

INTEGRATIVE CONCEPTUAL REVIEW

Evaluating the Effectiveness of Performance Management:
A 30-Year Integrative Conceptual ReviewDeidra J. Schleicher
Texas A&M UniversityHeidi M. Baumann
Bradley UniversityDavid W. Sullivan and Junhyok Yim
Texas A&M University

This integrative conceptual review is based on a critical need in the area of performance management (PM), where there remain important unanswered questions about the effectiveness of PM that affect both research and practice. In response, we create a theoretically grounded, comprehensive, and integrative model for understanding and measuring PM effectiveness, comprising multiple categories of evaluative criteria and the underlying mechanisms that link them. We then review more than 30 years (1984–2018) of empirical PM research vis-à-vis this model, leading to conclusions about what the literature has studied and what we do and do not know about PM effectiveness as a result. The final section of this article further elucidates the key “value chains” or mediational paths that explain how and why PM can add value to organizations, framed around three pressing questions with both theoretical and practical importance (*How do individual-level outcomes of PM emerge to become unit-level outcomes? How essential are positive reactions to the overall effectiveness of PM? and What is the value of a performance rating?*). This discussion culminates in specific propositions for future research and implications for practice.

Keywords: performance management, performance appraisal, evaluation, integrative conceptual review

Despite the popularity of performance appraisal (PA) and performance management (PM) in both research and practice, there is a great deal yet to know about the *effectiveness* of these practices. Consider, for example, the following observations.

These systems constitute a ‘human resource management paradox and their effectiveness an elusive goal’ (Taylor, Tracy, Renard, Harrison, & Carroll, 1995). (Nurse, 2005, p. 1178)

The formula for effective [PM] remains elusive. (Pulakos & O’Leary, 2011, p. 146)

There is no shortage of recommendations in the practitioner literature about what makes for effective PM systems. . . . The problem is that

few studies support the many claims about the actual contributions of various practices to the overall effectiveness of PM systems. (Haines & St-Onge, 2012, p. 1171)

It is not clear that [PM] will lead to more effective organizations. . . . Identifying how (if at all) the quality and the nature of performance appraisal programs contribute to the health and success of organizations is a critical priority. (DeNisi & Murphy, 2017, p. 429)

The lack of clear and compelling evidence for the effectiveness of PM (defined as “a continuous process of identifying, measuring, and developing the performance of individuals and teams and aligning performance with the strategic goals of the organization,” Aguinis, 2013, p. 2) has given rise to recent debates about whether or not formal PM is even necessary (e.g., Adler et al., 2016; Pulakos & O’Leary, 2011). Addressing these sorts of issues, as well as making informed judgments about PM research and practice in general, requires a fuller articulation of the evaluative space of PM than available in the extant literature. This is the primary purpose of this article, which identifies a particularly pressing need based on our extensive review of the PM literature: a theoretically grounded, comprehensive, and integrative framework for PM effectiveness.¹

This article was published Online First January 24, 2019.

Deidra J. Schleicher, Department of Management, Texas A&M University; Heidi M. Baumann, Department of Management and Leadership, Bradley University; David W. Sullivan and Junhyok Yim, Department of Management, Texas A&M University.

We wish to express our sincere appreciation to Murray Barrick, Wendy Boswell, and Matt Call for their very helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Deidra J. Schleicher, who is now at Ivy College of Business, Iowa State University, 2167 Union Drive, Ames, IA 50011-2027. E-mail: deidra@iastate.edu

¹ We thank, and agree with, a reviewer who pointed out that this issue within PM is actually a more specific instance of an issue that has been around a long time: the “criterion problem” (see Austin & Villanova, 1992).

The need for such a framework is highlighted by recent discussions within practice. For example, [Pulakos and O'Leary \(2011, p. 154\)](#) ask whether PM systems "provide a sufficient return to justify their use." Related, there has been a push to simplify PM by streamlining its "low value" aspects (see [Effron & Ort, 2010](#); and [Buckingham & Goodall's, 2015](#) discussion of Deloitte's changes in this regard). More generally, [Lawler and McDermott \(2003\)](#) find "little research data to establish the impact of the many practices recommended in the writings on PM" (p. 50). One key challenge is that there are myriad ways to define what terms like "return," "value," and "impact" mean in this context. Indeed, different research streams historically have argued (implicitly or explicitly) for different evaluative foci. For example, an ability-based or cognitive perspective on PA privileges the rating task and argues for an emphasis on psychometric criteria (e.g., [Cardy & Dobbins, 1994](#)); a motivational view privileges PM as a vehicle for improving employee performance and argues that "the proper focus . . . is to change employee behavior on the job" ([DeNisi & Pritchard, 2006](#)); and strategic views privilege unit-level outcomes and argue for firm performance as the ultimate criterion ([DeNisi & Smith, 2014](#)).

