Jurnal Economia, Vol. 20, No.3, October 2024, 328-347 P-ISSN: 1858-2648
Website: https://journal.uny.ac.id/index.php/economia E-ISSN: 2460-1152

A Systematic Literature Review of Subjective Performance Evaluation Research Over Three Decades

Afrida Putritama [1*], Sony Warsono [2]
[1,2]Department of Accounting, Faculty of Economics and Business, Universitas Gadjah Mada, Yogyakarta, Indonesia
[1]Department of Accounting, Faculty of Economics and Business, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia
[1]aputritama@uny.ac.id, [2]swarsono@ugm.ac.id
*Corresponding Author

Abstract
Subjective performance evaluation plays a crucial role in accounting research, leveraging human judgments or perceptions to assess employees' or managers' performance. This study performs a systematic literature review of 44 articles on subjective performance evaluation published in leading journals over the past three decades. It examines trends, research methods, and the geographical contexts of studies within this field. Additionally, the study outlines significant findings and implications, proposing directions for future research. By offering a comprehensive update on subjective performance evaluation research, it identifies existing research gaps and potential areas for further investigation. The study advocates for interdisciplinary collaboration and methodological diversity, aiming to spark increased interest and understanding of this complex, relevant topic.

Keywords: Subjective Performance Evaluation, Literature Review, Accounting Research

Tinjauan Literatur Sistematis Penelitian Evaluasi Kinerja Subjektif Selama Tiga Dekade

Abstrak
Evaluasi kinerja subyektif merupakan topik penting dalam penelitian akuntansi, karena melibatkan penggunaan penilaian atau persepsi manusia untuk menilai kinerja karyawan atau manajer. Studi ini melakukan tinjauan literatur sistematis terhadap 44 artikel tentang evaluasi kinerja subjektif yang diterbitkan di jurnal terkemuka dalam tiga dekade terakhir, dan menganalisis tren dan perubahan dalam topik, metode penelitian, dan negara lokasi penelitian evaluasi kinerja subjektif tersebut. Studi ini juga membahas temuan utama dan implikasi penelitian evaluasi kinerja subjektif, dan menyarankan beberapa arah penelitian di masa depan. Studi ini berkontribusi pada literatur dengan memberikan gambaran komprehensif dan terkini mengenai penelitian evaluasi kinerja subjektif, dengan menyoroti kesenjangan penelitian saat ini dan peluang untuk eksplorasi lebih lanjut. Studi ini juga mendorong kolaborasi lintas disiplin dan keragaman metodologi dalam penelitian evaluasi kinerja subjektif, dan berharap dapat merangsang lebih banyak minat dan perhatian terhadap topik yang relevan dan kompleks ini.

Kata Kunci: Evaluasi Kinerja Subjektif, Tinjauan Literatur, Penelitian Akuntansi

History: Received: 27 July 2023 Revised: 16 February 2024 Accepted: 13 May 2024
Citation (APA 6th): Putritama, A., & Warsono, S. (2024). A Systematic Literature Review of Subjective Performance Evaluation Research Over Three Decades. *Jurnal Economia,* 20 (3), 328-347. https://doi.org/10.21831/economia.v20i1.64496

INTRODUCTION

Performance evaluation, a fundamental aspect of human resource practices within organizations, intertwines with complex social contexts. It is integral to numerous employment practices, including compensation, promotion, and training, due to its association with organizational accountability (Ferris et al., 2008). The appeal of multi-attribute performance evaluation tools lies in their ability to offer numerous benefits, prompting research into biases present in performance evaluation (Liedtka et al., 2008a; Dilla & Steinbart, 2005). However, the lack of explicit weights for each measure often leads evaluators to subjectively merge various measures to glean relevant insights, given that subjectivity is a predominant feature of most performance appraisal systems (Kaplan & Norton, 1996; Gibbs et al., 2004; Johnson et al., 2014).

The significance of this systematic literature review on subjective performance evaluation is multifold. Firstly, it analyzes trends and future research directions within the scant number of articles on this topic in premier accounting journals. Secondly, it identifies existing research gaps and opportunities, emphasizing subjective performance evaluation's vital role in accounting research. This importance stems from its predictive value for employee promotions and performances across all job levels, its key role in incentive alignment, interest harmonization, and risk reduction for employees. Thirdly, the study promotes interdisciplinary collaboration between accounting and psychology academics, recognizing the intricate social, emotional, cognitive, and political factors influencing the evaluation process and outcomes. Lastly, it offers fresh perspectives and recommendations to energize future research on subjective performance evaluation. This endeavor aims to enhance the development of equitable, effective performance measurement systems that harmonize objective and subjective criteria.

Subjective performance evaluation plays a pivotal role across various company contexts, underscoring the importance for researchers to discern which factors enhance or diminish its effectiveness (Taylor & Yildirim, 2011). In accounting research, subjective performance evaluation is critical not only for predicting an employee's future promotion and performance but also because it pervades all job levels within an organization, significantly influencing incentives, fostering alignment of interests between employees and the firm, and mitigating employee risk (Hao, 2021). The subjectivity in performance appraisals is highly valued for enabling evaluators to consider all pertinent information about an employee's performance throughout their contract period (Du et al., 2018).

Subjective performance measures allow companies to integrate any performance-related information into their evaluations. This flexibility enables the discretionary adjustment of objective performance measures and grants evaluators the authority to determine the appropriate weighting for each measure (Gibbs et al., 2004; Bailey et al., 2011). However, reliance on subjective evaluation processes can exacerbate bias, as documented by previous studies (Kaplan et al., 2018; Kaplan et al., 2012; Liedtka et al., 2008a, Lipe & Salterio, 2000; Ghosh, 2005).

This literature review addresses two main research questions. First, what trends have emerged in subjective performance evaluation concerning topics, research methods, and country settings?. Second, what future research directions exist for subjective performance evaluation? Aimed at summarizing published research in leading journals, this review intends to offer fresh insights that will foster future investigations. By encouraging interdisciplinary collaboration between management accounting and psychology academics, this review aims to broaden the understanding and application of subjective performance evaluation. The structure of this literature review article is as follows: it begins with the methodology, followed by an analysis and discussion, and concludes with the final observations.

METHOD

This study identified and analyzed 44 articles from leading journals in the Scopus database up to February 16, 2024, utilizing a systematic literature review method. This approach involved organising methods, theories, and constructs into tables or figures, facilitating the reader's comprehension of the valuable insights from the data presented (Paul and Criado, 2020). The research methodology unfolds in three primary stages, as depicted in Figure 1: 1) Stage 1—defining the research questions and objectives; 2) Stage 2—developing the research design; and 3) Stage 3—analyzing the results.

