objectify’ performance (what is measured) and ‘quantify’ in a precise manner the desired levels of performance (targets set for the objective performance measures).

Our study has two main implications for practice. First, top-level managers can mitigate the negative effects of SPE by adding objective performance measures to the evaluation system. Hence, SPE with its particular advantages does not need to be abandoned by SMEs but rather complemented by objective performance measures to achieve better managerial work outcomes. Additionally, our study reinforces the importance of supervisor behaviour for subordinates’ work outcomes as these outcomes can also be improved if supervisors invest in providing useful and meaningful feedback and in developing a supportive and open atmosphere.

Our results should be interpreted in light of the limitations of this study. First, our study has a (non-) response bias in terms of firm size. This limitation may be explained by the greater propensity of larger organisations to participate in surveys, while smaller firms prefer to ignore them. Larger firms are also more likely to have performance measurement systems and, therefore, be more inclined to participate in the surveys concerning this issue. Therefore, our results should be generalised with caution to small firms. Second, this study does not consider the possibility that SPE may not be an exogenous, but rather an endogenous choice by the firm. For instance, firms are motivated or even forced to use SPE (as substitute for or complement to objective performance measures), when objective performance measures are noisy or difficult to obtain when evaluating managers’ efforts. Thus, the investigation of SPE as an endogenous variable may represent a fruitful avenue for subsequent studies. Third, we do not consider

how performance evaluation practices and managerial turnover intention, organisational identification, and performance evolve over time, which limits our ability to make causal claims regarding the relationships we document. Fourth, we concentrate our research in SMEs and, as such, our results may not be directly generalisable to other types of companies, such as micro- or large-sized enterprises. Finally, we use SPE as a unidimensional construct (percentage of the subjective elements in the overall assessment of the respondent's performance) and, as such, our results may not be generalisable to other types of subjectivity used in performance evaluation, such as rule-driven subjectivity and supervisor-driven subjectivity (Bicudo de Castro 2017).

By using longitudinal data, future research may be able to unlock causal mechanisms that are difficult to identify when using cross-sectional data. Our research could also be extended by directly including the effect of biases (e.g. central, leniency and spill-over biases, and favouritism) in analyses regarding the relationship between SPE and feedback quality and trust. This would provide further insight into the mechanisms through which SPE affects these two variables.

Acknowledgements

This paper is based on Iryna Alves' PhD thesis completed at the Lisbon School of Economics and Management of the Universidade de Lisboa, which was supervised by Sofia M. Lourenço. This paper was distinguished with the Best Accounting Paper Award at the XXVIII Portuguese-Hispanic Conference on Scientific Management (2018). We thank Aldónio Ferreira, António Samagaio, Bianca Groen, Patrícia Quesado, Peter Kroos, Rolf Bruhl and participants at the XV Grudis Conference and Doctoral Colloquium, 1st CSG Research Forum, Manufacturing Accounting Research Conference 2016, ISEG Doctoral Seminar 2016, XVI Grudis Conference and Doctoral Colloquium, 40th Annual Congress of the European Accounting Association, 9th Conference on Performance Measurement and Management Control, 2018 Management Accounting Section Midyear Meeting: Research Conference of the American Accounting Association, and XXVIII Portuguese-Hispanic Conference on Scientific Management for helpful comments in earlier drafts. Furthermore, we are grateful for the fruitful comments of the editor, Mark Clatworthy, the associate editor, Frank Hartmann, and the anonymous reviewer. All errors remain our own.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by FCT - Fundação para a Ciência e a Tecnologia (Portugal) [national funding through research Grant PTDC/EGE-GES/119607/2010,UID/SOC/04521/2013,UID/SOC/04521/2019, UIDB/04521/2020].

ORCID

Iryna Alves http://orcid.org/0000-0002-3706-1152
Sofia M. Lourenço http://orcid.org/0000-0001-6565-2155

References

Acock, A. C., 2013. *Discovering Structural Equation Modeling Using Stata*. Revised ed. Texas, USA: Stata Press.

Ahn, T. S., Hwang, I., and Kim, M., 2010. The impact of performance measure discriminability on ratee incentives. *The Accounting Review*, 85 (2), 389–417.

Allen, D. G., Weeks, K. P., and Moffitt, K. R., 2005. Turnover intentions and voluntary turnover: the moderating roles of self-monitoring, locus of control, proactive personality, and risk aversion. *Journal of Applied Psychology*, 90 (5), 980–990.