Importantly, our review of the PM literature reveals no previous attempts to systematically and comprehensively map (let alone integrate) the full evaluative criterion space of PM implied by these disparate research streams. This is likely one of the key contributors to some of the issues noted above. Specifically, our review suggests that cumulative and actionable knowledge about PM effectiveness has been significantly hindered by lack of attention to articulating and studying the multiple types of PM evaluative criteria, how they interrelate (e.g., how do more proximal criteria such as reactions accumulate to create value for the organization?), and how they are differentially relevant for different questions. Both empirical research and conceptual models historically have focused on a disappointingly small number of PM criteria (e.g., rating errors and accuracy, rater reactions; [Cardy & Dobbins, 1994](#); [Levy & Williams, 2004](#); see [Table 1](#), which provides a summary of earlier work). There exist very few models of how multiple types of PM criteria are likely to interrelate, and no such models that are comprehensive. In response, as part of this integrative conceptual review, we created a comprehensive theoretical model for the criteria underlying PM effectiveness. This model combines empirical and theoretical work in multiple areas to identify the types of criteria that have been—or should be—used to evaluate the effectiveness of PM.

The creation of this comprehensive model and subsequent review of the literature vis-à-vis this model are our primary contributions, representing a significant step forward compared to prior work in several ways. We integrate PM effectiveness criteria relevant to both research and practice, a longstanding need in this area ([Bretz, Milkovich, & Read, 1992](#); [Ilgen, Barnes-Farrell, & McKellin, 1993](#)). Moreover, although we incorporate extant models, we go beyond these to add concepts from other literatures critical for understanding the mechanisms underlying PM effectiveness. Specifically, PM literature to date has either (a) had a very micro focus, not attempting to link individual criteria like rating quality or reactions to unit-level constructs (see earlier review by [Levy & Williams, 2004](#)); or (b) has adopted an exclusively macro focus (e.g., [DeNisi & Smith's, 2014](#) discussion of PM and firm performance). In contrast we argue that progress in

understanding PM effectiveness requires incorporation of *both* micro and macro constructs as well as specification of the processes that link them ([Ployhart & Moliterno, 2011](#)). Doing so allows us to articulate how the various criteria are interrelated, including a mapping of the key mediational paths (or what we term "value chains") underlying PM effectiveness.

This model (see [Figure 1](#)) in turn has several important implications for both research and practice. First, regarding implications for PA/PM researchers specifically, our review uses this model to distill cumulative knowledge from the empirical PM literature, in terms of what aspects of PM exert the biggest influence on which evaluative criteria. This allows us to synthesize what is currently known about the effectiveness of PM while simultaneously identifying a number of limitations in the extant literature, which in turn provides an important foundation for charting a specific and fruitful course for future research. Second, regarding implications for practice, the distilled knowledge from our review concisely identifies which aspects of PM make the biggest difference for specific evaluative criteria. This enables organizations interested in a particular outcome (e.g., improving employees' reactions to PM) to understand what levers are likely to be most impactful in that goal. Our model and review of relationships *among* criteria also help organizations identify the more proximal criteria that lead to more distal outcomes. It is often the latter (e.g., firm performance) in which organizations are most interested, but identifying a direct link between these and PM can be very difficult, given the many alternative explanations.

Third, regarding implications for literatures beyond PA/PM, we contribute to the strategic human resources (HR) literature, which has emphasized the importance of better understanding the "black box" linking HR practices to organizational performance ([Becker & Huselid, 2006](#); [Messersmith, Patel, Lepak, & Gould-Williams, 2011](#), or what macro researchers would label the "microfoundations" of organizational performance, [Coff & Kryscynski, 2011](#)). Our comprehensive model that incorporates both micro and macro evaluative criteria and specifies their interrelationships helps shed light here. Finally, in articulating how PM affects both proximal and more distal criteria and emerges from individual to unit-level phenomena, we contribute to important multilevel work in the area of human capital ([Ployhart & Moliterno, 2011](#); [Ployhart, Nyberg, Reilly, & Maltarich, 2014](#)). [Ployhart and Moliterno \(2011\)](#) note that "one of the most promising avenues for future research will be linking specific HR practices to human capital emergence" (p. 145), and our model depicts multiple ways in which PM specifically can affect such emergence.