Stage 1 entailed outlining the research questions and objectives, as detailed in the introduction section. Stage 2 focused on creating a research design, beginning with selecting the Scopus database for its extensive repository of high-quality articles. The selection criteria included the use of specific keywords: TITLE-ABS-KEY (“subjective AND performance AND evaluation") AND (LIMIT-TO (SRCTYPE, "j")) AND (LIMIT-TO (PUBSTAGE, "final")) AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (SUBJAREA, "BUSI")) AND (LIMIT-TO(LANGUAGE, "English")). The search strategy incorporated five filters to refine the search results: 1) The subject area must be Business, Management, and Accounting; 2) The document type must be an article; 3) The publication stage must be final; 4) Sources must be journals; and 5) The document must be in English. Following these criteria, the study successfully identified 44 relevant articles. Stage 3 involved a detailed analysis of these articles to address the research questions.

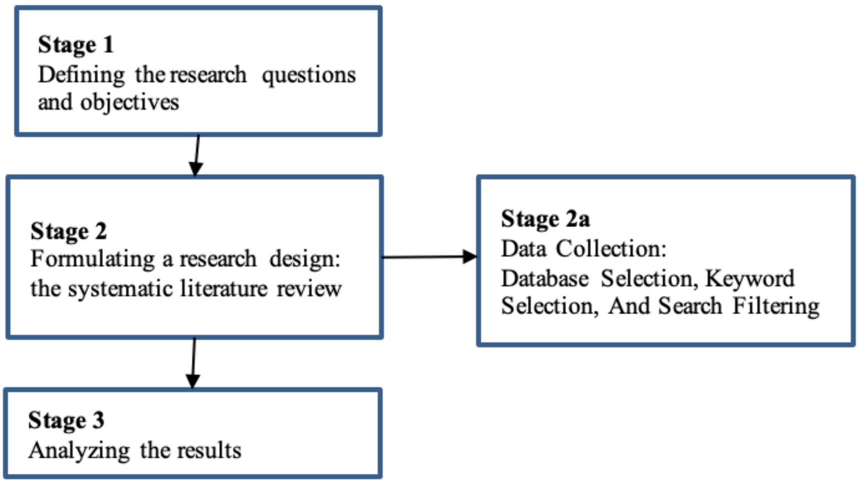

Figure 1. Research Methodology

FINDING AND DISCUSSION

Analysis

This section delves into the findings from the systematic literature review on subjective performance evaluation research, concentrating on three pivotal aspects: topics, research methods, and geographical settings. The array of topics within this field is broad, encompassing issues related to the design, implementation, and impacts of subjective performance evaluation systems. Certain topics have gained prominence in specific periods or within particular journals, reflecting the progression and spread of research on subjective performance evaluation over time and through various publications.

The distribution of articles by research topic over the past three decades (1996-2005, 2006-2015, and 2016-2024) is illustrated in Table 1. This table organizes the articles into nine distinct topics, revealing a trend of increasing publication volume on subjective performance evaluation, from eight articles between 1996-2005 to 24 from 2016-2024. This uptick signifies an escalating interest and importance of this topic within the accounting discourse. Furthermore, Table 1 highlights shifts in the research themes over time, providing valuable insights into the evolution and current research lacunae within subjective performance evaluation studies. Notably, recent years have seen the emergence or increased popularity of topics like team bias and leniency bias.

According to the data in Table 1, the juxtaposition of subjective versus objective performance evaluation is the most frequent topic, covered in 15 articles over the last three decades. This prevalence underscores the significant interest in exploring the merits and drawbacks of subjective versus objective metrics for appraising the performance of employees or managers. Subjective measures rely on human judgment or perceptions, whereas objective metrics are grounded in tangible data or formulas.

Following closely, the topic of compensation contracting has been addressed in seven articles. This area examines the structure and execution of compensation agreements between principals (e.g., shareholders or owners) and agents (e.g., managers or employees), aiming to synchronize their interests and incentivize the attainment of organizational objectives. Compensation contracts might encompass various incentive forms, such as salaries, bonuses, stock options, or promotions, based on various performance indicators, including financial, non-financial, objective, or subjective metrics.

The managerial discretion, cognitive bias, team bias, leniency bias, and negativity bias topics were also interesting topics. The topic of managerial discretion delves into the latitude managers possess in decision-making processes, such as resource allocation, performance evaluation, and employee compensation. This discretion is influenced by task nature, information availability, organizational structure, and external environment. Cognitive bias, another key area, refers to the systematic departures from rationality or objectivity in human cognition and decision-making, attributable to limitations in information processing, memory, attention, or motivation. These biases can impact managers' or supervisors' perceptions, interpretations, and evaluations of employee or managerial performance, influencing the allocation of rewards or penalties.

Table 1. *Frequency Distribution of Articles by Research Topic*

Frequency distribution of articles by topics	1996-2005	2006-2015	2016-2024	Total
Subjective performance evaluation versus objective performance evaluation: Kaplan & Norton (1996), Dilla & Steinbart (2005), Ghosh (2005b), Liedtka et al. (2008), Hartmann et al. (2010), Ahn et al. (2010), Maas et al. (2012), Dai et al. (2018), Fehrenbacher et al. (2018), Murphy (2020), Chen et al. (2020), Jayaraman et al. (2021), Maske et al. (2021), Alves & Lourenço (2023), Budde (2023)	3	4	8	**15**
Managerial discretion: Bailey et al. (2011), Taylor & Yildirim (2011), Johnson et al. (2014), Castro (2017), Tran & Jarvinen (2022)	-	3	2	**5**
Team Bias: Arnold et al. (2018), Mursita and Nahartyo (2021), Uribe et al. (2022)	-	-	3	**3**
Negativity bias: Baumeister et al. (2001), Kaplan et al. (2012), Kaplan et al. (2018)	1	1	1	**3**
Calibration Committees: Long et al. (2015), Demeré et al. (2019)	-	1	1	**2**
Cognitive bias: Nixon (1998), Bol & Smith (2011), Lipe & Salterio (2000), Cassar and Ko (2023)	2	1	1	**4**
Compensation contracting: Bol (2008), Cheng (2021), Gibbs et al. (2004), Lillis et al. (2022), Prendergast (1999), Mitsuhashi and Ohta (2024), Hao (2021)	2	1	4	**7**
Leniency Bias: Gong et al. (2021), Maas & Verdoorn (2017), Maske & Sohn (2023)	-	-	3	**3**
Escalation Bias: Angelovski et al. (2016), Kramer & Maas (2020)	-	1	1	**2**
Total	**8**	**12**	**24**	**44**