Anderson, S. W., Dekker, H. C., Sedatole, K. L., and Wiersma, E., 2020. When one size does not fit all: using ex post subjective ratings to provide parity in risk-adjusted compensation. *Management Accounting Research*, 49 (1), 1–12.
Arbuckle, J. L., and Wothke, W., 1999. *Amos 4.0 User's Guide*. Chicago, IL: SmallWaters Corporation.
Armstrong, J. S., and Overton, T. S., 1977. Estimating nonresponse bias in mail surveys. *Journal of Marketing Research*, 14, 396–402.
Ashforth, B. E., and Mael, F., 1989. Social identity theory and the organization. *Academy of Management Review*, 14 (1), 20–39.
Bagozzi, R. P., Yi, Y., and Phillips, L. W., 1991. Assessing construct validity in organizational research. *Administrative Science Quarterly*, 36 (3), 421–458.
Baiman, S., and Rajan, M. V., 1995. The informational advantages of discretionary bonus schemes. *The Accounting Review*, 70 (4), 557–579.
Baker, G. P., Gibbons, R., and Murphy, K., 1994. Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics*, 109 (4), 1125–1156.
Baker, G. P., Jensen, M. C., and Murphy, K. J., 1988. Compensation and incentives: practice vs. theory. *The Journal of Finance*, 43 (3), 593–616.
Bedford, D. S., and Speklé, R. F., 2018a. Construct validity in survey-based management accounting and control research. *Journal of Management Accounting Research*, 30 (2), 23–58.
Bedford, D. S., and Speklé, R. F., 2018b. Constructs in survey-based management accounting and control research: an inventory from 1996 to 2015. *Journal of Management Accounting Research*, 30 (2), 269–322.
Bellavance, F., Landry, S., and Schiehll, E., 2013. Procedural justice in managerial performance evaluation: effects of subjectivity, relationship quality, and voice opportunity. *The British Accounting Review*, 45 (3), 149–166.
Berger, J., Harbring, C., and Sliwka, D., 2013. Performance appraisals and the impact of forced distribution —an experimental investigation. *Management Science*, 59 (1), 54–68.
Bicudo de Castro, V., 2017. Unpacking the notion of subjectivity: performance evaluation and supervisor discretion. *The British Accounting Review*, 49 (6), 532–544.
Bol, J. C., 2008. Subjectivity in compensation contracting. *Journal of Accounting Literature*, 27, 1–32.
Bol, J. C., 2011. The determinants and performance effects of managers' performance evaluation biases. *The Accounting Review*, 86 (5), 1549–1575.
Bol, J. C., and Smith, S. D., 2011. Spillover effects in subjective performance evaluation: bias and the asymmetric influence of controllability. *The Accounting Review*, 86 (4), 1213–1230.
Bol, J. C., Kramer, S., and Maas, V. S., 2016. How control system design affects performance evaluation compression: the role of information accuracy and outcome transparency. *Accounting, Organizations and Society*, 51 (C), 64–73.
Bowen, D., 1982. Some unintended consequences of intention to quit. *Academy of Management Review*, 7 (2), 205–211.
Browne, M. W., and Cudeck, R., 1989. Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, 24 (4), 445–455.
Burdett, K., Carrillo-Tudela, C., and Coles, M. G., 2011. Human capital accumulation and labor market equilibrium. *International Economic Review*, 52 (3), 657–677.
Cammann, C., Fichman, M., Jenkins, D., and Klesh, J., 1979. *The Michigan organizational assessment questionnaire*. Unpublished manuscript, University of Michigan, Ann Arbor.
Cardon, M. S., and Stevens, C. E., 2004. Managing human resources in small organizations: what do we know? *Human Resource Management Review*, 14 (3), 295–323.
Cassell, C., Nadin, S., Gray, M., and Clegg, C., 2002. Exploring human resource management practices in small and medium sized enterprises. *Personnel Review*, 31 (6), 671–692.
Christian, J., and Ellis, A., 2014. The crucial role of turnover intentions in transforming moral disengagement into deviant behavior at work. *Journal of Business Ethics*, 119 (2), 193–208.
Coletti, A. L., Sedatole, K. L., and Towry, K. L., 2005. The effect of control systems on trust and cooperation in collaborative environments. *The Accounting Review*, 80 (2), 477–500.
Cook, J., and Wall, T., 1980. New work attitude measures of trust, organizational commitment and personal need non-fulfilment. *Journal of Occupational Psychology*, 53 (1), 39–52.
Costigan, R. D., Insinga, R. C., Berman, J. J., Kranas, G., and Kureshov, V. A., 2011. Revisiting the relationship of supervisor trust and CEO trust to turnover intentions: a three-country comparative study. *Journal of World Business*, 46 (1), 74–83.
Curry, L. A., Snyder, C. R., Cook, D. L., Ruby, B. C., and Rehm, M., 1998. Role of hope in academic and sport achievement. *Journal of Personality and Social Psychology*, 73 (6), 1257–1267.