In the sections that follow, we first explain the scope of this review, followed by a description of how our model of PM evaluative criteria was created, how we used it as a framework for systematically reviewing and coding more than 30 years of empirical PM work, and the meaning of each component. Then we synthesize the empirical PM research via this model (including criteria interrelationships), drawing conclusions about what the literature has studied and what we do and do not know about PM effectiveness as a result. The final section of our article further elucidates the key value chains or mediational paths that explain how and why PM processes can add value to organizations. Discussion of these specific mediational paths is organized around several pressing questions with both theoretical and practical im-

Table 1
Extant Models of PA/PM Effectiveness

Authors and year	Journal	Scope	Approach	Years reviewed	Criterion constructs examined	Relevant model components
Cardy and Dobbins (1994)	Book	PA	Theoretical model	Unspecified	<ul style="list-style-type: none">• Rater errors• Rating accuracy• Qualitative aspects• EE reactions	<ul style="list-style-type: none">• MGR learning (Rating quality)• EE and MGR reactions (Satisfaction and cognitive)• EE reactions (All)• EE learning (Motivational)• EE reactions (All)
Cawley, Keeping, and Levy (1998)	<i>Journal of Applied Psychology</i>	PA	Meta-analysis	1967–1998		
Keeping and Levy (2000)	<i>Journal of Applied Psychology</i>	PA	Literature review and single empirical study	All extant appraisal research	<ul style="list-style-type: none">• EE reactions	
den Hartog, Boselie, and Paauwe (2004)	<i>Applied Psychology: An International Review</i>	PM	Theoretical model	Unspecified	<ul style="list-style-type: none">• EE perceptions and attitudes• EE behavior and performance• Organizational performance	<ul style="list-style-type: none">• EE reactions (All)• EE transfer (Task performance)• Unit-level financial performance
Levy and Williams (2004)	<i>Journal of Management</i>	PA	Literature review	1995–2003	<ul style="list-style-type: none">• Rater errors and biases• Rating accuracy• Appraisal reactions• EE performance improvement• Agreement between evaluators• Fairness in evaluation procedures• Consistency across rates/time• Distributive justice	<ul style="list-style-type: none">• MGR learning (Rating quality)• EE and MGR reactions (All)• EE learning (Motivational)
DeNisi and Pritchard (2006)	<i>Management and Organization Review</i>	PA and PM	Theoretical model	Unspecified	<ul style="list-style-type: none">• Ratee reactions	<ul style="list-style-type: none">• EE transfer (Task performance)• MGR learning (Rating quality)• EE reactions (Fairness)
Pichler (2012)	<i>Human Resource Management</i>	PA	Meta-analysis	All extant appraisal research		<ul style="list-style-type: none">• EE reactions (All)
Kinicki, Jacobson, Peterson, and Prussia (2013)	<i>Personnel Psychology</i>	PM	Literature review and scale development	Unspecified	<ul style="list-style-type: none">• MGR performance• management behavior	<ul style="list-style-type: none">• MGR learning (Skills-based)• MGR transfer (Quality of decisions made about EEs)• Unit-level financial performance
DeNisi and Smith (2014)	<i>Academy of Management Annals</i>	PA and PM	Literature review and theoretical model	Unspecified	<ul style="list-style-type: none">• Firm-level performance	<ul style="list-style-type: none">• MGR learning (Rating quality)• EE and MGR reactions (All)
DeNisi and Murphy (2017)	<i>Journal of Applied Psychology</i>	PA and PM	Literature review	1917–2015 (Focus on 1970–2000 and on JAP articles)	<ul style="list-style-type: none">• Rater errors and biases• Rating accuracy• Ratee and rater reactions• EE performance improvement• Firm-level performance	<ul style="list-style-type: none">• EE transfer (Task performance)• Unit-level financial performance

Note. Italicized text represents constructs that were proposed as independent variables in the original source, but are categorized in the current model as criterion constructs. EE = employee; MGR = manager; PM = performance management; PA = performance appraisal.