Research on team bias examines the propensity of managers or supervisors to favor or disfavor certain employees or groups based on personal or social connections rather than actual performance. This bias can compromise the fairness and accuracy of performance evaluations and compensation systems, with potential positive or negative effects on employee motivation and behavior. The leniency bias topic investigates the tendency among managers or supervisors to assign higher performance ratings than the actual performance merits, driven by social pressure, personal preference, or information scarcity. Such bias undermines the validity and reliability of subjective performance evaluations, adversely affecting employee motivation and behavior. Negativity bias research explores the phenomenon where negative information or events disproportionately influence human cognition and decision-making compared to positive ones. This bias can alter managers' or supervisors' performance evaluations and employees' or managers' responses to feedback or incentives.

Table 2. *Frequency Distribution of Articles by Research Method*

Frequency distribution of articles by topics	1996-2005	2006 -2015	2016 -2024	Total
Experiment	3	7	9	**19**
Lipe & Salterio (2000), Ghosh (2005), Dilla & Steinbart (2005), Liedtka et al. (2008), Bailey et al. (2011), Bol & Smith (2011), Maas et al. (2012), Kaplan et al. (2012), Johnson et al. (2014), Long et al. (2015), Angelovski et al. (2016), Maas & Verdoorn, (2017), Arnold et al. (2018), Fehrenbacher et al. (2018), Kaplan et al. (2018), Kramer & Maas (2020), Maske et al. (2021), Mursita and Nahartyo (2021), Maske & Sohn (2023)				
Survey	1	1	3	**5**
Gibbs et al. (2004), Hartmann et al. (2010), Castro (2017), Alves & Lourenço (2023), Tran & Jarvinen (2022)				
Archival	-	1	5	**6**
Ahn et al. (2010), Chen et al. (2019), Wang & Xue (2020), Hao (2021), Jayaraman et al. (2021), Gong et al. (2021)				
Interview	-	-	2	**2**
Lillis et al. (2022), Uribe et al. (2022)				
Literature Review	2	1	1	**4**
Prendergast (1999), Baumeister et al. (2001), Bol (2008), Murphy (2020)				
Theoretical Approach	-	1	4	**5**
Taylor & Yildirim (2011), Dai et al. (2018), Cheng (2021), Budde (2023), Mitsuhashi & Ohta (2024)				
Case Study	2	-	-	**2**
Kaplan & Norton (1996), Nixon (1998)				
Field Study	-	-	1	**1**
Cassar & Ko (2023)				
Total	**8**	**11**	**25**	**44**

Less frequently discussed topics in subjective performance evaluation research include calibration committees and escalation bias. Calibration committees, groups tasked with reviewing and adjusting subjective performance ratings to ensure consistency and accuracy across organizational units or levels, can mitigate biases and errors in performance evaluations, enhancing supervisory and employee learning and feedback processes. Escalation bias investigates the biased evaluation of employees or managers based on hiring decisions, such as assigning higher ratings to recommended individuals or lower ratings to those rejected yet still hired. Stemming from self-justification or commitment to prior decisions, escalation bias can compromise the fairness and accuracy of subjective performance evaluations.

Table 2 provides an insightful overview of the methodologies employed in 44 articles on subjective performance evaluation across three distinct periods: 1996-2005, 2006-2015, and 2016-2024. Predominantly empirical, the research methods reveal a strong preference for experimental approaches, with 19 articles utilizing this method, followed by archival, survey, and theoretical approaches. Field studies were the least common, featured in only one article. The data also indicate evolving trends in methodological preferences over time. The use of experimental methods saw an increase from three articles in the 1996-2005 period to nine articles in 2016-2024, while archival methods rose from none in 1996-2005 to five in 2016-2024. Conversely, the application of case study methods declined, from two articles in the earlier period to none in the most recent. These trends suggest that while various research methods apply to subjective performance evaluation studies, experimental approaches have become increasingly predominant over the last three decades.

Table 3. *Frequency Distribution of Articles by Country Setting*

Frequency distribution of articles by topics	**1996-2005**	**2006-2015**	**2016-2024**	**Total**
United States of America	4	6	7	17
Kaplan & Norton (1996), Gibbs et al. (2004), Dilla & Steinbart (2005),, Ghosh (2005), Liedtka et al. (2008), Bailey et al. (2011), Bol & Smith (2011), Kaplan et al. (2012), Johnson et al. (2014), Long et al. (2015), Dai et al. (2018), Arnold et al. (2018), Kaplan et al. (2018), Demeré et al. (2019), Jayaraman et al. (2021), Cheng (2021), Uribe et al. (2022)				
China	-	-	2	2
Chen et al. (2020), Gong et al. (2021)				
Taiwan	-	-	1	1
Hao (2021)				
Korea	-	1	-	1
Ahn et al. (2010)				
Spain	-	-	1	1

Angelovski et al. (2016)				
Indonesia	-	-	1	1
Mursita and Nahartyo (2021)				
Singapore	-	-	1	1
Cassar and Ko (2023)				
Vietnam	-	-	1	1
Tran & Jarvinen (2022)				
United Kingdom	1	-	-	1
Nixon (1998)				
Netherlands	-	2	-	2
Hartmann et al. (2010), Maas et al. (2012)				
Portugal	-	-	1	1
Alves & Lourenço (2023)				
West Europe	-	-	2	2
Maas, & Verdoorn (2017), Kramer, & Maas (2020)				
Australia	-	-	2	2
Castro (2017), Fehrenbacher et al. (2018)				
German	-	-	2	2
Maske et al. (2021), Maske & Sohn (2023)				
Various countries	2	-	2	4
Lipe & Salterio (2000), Baumeister et al. (2001), Murphy (2020), Lillis et al. (2022)				
Not specify a particular country setting	1	2	2	5
Prendergast (1999), Bol (2008), Taylor & Yildirim (2011), Budde (2023), Mitsuhashi and Ohta (2024)				
Total	**8**	**11**	**25**	**44**

Table 3 delineates the geographic distribution of the 44 subjective performance evaluation studies, highlighting a significant concentration in the United States, followed by China, Western Europe, Germany, the Netherlands, Australia, Indonesia, Taiwan, Korea, Spain, Singapore, Vietnam, the United Kingdom, and Portugal. It also notes that some studies encompass multiple countries or do not specify a particular country setting. This distribution underscores that subjective performance evaluation is a globally relevant research topic, albeit with a pronounced focus on the United States. The varied country settings indicate the topic's applicability and interest across different cultural and business environments, suggesting its universal significance and potential for cross-cultural and international research exploration.