Dai, N. T., Kuang, X., and Tang, G., 2018. Differential weighting of objective versus subjective measures in performance evaluation: experimental evidence. *European Accounting Review*, 27 (1), 129–148.
Dirks, K. T., and Ferrin, D. L., 2001. The role of trust in organizational settings. *Organization Science*, 12 (4), 450–467.
Ellemers, N., Kortekaas, P., and Ouwerkerk, J. W., 1999. Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. *European Journal of Social Psychology*, 29 (2-3), 371–389.
European Commission, 2020. *User guide to the SME definition*. Luxembourg: Publications Office of the European Union. Available from: https://op.europa.eu/en/publication-detail/-/publication/756d9260-ee54-11ea-991b-01aa75ed71a1 [Accessed 25 July 2021].
Fehrenbacher, D. D., Schulz, A. K. D., and Rotaru, K., 2018. The moderating role of decision mode in subjective performance evaluation. *Management Accounting Research*, 41 (1), 1–10.
Fisher, J. G., Maines, L. A., Peffer, S. A., and Sprinkle, G. .B, 2005. An experimental investigation of employer discretion in employee performance evaluation and compensation. *The Accounting Review*, 80 (2), 563–583.
Franco-Santos, M., Lucianetti, L., and Bourne, M., 2012. Contemporary performance measurement systems: a review of their consequences and a framework for research. *Management Accounting Research*, 23 (2), 79–119.
Fulk, J., Brief, A. P., and Barr, S. H., 1985. Trust-in-supervisor and perceived fairness and accuracy of performance evaluations. *Journal of Business Research*, 13 (4), 301–313.
Gibbs, M., Merchant, K., Van Der Stede, W., and Vargus, M., 2004. Determinants and effects of subjectivity in incentives. *The Accounting Review*, 79 (2), 409–436.
Golman, R., and Bhatia, S., 2012. Performance evaluation inflation and compression. *Accounting, Organizations and Society*, 37 (8), 534–543.
Griffeth, R. W., Hom, P. W., and Gaertner, S., 2000. A meta-analysis of antecedents and correlates of employee turnover: update, moderator tests, and research implications for the next millennium. *Journal of Management*, 26 (3), 463–488.
Groen, B. A. C., Wouters, M. J. F., and Wilderom, C. P. M., 2017. Employee participation, performance metrics, and job performance: a survey study based on self-determination theory. *Management Accounting Research*, 36 (1), 51–66.
Hair, J., Black, W., Babin, B., and Anderson, R., 2014. *Multivariate Data Analysis*. 7th ed. London, UK: Pearson Education Limited.
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.
Haque, A., Fernando, M., and Caputi, P., 2017. The relationship between responsible leadership and organisational commitment and the mediating effect of employee turnover intentions: an empirical study with Australian employees. *Journal of Business Ethics*, 156 (3), 759–774.
Hartmann, F., and Slapnicar, S., 2009. How formal performance evaluation affects trust between superior and subordinate managers. *Accounting, Organizations and Society*, 34 (6-7), 722–737.
Hartmann, F., Naranjo-Gil, D., and Perego, P., 2010. The effects of leadership styles and use of performance measures on managerial work-related attitudes. *European Accounting Review*, 19 (2), 275–310.
Höppe, F., and Moers, F., 2011. The choice of different types of subjectivity in CEO annual bonus contracts. *The Accounting Review*, 86 (6), 2023–2046.
Hopwood, A. G., 1972. An empirical study of the role of accounting data in performance evaluation. Empirical research in accounting: selected studies. *Journal of Accounting Research*, 10 (Suppl.), 156–182.
Hui, C., Wong, A., and Tjosvold, D., 2007. Turnover intention and performance in China: the role of positive affectivity, Chinese values, perceived organizational support and constructive controversy. *Journal of Occupational and Organizational Psychology*, 80 (4), 735–751.
Ilgen, D. R., Fisher, C. D., and Taylor, M. S., 1979. Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64 (4), 349–371.
Ittner, C. D., Larcker, D. F., and Meyer, M. W., 2003. Subjectivity and the weighting of performance measures: evidence from a balanced scorecard. *The Accounting Review*, 78 (3), 725–758.
Jones, T. M., 1995. Instrumental stakeholder theory: a synthesis of ethics and economics. *Academy of Management Review*, 20 (2), 404–437.
Joo, B.-K., Hahn, H. J., and Peterson, S. L., 2015. Turnover intention: the effects of core self-evaluations, proactive personality, perceived organizational support, developmental feedback, and job complexity. *Human Resource Development International*, 18 (2), 116–130.