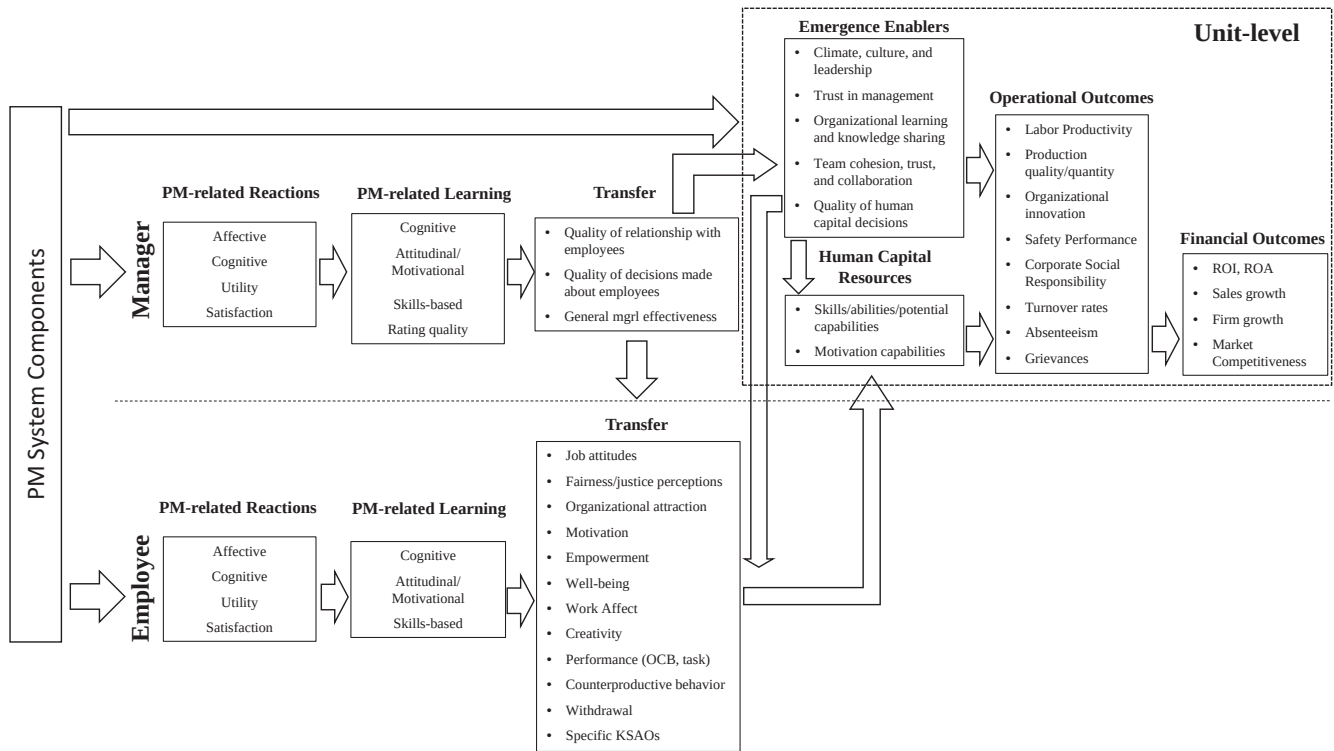


Figure 1. Model of evaluative criteria underlying performance management (PM) effectiveness.

portance, culminating in specific propositions for future research and implications for practice.

The Scope of This PM Review

There are several aspects related to scope that we would like to clarify. To start, our review focuses on PM. Whereas PA is generally understood to be a discrete, formal, organizationally sanctioned event, usually occurring just once or twice a year, PM is seen as a broader set of ongoing activities aimed at managing employee performance (DeNisi & Murphy, 2017; DeNisi & Pritchard, 2006; Williams, 1997). In other words, PA can be thought of as a subset of PM (see also Levy, Tseng, Rosen, & Lueke, 2017). We use the terms PA and PM somewhat interchangeably when referring to the body of literature only. The scope of our review (which is PM) necessarily includes work in both PA and PM, and to create a comprehensive evaluative model, it is necessary to include both the traditionally narrower practices of PA (constituting a longer and more voluminous tradition in the empirical literature) as well as the broader set of activities considered more recently to be part of PM. Thus, we discuss both in the ensuing review of the literature, which spans the last 30+ years of work in PA/PM (1984–2018).²

Our review is also not a “general” review of PM but instead is more specifically focused on the evaluative criteria of PM. This addresses what we see as a particularly important need in the literature (as articulated above); it also makes this review substantively unique from others in the literature (see Table 1), including the very recent literature. For example, DeNisi and Murphy (2017), in the Centennial Issue of *Journal of Applied Psychology*