Discussion

This discussion synthesizes findings from several pivotal studies on subjective performance evaluation, highlighting their contributions and implications for the field. Ahn et al. (2010) investigated the concept of performance measure discriminability—how well performance ratings differentiate between high and low performers—within the context of employees evaluated by multiple measures. Their comparison of subjective versus objective performance measures revealed that subjective measures tend to have lower discriminability and provide less potent incentives than their objective counterparts. This insight underscores the potential limitations of subjective evaluations in effectively motivating employee performance enhancement. Meanwhile, Angelovski et al. (2016) delved into the impact of hiring decisions on subjective performance evaluations, particularly examining the presence of escalation bias. This bias manifests as a tendency to favorably evaluate employees recommended by the evaluator or unfavorably assess those rejected but nevertheless hired. The study found that such biases are not only present but are also shaped by factors like incentives, managerial experience, and the manager's performance in tasks similar to those of the evaluated employees. These findings suggest the need for awareness and mitigation strategies against biases in performance evaluations.

Arnold et al. (2018) investigated the impact of subjective communication among team members and the diversity of their abilities—referred to as ability heterogeneity—on overall team performance. Subjective communication is characterized as the unverifiable information shared by team members with their manager regarding individual contributions to team outcomes. Ability heterogeneity, meanwhile, denotes the range of productivity levels across team members or the varying extents of effort they contribute to team output. The study discovered that subjective communication generally enhances team performance. However, this positive impact is less pronounced in teams with high ability heterogeneity. The reason for this diminished effect is attributed to divergent perceptions of fairness in bonus distribution among team members with varied abilities, leading to inconsistent communication and causing managers to deviate from relying solely on subjective input.

Bailey et al. (2011) focused on how the degree of managerial discretion in bonus pool allocation influences the incorporation of subjective information. The study outlined two primary approaches for allocating bonus pools: the integrative approach, which amalgamates contractible (verifiable) and noncontractible (non-verifiable) information into a unified performance metric, and the piecemeal approach, which assesses the impact of each type of information on compensation independently. The research also examined how different levels of discretion (full versus partial) affect the usage of subjective information in the allocation process. It was found that partial discretion more effectively communicates the intent behind using discretion, primarily viewed as noncontractible information, thereby heightening the importance of such information. The study observed a tendency among managers to favor an anchoring approach, wherein those employing it integrate noncontractible information to a lesser extent than those adopting the integrative method. However, this inclination can be counteracted by granting managers partial discretion. Additionally, the findings indicated that managers with partial discretion are better at reflecting noncontractible information in their overall bonus pool allocations compared to those with full discretion, highlighting how varying levels of discretion can significantly impact the relevance and treatment of subjective information in compensation decisions.

Baumeister et al. (2001) delve into a wide-ranging exploration of why negative experiences, emotions, and outcomes exert a more potent and enduring influence than positive ones, drawing evidence across several research domains. The study offers multiple explanations for this phenomenon, such as the heightened survival threat posed by negative occurrences, the increased attention and cognitive processing dedicated to negative events, and the asymmetric nature of positive and negative feedback loops. It further examines the implications and applications of this principle in various fields, including clinical psychology, interpersonal relationships, social behavior, and organizational behavior, providing a comprehensive overview of how the negativity bias impacts various aspects of human life and society.

Bol (2008) presents a literature review focusing on the role of subjectivity in compensation contracting. The review compiles and analyzes findings from prior studies on the motivations for, designs of, and outcomes resulting from incorporating subjective performance measures, weights, or adjustments into incentive schemes. The review identifies several factors that influence the optimal level and form of subjectivity in these contexts, such as the availability of information, the preferences and incentives of shareholders and managers, and the costs associated with verification and litigation. The review also highlights potential avenues for future research, including the impact of subjectivity on management control systems, its relationship with organizational culture, and the influence of psychological and behavioral factors on the application of subjectivity. The review concludes by emphasizing the significance and complexity of subjectivity in compensation contracting, advocating for increased attention from both accounting researchers and practitioners.

Dai (2021) explores the design of optimal contracts within teams, particularly when team performance evaluation hinges on both objective outputs and subjective assessments based on private signals regarding individual and peer contributions. The analysis reveals that as evaluations grow more subjective, the sensitivity of workers' wages to both their own assessments and the overall team output should decrease. This finding suggests that objective outputs and subjective evaluations act as complements rather than substitutes within incentive schemes. Additionally, the study indicates that subjective evaluations are most effective in relative terms when the degree of subjectivity is maintained within a moderate range, shedding light on the nuanced role of subjectivity in shaping effective incentive mechanisms.

On the other hand, Bol and Smith (2011) delved into the dynamics between the level and controllability of objective measures for one task and their influence on the subjective evaluation of another task. Drawing on theories from psychology, organizational behavior, and economics, the study explored the concept of spillover effects—where the evaluation of one task is influenced by performance in another, unrelated task. The research discovered that subjective performance evaluations often mirror objective measures, regardless of their relevance to the performance dimension being subjectively assessed. Moreover, Bol and Smith (2011) observed that the nature of these spillover effects varies with the controllability of the objective measure. Specifically, when objective measures are influenced by factors beyond an individual's control, managers tend to adjust their evaluations to account for bad luck but refrain from penalizing good luck. This study enriches the accounting literature and practice by illustrating the complex interplay between objective and subjective performance measures within a multitask environment.

Cheng (2021) investigated how evaluators, particularly those familiar with the Balanced Scorecard (BSC), employ both common and unique measures within the BSC framework to assess performance and allocate bonuses across two business divisions. The study also examined the evaluators' propensity to prioritize negative over equally significant positive information—a phenomenon known as negativity bias—and how this bias is affected by the evaluator's role and the significance of the performance measure reporting a negative result. Cheng (2021) found that managers with mixed performance outcomes tend to de-emphasize crucial performance measures with negative results compared to supervisors and managers with consistently positive outcomes. This suggests that a self-enhancement motivation, limited by negativity bias, is at play. Specifically, negativity bias is more pronounced when a strategically significant performance measure underperforms compared to when a less critical measure does. This insight into the interplay between evaluator roles, the importance of performance measures, and the presence of negativity bias offers valuable nuances to the understanding of performance evaluation processes.