Keasey, K., Moon, P., and Duxbury, D., 2000. Performance measurement and the use of league tables: some experimental evidence of dysfunctional consequences. *Accounting and Business Research*, 30 (4), 275–286.
Kenny, D. A., and Judd, C. M., 2014. Power anomalies in testing mediation. *Psychological Science*, 25 (2), 334–339.
Klein, H. J., and Kim, J. S., 1998. A field study of the influence of situational constraints, leader-member exchange, and goal commitment on performance. *Academy of Management Journal*, 41 (1), 88–95.
Kline, R. B., 2016. *Principles and Practice of Structural Equation Modeling*. 4th ed. New York: Guilford Press.
Kunz, J., 2015. Objectivity and subjectivity in performance evaluation and autonomous motivation: an exploratory study. *Management Accounting Research*, 27 (1), 27–46.
Lai, Y., Saridakis, G., and Johnstone, S., 2017. Human resource practices, employee attitudes and small firm performance. *International Small Business Journal: Researching Entrepreneurship*, 35 (4), 470–494.
Lau, C. M., and Buckland, C., 2001. Budgeting-the role of trust and participation: a research note. *ABACUS*, 37 (3), 369–388.
Lau, C. M., and Shohilin, M., 2005. Financial and nonfinancial performance measures: how do they affect job satisfaction? *The British Accounting Review*, 37 (4), 389–413.
Lau, C. M., Wong, K. M., and Eggleton, I. R. C., 2008. Fairness of performance evaluation procedures and job satisfaction: the role of outcome-based and non-outcome-based effects. *Accounting and Business Research*, 38 (2), 121–135.
Lawrence, R. J., Mulaik, S. A., and Brett, J. M., 1982. *Causal Analysis: Assumptions, Models, and Data*. Beverly Hills, CA: Sage.
Lillis, A. M., Malina, M. A., and Mundy, J., 2017. Rendering subjectivity informative: field study evidence of subjectivity and bias mitigation in performance measurement and reward systems. Available from SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2998471.
Liu, Y., Loi, R., and Lam, L. W., 2011. Linking organizational identification and employee performance in teams: the moderating role of team-member exchange. *The International Journal of Human Resource Management*, 22 (15), 3187–3201.
López, O. L., and Hiebl, M. R. W., 2015. Management accounting in small and medium-sized enterprises: current knowledge and avenues for further research. *Journal of Management Accounting Research*, 27 (1), 81–119.
MacCrimmon, K. R., and Wehrung, D. A., 1985. A portfolio of risk measures. *Theory and Decision*, 19 (1), 1–29.
Mahoney, T. A., Jerdee, T. H., and Carroll, S. J., 1965. The job(s) of management. *Industrial Relations*, 4 (2), 97–110.
Manthei, K., and Sliwka, D., 2019. Multitasking and subjective performance evaluations: theory and evidence from a field experiment in a bank. *Management Science*, 65 (12), 5861–5883.
Martin, R., and Edwards, H. R., 2009. Perceived organizational support and organizational identification: an analysis after organizational formation. *Human Resource Management Journal*, 19 (1), 110–118.
Mayer, R. C., and Davis, J. H., 1999. The effect of the performance appraisal system on trust for management: a field quasi-experiment. *Journal of Applied Psychology*, 84 (1), 123–136.
Merchant, K. A., 1989. *Rewarding Results: Motivating Profit Center Managers*. Boston, MA: Harvard Business School Press.
Merchant, K. A., Stringer, C., and Theivananthampillai, P., 2010. Relationships between objective and subjective performance ratings. Working paper series number 17, University of Southern California, Department of Accountancy and Business Law.
Moers, F., 2005. Discretion and bias in performance evaluation: the impact of diversity and subjectivity. *Accounting, Organizations and Society*, 30 (1), 67–80.
Mossholder, N., Bennett, E., Kemery, R., and Wesolowski, M. A., 1998. Relationships between bases of power and work reactions: the mediational role of procedural justice. *Journal of Management*, 24 (4), 533–552.
Neubauer, H., Mayr, S., Feldbauer-Durstmuller, B., and Duller, C., 2012. Management accounting systems and institutionalization in medium-sized and large family businesses: empirical evidence from Germany and Austria. *European Journal of Management*, 12 (2), 41–60.
Nunnally, J. C., 1978. *Psychometric Theory*. 2nd ed. New York: McGraw-Hill.
Pennings, J. M. E., and Smidts, A., 2000. Assessing the construct validity of risk attitude. *Management Science*, 46 (10), 1337–1348.
Perren, L., and Grant, P., 2000. The evolution of management accounting routines in small businesses: a social construction perspective. *Management Accounting Research*, 11 (4), 391–411.