(JAP), summarize PA/PM research published in JAP specifically, during the “heyday” of PA research (1970–2000), in eight areas: rating scale formats, criteria for evaluating ratings (primarily rating quality and rater and ratee reactions, see Table 1), PA training, reactions to appraisal, purpose of rating, rating sources, demographic differences in ratings, and cognitive processes in PA. Another review on the topic of PM was recently published in the *Journal of Management* (Schleicher et al., 2018). Whereas the current review can be thought of as comprehensively articulating what is known about the outcomes or dependent variables (“DVs”) of PA and PM, Schleicher et al. (2018) focus squarely on the independent variables (“IVs”) of PM, categorizing all of the components of PM systems to help shed light on what the most relevant “moving pieces” are of PM practices and systems. Importantly, neither of these two recent reviews, nor any that came before them, have explicitly and comprehensively focused on the evaluative criteria of PM, as the current review does.

Finally, it is admittedly difficult to discuss the “DVs” of PM without also referencing the “IVs,” as it is useful to summarize which aspects of PM are particularly influential in affecting the various evaluative criteria. Schleicher et al. (2018) take a systems-based approach to understanding the various IVs of PM. Because their taxonomy is the most recent and most comprehensive ap-

² This timeframe seemed appropriate given that DeNisi and Murphy (2017) identified the year 2000 as the end of the “heyday” of PA research. Our timeframe of 1984–2018 brings us to the most recent research and also allows for a nearly even split (17–18 years on either side) regarding the ending of this heyday.

proach to date of the IVs of PM—and also because we built our DV model with the assumption that PM in organizations is a system—we adopt their IV framework for facilitating our synthesis of the empirical research, as we discuss in that later section.

Creation and Overview of Our Model of PM Evaluative Criteria

In creating our model, we took an iterative (inductive-deductive-inductive) approach. First, we reviewed the last 30+ years of work in PA/PM, including empirical and conceptual articles in both the research and practice, and micro and macro literatures, to uncover the types of evaluative criteria being measured and discussed. By “criteria,” we mean the categories of constructs used to measure the effectiveness of PM (see Kirkpatrick, 1987). We wanted our model to be explicitly comprehensive with regard to (a) the content existing in the variety of (narrower) evaluative frameworks in the extant literature; (b) criteria of interest to both research and practice; and (c) both micro and macro constructs. Regarding (a), we incorporated definitions of PA effectiveness by Cardy and Dobbins (1994), Keeping and Levy (2000), and Levy and Williams (2004) and frameworks from other authors (e.g., den Hartog, Boselie, & Paauwe, 2004; DeNisi & Smith, 2014; Toegel & Conger, 2003). Table 1 provides a summary of this prior (and notably narrower) work. Regarding (b), we know from longstanding discussions of the “research-practice gap” in PA that researchers and practitioners tend to be interested in different criteria (Banks & Murphy, 1985; Bretz et al., 1992). For example, while issues of validity and other psychometrics are focal evaluative criteria in research, issues of acceptability to users are key in practice. Wanting to reflect both sides of this “gap,” we explicitly incorporated criteria important to research *and* practice. Regarding (c), a comprehensive and generative model also must incorporate both “micro” and “macro” criteria, as full understanding can only come by examining both what PM can do to and for individuals as well as what it can do to and for organizations. Although extant writing in PM (and certainly PA) has had a decidedly more micro feel (notable exceptions include Bhawe & Brutus, 2011; DeNisi & Smith, 2014), the evaluation of PM is inherently multilevel. In fact, we would argue that this is likely more true for PM than for other areas of HR, given the integral role of the manager in PM (den Hartog et al., 2004). PM processes and policies affect organization-level outcomes not only through employees (“ratees” in traditional PA research) but also through the actions and attitudes of managers (“raters” in traditional PA research). For this reason, our model maps the evaluative criteria at both employee/ratee and manager/rater levels as well as how these individual-level constructs aggregate and emerge to affect unit-level outcomes (see Figure 1).³