Meanwhile, Fehrenbacher et al. (2018) investigated the causes and implications of compression bias—a tendency among raters to avoid using extreme ratings—demonstrating how this bias diminishes the discriminability of subjective measures and adversely affects employee motivation to improve performance. Utilizing a two-level within-subject experimental design, the study differentiated between independent variables (division's relative performance on common and unique Balanced Scorecard measures) and dependent variables (performance evaluations and bonus allocations for division managers). The findings indicated that compression bias is more evident when managers rely on intuitive rather than deliberate decision-making processes. Furthermore, this bias is more significant for unique measures compared to common measures. The study also highlighted the detrimental effects of compression bias on employee satisfaction and effort, mediated by perceptions of evaluation fairness.

Maas et al. (2012) executed two experiments to examine the impact of superiors' performance evaluation behaviors on subordinates' work-related attitudes. The research revealed that specifying subjective evaluation criteria in advance positively influences goal clarity but only in highly formalized settings. Conversely, subjective evaluations conducted after the fact positively affect perceptions of evaluation fairness, particularly in roles characterized by high job autonomy. The study did not find that formalization moderated the relationship between ex post subjective evaluation and perceived evaluation fairness, arguing instead that ex ante specified subjective evaluations improve goal clarity by aligning the evaluation process with a formalized environment. On the other hand, ex post subjective evaluations enhance perceptions of fairness by fostering a social exchange dynamic between superiors and subordinates within an environment that values autonomy.

Dilla and Steinbart (2005) revealed that decision-makers tend to prioritize common BSC measures over unique ones, independent of their relevance to the performance dimension under evaluation. The study also found that decision-makers are swayed by the performance indicated by common measures, even when these measures do not directly relate to the performance dimension being assessed. This suggests a potential oversight in the nuanced use of performance measures, emphasizing the influence of common measures in decision-making processes. In addition, Chen et al. (2020) discovered that firms adjust the emphasis placed on objective versus subjective performance measures for top management and middle managers based on a variety of factors, including competitive intensity, environmental uncertainty, and CEO power. Their research indicates that mismatches in these adjustments can lead to higher management turnover, underscoring the significance of appropriate measure tailoring in enhancing manager satisfaction and retention. This finding highlights the complex interplay between performance measurement and organizational dynamics, pointing to the strategic importance of tailored performance evaluation systems.

Ghosh (2005) identified that managers employ alternative performance measures to counteract the outcome effect—the tendency to base performance evaluations solely on the end result, disregarding uncontrollable factors. The study further observed that the reliance on alternative measures grows with the level of uncontrollability and the criticality of the performance dimension. This underscores the importance of comprehensive evaluation approaches that consider factors beyond direct outcomes. Meanwhile, Gong et al. (2021) explored how contextual uncertainty augments leniency bias and how past employee performance affects this bias. Their findings suggest that leniency bias can demotivate employees, diminishing effort and productivity, and can lead to increased turnover. This insight underscores the detrimental impact of leniency bias on organizational effectiveness and employee engagement. On another note, Demeré et al. (2019) investigated the role of calibration committees in performance evaluations. They found that such committees enhance rating consistency across supervisors and mitigate leniency bias, albeit at the expense of introducing a central tendency bias. The research also noted that calibration committees acknowledge supervisors' informational advantage in rating adjustments, thereby facilitating supervisor learning regarding organizational expectations for performance ratings. This study highlights the nuanced benefits and challenges of using calibration committees to improve the objectivity and fairness of performance evaluations.

Alves and Lourenço (2023) delve into the effects of subjective performance evaluation (SPE) on financial reporting quality and audit fees, discovering a positive correlation between SPE and higher earnings quality alongside lower audit fees. This indicates that SPE can enhance the reliability and credibility of financial information, serving as a valuable tool in financial reporting and auditing processes. Meanwhile, Budde (2023) investigates the impact of SPE on employee motivation and performance within a public sector context. The study finds that while SPE positively influences employee motivation, it paradoxically has a negative effect on employee performance. This highlights a complex interplay between intrinsic and extrinsic incentives, suggesting a potential trade-off that organizations might face when implementing SPE strategies.

Castro (2017) examines subjective performance evaluation 's role in resource and reward allocation within multi-divisional firms. The findings suggest that SPE is often employed to supplement objective performance measures, allowing for adjustments due to interdependencies and external factors affecting different divisions. Furthermore, SPE's application appears to be influenced by the power and reputation of division managers, pointing to the nuanced dynamics of performance evaluation in complex organizational structures. Contrasting with Castro (2017), Tran & Jarvinen (2022) analyze SPE in the context of Vietnamese rice farming, focusing on its determinants and impacts. The study reveals that SPE is significantly shaped by the social distance and trust between evaluators and evaluatees. Moreover, SPE is found to enhance both productivity and profitability among farmers, indicating its effectiveness beyond conventional corporate settings and highlighting the importance of interpersonal relationships in the evaluation process.

Mursita and Nahartyo (2021) delve into the impact of subjective performance evaluation on employee creativity and innovation within the creative industry. Their findings indicate a positive correlation between subjective performance evaluation and these outcomes, with employee intrinsic motivation and psychological empowerment acting as mediators. This suggests that subjective evaluations, by focusing on individual contributions and potentials, can significantly foster creative and innovative work environments. Uribe et al. (2022) examined the application of subjective performance evaluation in assessing engineering design projects. The study highlights the importance of subjective evaluation in capturing non-technical aspects of design work, such as teamwork, communication skills, and customer satisfaction. Furthermore, the findings suggest that the effectiveness of subjective performance evaluations in this context is influenced by the evaluator's expertise, experience, and feedback style, underscoring the value of nuanced, human-centered assessment methods in technical disciplines.

Kaplan et al. (2018) investigated the relationship between subjective performance evaluation and employee turnover within a large professional services firm. Their research reveals a negative association between subjective performance evaluation and turnover, particularly among high-performing employees. Additionally, the study presents evidence that subjective performance evaluations can moderate the impact of objective performance measures on turnover, suggesting that subjective evaluations may help retain top talent by acknowledging their contributions beyond quantifiable metrics. Conversely, Cassar and Ko (2023) explored the effects of subjective performance evaluation on entrepreneurs in a randomized controlled trial. The study found that subjective performance evaluations not only increase entrepreneurs' effort and persistence but also their risk-taking behaviors and overconfidence. Despite these mixed outcomes, subjective performance evaluation positively impacts the survival and growth of entrepreneurial ventures. This indicates that while subjective evaluations can drive greater commitment and ambitious pursuits, they may also introduce biases toward risk and confidence levels, highlighting the complex implications of evaluation practices on entrepreneurial success.