Podsakoff, P. M., and Organ, D. W., 1986. Self-reports in organizational research: problems and prospects. *Journal of Management*, 12 (4), 531–544.
Podsakoff, P. M., MacKenzie, S. B., Lee, J., and 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.
Prendergast, C., and Topel, R., 1993. Discretion and bias in performance evaluation. *European Economic Review*, 37 (2-3), 355–365.
Prendergast, C., and Topel, R., 1996. Favoritism in organizations. *Journal of Political Economy*, 104 (5), 958–978.
Rajan, M., and Reichelstein, S., 2009. Objective versus subjective indicators of managerial performance. *The Accounting Review*, 84 (1), 209–237.
Riketta, M., 2005. Organizational identification: a meta-analysis. *Journal of Vocational Behavior*, 66 (2), 358–384.
Roszkowski, M. J., and Soven, M., 2010. Shifting gears: consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment & Evaluation in Higher Education*, 35 (1), 113–130.
Saridakis, G., Torres, R. M., and Johnstone, S., 2013. Do human resource practices enhance Organizational commitment in SMEs with Low employee satisfaction? *British Journal of Management*, 24 (3), 445–458.
Sholihin, M., and Pike, R., 2009. Fairness in performance evaluation and its behavioural consequences. *Accounting and Business Research*, 39 (4), 397–413.
Speklé, R. F., and Widener, S. K., 2018. Challenging issues in survey research: discussion and suggestions. *Journal of Management Accounting Research*, 30 (2), 3–21.
Talacchi, S., 1960. Organization size, individual attitudes and behavior: an empirical study. *Administrative Science Quarterly*, 5 (3), 398–420.
Tett, R. P., and Meyer, J. P., 1993. Job satisfaction, organizational commitment, turnover intention and turnover: path analyses based on meta-analytic findings. *Personnel Psychology*, 46 (2), 259–293.
Tziner, A., and Latham, G., 1989. The effects of appraisal instrument, feedback and goal-setting on worker satisfaction and commitment. *Journal of Organizational Behavior*, 10 (2), 145–153.
Van der Stede, W. A., Chow, C. W., and Lin, T. W., 2006. Strategy, choice of performance measures, and performance. *Behavioral Research in Accounting*, 18 (1), 185–205.
Van Dick, R., Grojean, M. W., Christ, O., and Wieseke, J., 2006. Identity and the extra mile: relationships between organizational identification and organizational citizenship behaviour. *British Journal of Management*, 17 (4), 283–301.
Van Rinsum, M., and Verbeeten, F. H. M., 2012. The impact of subjectivity in performance evaluation practices on public sector managers' motivation. *Accounting and Business Research*, 42 (4), 377–396.
Walumbwa, F. O., Avolio, B. J., and Zhu, W., 2008. How transformational leadership weaves its influence on individual job performance: the role of identification and efficacy beliefs. *Personnel Psychology*, 61 (4), 793–825.
Wapshott, R., and Mallett, O., 2015. *Managing Human Resources in Small and Medium-Sized Enterprises: Entrepreneurship and the Employment Relationship*. London: Routledge.
Woods, A., 2012. Subjective adjustments to objective performance measures: the influence of prior performance. *Accounting, Organizations and Society*, 37 (6), 403–425.
Wright, B. E., 2004. The role of work context in work motivation: a public sector application of goal and social cognitive theories. *Journal of Public Administration Research and Theory*, 14 (1), 59–78.
Zábojník, J., 2014. Subjective evaluations with performance feedback. *The RAND Journal of Economics*, 45 (2), 341–369.