Second, we identified models and theories from other literatures that would be useful for classifying all the criteria uncovered in the previous step, suggesting additional relevant criteria, and perhaps most important, understanding how all of these criteria might interrelate in theoretically meaningful ways. Thus, our model includes both criteria measured in the extant PM literature as well as those that are not currently measured but are theoretically relevant. The latter may denote mechanisms that explain how some criteria link to other more distal criteria. We believe these are important to identify, given the goals of a more comprehensive

model, which include understanding *how* PM results in effectiveness. For this deductive phase we relied in particular on work in the training evaluation area, including Kirkpatrick’s (1987) taxonomy, Alliger, Tannenbaum, Bennett, Traver, and Shotland’s (1997) model of the relations among training criteria, and the Kraiger, Ford, and Salas (1993) model of cognitive, skill-based, and affective learning criteria; and theories within strategic HR, including the ability-motivation-opportunity (AMO) framework (Becker & Huselid, 1998; Delery & Shaw, 2001; Jiang, Takeuchi, & Lepak, 2013) and multilevel work on the construct of human capital resources and the emergence process (Ployhart & Moliterno, 2011; Ployhart et al., 2014).

Third, we then systematically coded all criterion variables found in the empirical PM literature, identified through a search that used Business Source Ultimate and PsycINFO for the years 1984–2018 and the terms *performance management*, *performance appraisal*, and *performance evaluation*. After removing all irrelevant articles, there were a total of 488 empirical PM articles (544 separate studies, with 768 instances of criteria across all studies). We coded each study vis-à-vis the components of our model and also recorded findings and methodological details. This final step ensured completeness of the model and also gave us important summative information about what the literature is and is not investigating with regard to evaluative criteria and what we know about PM as a result. The resulting model is depicted in Figure 1, with each component explained below. Here we discuss linkages between components at a general level, to establish the relevance of various components; in the final section of the article we articulate these links in greater detail and explicate specific propositions.

PM-Related Reactions

Because PM practices first affect employees’ perceptions (den Hartog et al., 2004), reactions are the first component of our model (see Figure 1). This refers to how employees and managers feel or think about the overall PM system and/or its specific aspects (e.g., rating, the appraisal interview, a feedback meeting); for employees, this would include managers as a target of reactions, given they are enactors of these processes. Theoretically, reactions play an important role as they can relate to learning (Alliger, Tannenbaum, Bennett, Traver, & Shotland, 1997; Kirkpatrick, 1987), and they have been found to be important in the social exchange between PM partners (i.e., managers and employees; Masterson, Lewis, Goldman, & Taylor, 2000; Pichler, 2012), suggesting they may be related to attitudes and behaviors as well.

Although the majority of PM research has focused on reactions of employees (especially ratees), reactions of managers are also key to understanding PM. Because such practices “are facilitated and implemented by direct supervisors or front-line managers” (den Hartog et al., 2004, p. 565), their reactions are critical in any model of PM effectiveness. In addition, there is evidence that raters’ attitudes and beliefs about PM are related to their rating behavior and that these PM-specific reactions are stronger predictors of such behavior than are general job or organizational attitudes (Tziner, Murphy, Cleveland, & Roberts-Thompson, 2001). Although the structure of this category (see next paragraph) par-

³ From here on out we use the more general terms of *employees* and *managers*, respectively.

allels that of employee reactions, manager reactions likely have different implications for downstream criteria (and operate through different mediators) than employee reactions (Seiden & Sowa, 2011), as we develop later.

Like Alliger et al.'s (1997) augmentation of Kirkpatrick's taxonomy, our model distinguishes between affective, cognitive, and utility reactions to PM; we also add satisfaction as a subcategory to capture overall evaluations of PM (Keeping & Levy, 2000). Affective reactions refer to how the employee or manager *feels* about the PM event or system and include discomfort, frustration, anxiety/stress, or other emotional reactions to PM (e.g., David, 2013; Smith, Harrington, & Houghton, 2000). Cognitive reactions refer to how the employee or manager *thinks* about the PM event or system and include perceived justice or fairness, perceived acceptability or appropriateness, and perceived accuracy of the evaluation (e.g., Erdogan, 2002; Erdogan, Kraimer, & Liden, 2001; Hedge & Teachout, 2000). Utility reactions more directly ask about the perceived *usefulness* or *value* of the PM event or system (e.g., Burke, 1996; Keaveny, Inderrieden, & Allen, 1987; Nathan, Mohrman, & Milliman, 1991). Satisfaction reactions are typically measured as a general evaluation of the PM system or event (Cawley, Keeping, & Levy, 1998). Although satisfaction can be affective or cognitive (see Schleicher, Smith, Casper, Watt, & Greguras, 2015; Schleicher, Watt, & Greguras, 2004), many reactions in the PM literature measure more general satisfaction and cannot be cleanly categorized as just affective/cognitive. Thus, we retained overall satisfaction as a subcategory. Keeping and Levy (2000) found that PA reactions (e.g., satisfaction, utility) are best modeled as distinct constructs that are related to one another through a higher-order factor. Moreover, we know from the training evaluation literature that affective versus cognitive versus utility-based reactions can have differential effects on other criteria (Alliger et al., 1997). Thus, we believe it is important to differentiate reactions in this way in our model. Finally, we found in our review that what the PM literature sometimes casually refers to as reactions (e.g., "buy-in," acceptance, or commitment to the PM system) may be more accurately classified as learning, as described in the next section.