Mitsuhashi and Ohta (2024) assessed the impact of subjective performance evaluation on health care services' quality and efficiency within a Japanese hospital setting. Their study revealed that such evaluations contribute to enhanced service quality and efficiency, manifested through decreased medical errors, heightened patient satisfaction, and reduced costs. These benefits were partly attributed to the influence of the hospital's organizational culture and leadership. Similarly, Hao (2021) explored the significance of subjective performance evaluations in the governance and performance of family-owned firms in China. His findings indicated that these evaluations help to align the interests and values between family members and non-family managers, effectively addressing potential agency issues and conflicts within family businesses. Furthermore, Hao noted that subjective performance evaluations played a crucial role in improving the performance and ensuring the long-term sustainability of these firms.

Social interaction and informal communication between superiors and subordinates influence their subjective performance evaluations (Du et al., 2012). In diverse teams, the impact of subjective communication among team members is less pronounced (Arnold et al., 2018). The way mixed-gender managers conduct performance evaluations is similar to supervisors, where negative outcomes are linked to less critical metrics (Kaplan et al., 2012). The leniency bias in evaluations is affected by the supervisor's span of control and the non-routine nature of the job, which then differently influences how previous employee performance moderates this bias (Gong et al., 2021). Supervisors tend to make more lenient adjustments when new, unfavorable performance data contrasts with old, favorable data. Instead of simply adjusting scores upwards, they prefer to shift the emphasis in their evaluations from the new, unfavorable measures to the old, favorable ones (Du et al., 2018).

Companies frequently encounter issues with misleading information in subordinate performance evaluations. This misinformation stems from supervisors lacking the necessary assessment skills or intentionally inflating evaluations. To align performance measurements with corporate goals, companies can recommend specific weightings for evaluation criteria while still permitting some degree of subjectivity for factors not explicitly regulated (Long et al., 2015). By defining these weightings, the reliance on subjective judgment for incorporating uncontracted information can be minimized, leading to more significant effects on outcomes. The balance between objective and subjective performance metrics, especially among senior and mid-level managers, is influenced by various factors including market competition, CEO authority, and environmental uncertainty, but is negatively affected by growth prospects and organizational stability (Chen et al., 2020). Managers' control over outcome determinants affects their performance evaluations, whereas uncontrollable environmental factors do not (Ghosh & Lusch, 2000). To incentivize employees, companies may offer performance-based bonuses. However, subjective appraisals may lead employers to unjustly deem an employee's performance as unsatisfactory, thereby withholding bonuses. It is suggested that employees showing consistent performance improvement should receive greater rewards than those with subpar performance (Chan & Zheng, 2011; Jayaraman et al., 2021). The management of risks associated with unexpected behavioral responses to performance measures is crucial in subjective performance evaluations (Lillis et al., 2022). Longer managerial tenure correlates positively with employee satisfaction regarding wage and subjective bonuses (Gibbs et al., 2004).

Managerial heuristic reasoning influences the balance between objective and subjective criteria in performance evaluation decisions (Dai et al., 2018). Subjective evaluations are more favorable when a strong supervisor-subordinate relationship exists, whereas they tend to be less favorable when there is a significant age difference between supervisors and subordinates (Hao, 2021). Managers often rate employees higher when they have a positive impression of them, compared to those they view negatively (Kramer & Maas, 2020). There is a tendency among managers to provide high, yet not widely varied, subjective performance evaluations. This leniency, and the lack of variability (compression), is evident in the observation that performance ratings are higher when criteria are alphabetically listed rather than presented in a four-category balanced scorecard format (Maas & Verdoorn, 2017). Additionally, the reliance on managerial outcome measures intensifies as the impact of outcomes increases, more so for subjective non-financial measures than for financial ones (Ghosh, 2005).

One approach companies can employ to reduce subjective biases in performance evaluations is to empower calibration committees. These committees adjust the performance evaluation process and its impact on compensation to specifically deter opportunistic appraisal behaviors by lowering the ratings of supervisors found to be biased (Grabner et al., 2020). Calibration committees typically reduce the ratings of supervisors who assign subjective evaluations higher than the average initial rating, thereby enhancing the consistency of these evaluations across all supervisors and mitigating leniency bias. However, this approach tends to exacerbate centrality bias, where evaluations cluster around a central point (Demeré et al., 2019). Moreover, requiring supervisors and managers to assess employee behavior, judge performance, reward compliance, and penalize actions can hinder their effectiveness as leaders (Murphy, 2020). Consequently, the accounting profession needs to develop a performance measurement system capable of defining and measuring subjective evaluations with sufficient precision. This would allow subjective and objective appraisal measures to support each other effectively (Ahn et al., 2010).

The analysis provided addresses the second research question on future directions for subjective performance evaluation research. It highlights the need for increased attention to country settings and the influence of cultural, institutional, and environmental factors on the design and effectiveness of such systems. It suggests a call for more cross-country or cross-cultural studies to understand the variations in practices and outcomes of subjective performance evaluations across different contexts. The recommendation extends to diversifying data sources and types, employing rigorous and robust methods for hypothesis testing and result validation. This study also emphasizes the importance of clearly stating the theoretical frameworks of future studies and considering the integration of multiple theories for a deeper, more nuanced understanding of subjective performance evaluation. Additionally, future research should investigate various system aspects, including design alternatives, communication channels, and the complexity of incentives. The role of supervisors in ensuring fairness, the organization of information in performance reports, and its impact on judgment warrant further exploration. The study calls for examining cognitive processes affecting decision-making and strategic actions by supervisors and subordinates over time, the effectiveness of employee rewards, governance in collaboration with market partners, and susceptibility to psychological biases in evaluations. Other recommended areas include the manipulation of performance measurement weights, the influence of contingency factors on employee selection and compensation risk, and the effects of information types and sources on outcomes. The investigation should also cover leadership styles, performance measures, and personality traits like narcissism and Machiavellianism in performance evaluations. Finally, future research should consider how contextual factors like organizational culture, trust, feedback quality, and goal difficulty affect subjective evaluations and procedural justice.