Appendix A. Measurement instrument

The presence of the subjective performance evaluation *(Yes / No)*
In the performance evaluation of managers there may be objective elements (quantified performance measures) as well as subjective elements (the evaluator's personal judgment is used, for example, in the evaluation of qualitative criteria, in the choice of performance measures, or in the weights assigned to them).
In the last evaluation event, did your supervisor use subjective elements to evaluate your performance?

The weight of the subjective performance evaluation *(from 0 to 100%)*
What was the weight of the subjective elements in the overall evaluation of your performance?

Family ownership *(from 0 to 100%)*
Indicate the percentage of the company's equity held by family members.

Feedback quality *(1 = I completely disagree, 5 = I completely agree)*
Indicate your level of agreement with the statements:

Item	Label
My supervisor gives me useful feedback about my job performance.	fbq1
The performance information I get from my supervisor is generally not very meaningful. (R)	fbq2
The feedback I receive from my supervisor helps me do my job.	fbq3
I value the feedback I receive from my supervisor.	fbq4

Trust in supervisor *(1 = I completely disagree, 5 = I completely agree)*
Indicate your level of agreement with the statements:

Item	Label
My supervisor will always act in my favour if he has the chance.	trust1
I am convinced that my supervisor will always fully and honestly keep me up to date of everything that is important to me.	trust2
If my supervisor takes a decision that is against my interest, I am convinced that this decision is justified for other reasons.	trust3

Turnover intention *(1 = Strongly disagree, 7 = Strongly agree)*
Indicate your level of agreement with the following statements:

Item	Label
I often think about quitting.	turnov1
I will probably look for a new job in the next year.	turnov2
(1 = Not at all likely, 7 = Extremely likely)	
How likely is it that you will actively look for a new job in the next year?	turnov3

Organisational identification *(1 = No, I strongly disagree, 7 = Yes, I strongly agree)*
Indicate your level of agreement with the following statements:

Item	Label
I am quite proud to be able to tell people who it is I work for.	ident1
I feel myself to be part of the organisation.	ident2
I would not recommend a close friend to join our staff. (R)	ident3

Managerial performance *(1 = Well below average, 7 = Well above average)*
Indicate to what extent your performance, compared with other managers in your company in similar positions, was below or above the average in the following dimensions:

Item	Label
Planning: determining goals, policies, and courses of action such as work scheduling, budgeting, and programming.	perf1
Investigating: collecting and preparing of information usually in the form of records, reports, and accounts (measuring output, record keeping, and job analysis).	perf2
Coordinating: exchanging information with people in the organisation other than my subordinates in order to relate and adjust procedures, policies, and programs.	perf3
Evaluating: assessment and appraisal of proposals or of reported/ observed performance (e.g. employee appraisals, judging financial performance, and product inspection).	perf4
Supervising: directing, leading, and developing your subordinates.	perf5
Staffing: maintaining the work force of your responsibility area (e.g. selecting and promoting your subordinates).	perf6
Overall, how do you rate your performance?	perf7

Objective performance measures (from 0 to 100%)
Indicate for each of the performance measures below the percentage attached to your variable compensation, indexed to performance.

Performance measure	%
Net income	
Sales	
Sales growth	
EBITDA – Earnings before interest, taxes, depreciation, and amortization	
Operating income	
EVA – Economic value added	
Budget	
Cash-flow	
ROA – Return on assets	
ROI – Return on investment	
ROE – Return on equity	
Market share	
Customer satisfaction	
No. of customer complains	
Quality of the product/service	
Employee satisfaction	
Rotation of personnel	
Production volume	
Productivity	
Compliance with the processes	
Innovation	
Other metrics	

Risk Attitude *(1 = Strongly disagree, 7 = Strongly agree)*
Indicate your level of agreement with the following statements:

Item	Label
I am willing to take high financial risks in order to realise higher average yields.	risk1
I like taking big financial risks.	risk2
In my professional life, I am willing to take risks in order to realise higher average yields.	risk3