CONCLUSION

This study undertook a systematic literature review of 44 articles focused on subjective performance evaluation, published in leading journals over the last three decades. It examined the evolution and current trends in topics, research methods, and geographical focus within SPE research, presenting key findings and their implications. The study identified a diversification of themes over time, with recent emphasis on biases such as team, leniency, and negativity biases, indicating a shift in research focus towards understanding the nuances of subjective performance evaluation systems.

The predominance of empirical research methods, particularly experiments, was noted, highlighting a methodological preference in the field. Additionally, the study observed a significant concentration of research within the United States, suggesting a potential geographical bias in subjective performance evaluation studies. It pinpointed several research gaps, including the impact of subjectivity in management control systems, the interplay between subjectivity and organizational culture, and the influence of psychological and behavioral factors on subjective evaluations.

By offering a detailed and current synthesis of subjective performance evaluation research, this study contributes valuable insights into ongoing challenges and opportunities for future research. It calls for increased cross-disciplinary collaboration and a broader methodological approach to deepen understanding of subjective performance evaluation. The study aims to spark further interest in subjective performance evaluation, a topic of significant relevance to both accounting theory and practice, encouraging exploration beyond the conventional boundaries of the field.

REFERENCES

Ahn, T. S., Hwang, I., & Kim, M. I. (2010). The impact of performance measure discriminability on ratee incentives. *Accounting Review*, *85*(2), 389–417. https://doi.org/10.2308/accr.2010.85.2.389

Alves, I. & Lourenço. S.M. (2023). Subjective performance evaluation and managerial work outcomes. Accounting and Business Research, 53:2, 127-157, DOI: 10.1080/00014788.2021.1959292

Angelovski, A., Brandts, J., & Sola, C. (2016). Hiring and escalation bias in subjective performance evaluations: A laboratory experiment. Journal of Economic Behavior and Organization, 121, 114–129. https://doi.org/10.1016/j.jebo.2015.10.012

Arnold, M. C., Hannan, R. L., & Tafkov, I. D. (2018). Team member subjective communication in homogeneous and heterogeneous teams. *Accounting Review*, *93*(5), 1–22. https://doi.org/10.2308/accr-52002

Bailey, W. J., Hecht, G., & Towry, K. L. (2011). Dividing the pie: The influence of managerial discretion extent on bonus pool allocation. *Contemporary Accounting Research*, *28*(5), 1562–1584. https://doi.org/10.1111/j.1911-3846.2011.01073.x

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad Is Stronger Than Good. Review of General Psychology, 5(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Bol, J. C. (2008). Subjectivity in Compensation Contracting. *Journal of Accounting Literature* (Vol. 27).

Bol, J. C., & Smith, S. D. (2011). Spillover effects in subjective performance evaluation: Bias and the asymmetric influence of controllability. *Accounting Review*, 86(4), 1213–1230. https://doi.org/10.2308/accr-10038

Budde, J. (2023). The interaction of subjective and objective performance information in fixed payment schemes. *Management Accounting Research, 61* https://doi.org/10.1016/j.mar.2023.100853

Cassar, G. and Ko, T. (2023). Peer effects in subjective performance evaluation. *Contemporary Accounting Research*, 40, 1704–1732.

Castro, Vincent Bicudo de (2017). Unpacking the notion of subjectivity: Performance evaluation and supervisor discretion. *The British Accounting Review*, 49 http://dx.doi.org/10.1016/j.bar.2017.08.003

Chan, J., & Zheng, B. (2011). Rewarding improvements: optimal dynamic contracts with subjective evaluation. *Journal of Economics* (Vol. 42, Issue 4).

Chen, C. X., Gao, Y., Wang, Y., & Xue, S. (2020). Tailoring the weights on objective versus subjective performance measures between top management and middle managers: Evidence from performance-based equity incentive plans. *Journal of Management Accounting Research*, *32*(3), 49–70. https://doi.org/10.2308/jmar-18-042

Cheng, C. (2021). Moral hazard in teams with subjective evaluations. Journal of Economics (Vol. 52, Issue 1). http://www.quora.com/How-are-performance-re

Dai, N. T., Kuang, X. (Jason), & Tang, G. (2018). Differential Weighting of Objective Versus Subjective Measures in Performance Evaluation: Experimental Evidence. *European Accounting Review*, *27*(1), 129–148. https://doi.org/10.1080/09638180.2016.1234402

Demeré, B. W., Sedatole, K. L., & Woods, A. (2019). The role of calibration committees in subjective performance evaluation systems. *Management Science*, *65*(4), 1562–1585. https://doi.org/10.1287/mnsc.2017.3025

Dilla, W. N., & Steinbart, P. J. (2005). Relative Weighting of Common and Unique Balanced Scorecard Measures by Knowledgeable Decision Makers. In *BEHAVIORAL RESEARCH IN ACCOUNTING* (Vol. 17). http://www.bscol.com

Du, F., Erkens, D. H., Young, S. M., & Tang, G. (2018). How adopting new performance measures affects subjective performance evaluations: Evidence from EVA adoption by Chinese State-Owned Enterprises. *Accounting Review*, *93*(1), 161–185. https://doi.org/10.2308/accr-51763

Du, F., Tang, G., & Mark Young, S. (2012). Influence activities and favoritism in subjective Performance Evaluation: Evidence from Chinese state-owned enterprises. *Accounting Review*, *87*(5), 1555–1588. https://doi.org/10.2308/accr-50196

Fehrenbacher, D. D., Schulz, A. K. D., & Rotaru, K. (2018). The moderating role of decision mode in subjective performance evaluation. Management Accounting Research, 41, 1– 10. https://doi.org/10.1016/j.mar.2018.03.001

Ferris, G. R., Munyon, T. P., Basik, K., & Buckley, M. R. (2008). The performance evaluation context: Social, emotional, cognitive, political, and relationship components. *Human Resource Management Review*, *18*(3), 146–163. https://doi.org/10.1016/j.hrmr.2008.07.006

Ghosh, D. (2005). Alternative Measures of Managers' Performance, Controllability, and the Outcome Effect. *Behavioral Research In Accounting* (Vol. 17).

Ghosh, D., & Lusch, R. F. (2000). Outcome effect, controllability and performance evaluation of managers: some field evidence from multi-outlet businesses. *Organizations and Society* (Vol. 25). www.elsevier.com/locate/aos

Gibbs, M., Merchant, K. A., van der Stede, W. A., Vargus, M. E., (2004). Determinants and Effects of Subjectivity in Incentives. The Accounting Review (Vol. 79, Issue 2).

Gong, N., Boh, W. F., Wu, A., & Kuo, T. (2021). Leniency Bias in Subjective Performance Evaluation: Contextual Uncertainty and Prior Employee Performance. *Emerging Markets Finance and Trade*, *57* (8), 2176–2190. https://doi.org/10.1080/1540496X.2019.1660161

Grabner, I., Künneke, J., & Moers, F. (2020). How calibration committees can mitigate performance evaluation bias: An analysis of implicit incentives. *Accounting Review* (Vol. 95, Issue 6, pp. 213–233). American Accounting Association. https://doi.org/10.2308/TAR-2016-0662

Hao, J. Y. P. (2021). Subjective performance evaluation and forward-looking implications: The role of supervisor incentives. *Journal of Management Accounting Research*, *33*(2), 109–127. https://doi.org/10.2308/JMAR-19-023

Jayaraman, S., Milbourn, T., Peters, F., & Seo, H. (2021). Product market peers and relative performance evaluation. *Accounting Review*, *96*(4), 341–366. https://doi.org/10.2308/TAR-2018-0284

Johnson, E. N., Reckers, P. M. J., & Bartlett, G. D. (2014). Influences of timeline and perceived strategy effectiveness on balanced scorecard performance evaluation judgments. *Journal of Management Accounting Research*, *26*(1), 165–184. https://doi.org/10.2308/jmar-50639

Kaplan, R. S., & Norton, D. P. (1996). strategic learning & the balanced scorecard. In *Strategy & Leadership* (Vol. 24, Issue 5, pp. 18–24). https://doi.org/10.1108/eb054566

Kaplan, S. E., Petersen, M. J., & Samuels, J. A. (2012). An examination of the effect of positive and negative performance on the relative weighting of strategically and non-strategically linked balanced scorecard measures. *Behavioral Research in Accounting*, *24*(2), 133–151. https://doi.org/10.2308/bria-50114

Kaplan, S. E., Petersen, M. J., & Samuels, J. A. (2018). Further evidence on the negativity bias in performance evaluation: When does the evaluator's perspective matter? *Journal of Management Accounting Research*, *30*(1), 169–184. https://doi.org/10.2308/jmar-51698

Kramer, S., & Maas, V. S. (2020). Selective attention as a determinant of escalation bias in subjective performance evaluation judgments. *Behavioral Research in Accounting*, *32*(1), 87–100. https://doi.org/10.2308/bria-18-021

Liedtka, S. L., Church, B. K., & Ray, M. R. (2008a). Performance Variability, Ambiguity Intolerance, and Balanced Scorecard-Based Performance Assessments. In *BEHAVIORAL RESEARCH IN ACCOUNTING* (Vol. 20, Issue 2).

Liedtka, S. L., Church, B. K., & Ray, M. R. (2008b). Performance Variability, Ambiguity Intolerance, and Balanced Scorecard-Based Performance Assessments. In *BEHAVIORAL RESEARCH IN ACCOUNTING* (Vol. 20, Issue 2).

Lillis, A. M., Malina, M. A., & Mundy, J. (2022). The Role of Subjectivity in Mitigating Incentive Contracting Risks. *Accounting Review*, *97*(1), 365–388. https://doi.org/10.2308/TAR-2017-0652

Lipe, M. G., & Salterio, S. E. (2000). The Balanced Scorecard: Judgmental Effects of Common and Unique Performance Measures. In *THE ACCOUNTING REVIEW* (Vol. 75, Issue 3).

Long, J. H., Mertins, L., & Vansant, B. (2015). The effects of firm-provided measure weightings on evaluators' incorporation of non-contractible information. *Journal of Management Accounting Research*, *27*(1), 47–62. https://doi.org/10.2308/jmar-50837

Maas, V. S., & Verdoorn, N. (2017). The effects of performance report layout on managers' subjective evaluation judgments. *Accounting and Business Research*, *47*(7), 731–751. https://doi.org/10.1080/00014788.2017.1324756

Maas, V.S., Rinsum, M.V., and Towry, K.L. (2012). In Search of Informed Discretion: An Experimental Investigation of Fairness and Trust Reciprocity. The Accounting Review, Vol. 87, No. 2. DOI: 10.2308/accr-10205 2012 pp. 617–644

Maske, M.K., Sohn, M., and Hirsch, B. (2021). How managerial accountability mitigates a halo effect in managers' ex-post bonus adjustments. Management Accounting Research, 51. https://doi.org/10.1016/j.mar.2021.100738

Maske, M.K., Sohn, M., (2023). Do Birds of a Feather Flock Together? The Joint Effects of Manager and Subordinate Narcissism on Performance Evaluation. *European Accounting Review.* DOI: 10.1080/09638180.2023.2235379

Mitsuhashi, H. and Ohta, Y. (2024). The bonding cost in subjective performance evaluation. *Economics Letters,* 234 https://doi.org/10.1016/j.econlet.2023.111465

Murphy, K. R. (2020). Performance evaluation will not die, but it should. *Human Resource Management Journal*, *30*(1), 13–31. https://doi.org/10.1111/1748-8583.12259

Mursita, L.Y., and Nahartyo, E. (2021). How centrality bias in subjective evaluation affects positive and negative employee work behavior: a real-effort task experiment. *Journal of Accounting & Organizational Change*, Vol. 18 No. 5, 2022 pp. 789-810

Nixon, B. (1998). Research and development performance measurement: a case study. *Management Accounting Research*, 9, 329-355

Prendergast, C. (1999). The Provision of Incentives in Firms. *Journal of Economic Literature*, Vol. XXXVII.

Taylor, C. R., & Yildirim, H. (2011). Subjective performance and the value of blind evaluation. In *Review of Economic Studies* (Vol. 78, Issue 2, pp. 762–794). Oxford University Press. https://doi.org/10.1093/restud/rdq005

Tran, T., and Jarvinen, J. (2022). Understanding the concept of subjectivity in performance evaluation and its effects on perceived procedural justice across contexts. *Accounting & Finance,* 62, 4079–4108

Uribe, J., Carnahan, S., Meluso, J. and Austin-Breneman, J. (2022) How do managers evaluate individual contributions to team production? A theory and empirical test. *Strategic Management Journal,* 43:2577–2601. DOI: 10.1002/smj.3433