PM-Related Learning

We argue that multifaceted learning, by both employees and managers, is an expected outcome of PM, yet one that has never been fully articulated in extant models (see Table 1). The training literature describes learning as "the extent to which trainees have acquired relevant principles, facts, or skills" (Kraiger, Ford, & Salas, 1993, p. 311), and the learning components of our model reflect what employees and managers may have gained—in terms of proximal PM-related knowledge, skills, attitudes, and motivation—as a result of PM. This necessarily includes both learning things about PM itself (e.g., for employees, awareness of development opportunities; for managers, awareness of what behaviors comprise effective feedback meetings or effective note-taking) as well as learning things about oneself (e.g., increased self-awareness regarding strengths and areas for improvement). By "proximal," we mean that the learning occurred as a direct result of participating in a PM task (e.g., the employee's increased awareness of and greater intent to engage in development opportunities after participating in a formal performance evaluation; Boswell & Boudreau, 2002) or is

in reference to the PM aspects themselves (e.g., managers' increased understanding of what goes into effective feedback and beliefs about its importance); they are also often measured in close proximity to the PM event.

To build out this component, we rely on Kraiger et al.'s (1993) multidimensional model of learning criteria and differentiate between cognitive, skills-based, and attitudinal/motivational learning (see Figure 1). Cognitive PM-related learning includes knowledge (declarative, procedural, and tacit), knowledge organization, or cognitive strategies resulting from participation in PM. Skills-based learning represents behavioral changes related to skill compilation and skill automaticity resulting from PM (e.g., effective note-taking, Mero, Guidice, & Brownlee, 2007; employee feedback-seeking, Moss, Valenzi, & Taggart, 2003). Attitudinal/motivational PM-related learning includes attitudinal changes and motivational tendencies resulting from PM. These are attitudes about PM specifically, formed by participation in the PM system, not job attitudes more generally; and motivation for PM tasks (e.g., acceptance and commitment of goals set during PM; buy-in or acceptance of the PM system as a whole), not general motivation related to one's job. As Kraiger et al. (1993) have noted "an emphasis on behavioral or cognitive measurement at the expense of attitudinal or motivational measurement provides an incomplete profile of learning" (p. 318). In addition, its inclusion in both their model and in ours reflects the fact that training programs and PM systems in organizations go beyond impacting knowledge and skills to also act as "powerful socialization agent[s]" (p. 319), indoctrinating employees and managers to the importance of various aspects of the training content or PM systems. For example, in the PM literature, attitudinal/motivational learning variables include agreement with the theories of performance espoused by the organization (which increases as a result of rater training, Schleicher & Day, 1998) and rater self-efficacy (Tziner et al., 2001) for managers; and intentions to engage in future development (Boswell & Boudreau, 2002) and acceptance of and commitment to goals discussed in the feedback meeting (Tziner & Kopelman, 1988) for employees.

Learning criteria involve PM-related knowledge, skills, attitudes, and motivations that employees and especially managers need to "do PM well" and that should theoretically improve as a result of experience with PM (e.g., understanding what good performance is, learning to more constructively receive feedback, felt accountability for PM, avoidance of intentional distortion). This is an important component of the model because the extent to which managers do PM well is likely to directly affect employees' reactions to PM (Jawahar, 2010; Waung & Jones, 2005), setting off the evaluative chain in the bottom row of our model. It has been suggested that managers who do such things well should also produce employees who are more engaged and motivated (Ladyshewsky, 2010). Unfortunately, these manager learning criteria have been largely ignored in the extant PM literature, with one major exception. Related to this exception, we categorize the quality of ratings under this category because, like the other constructs included here, rating quality represents tangible and proximal manifestations of managers' knowledge, skills, abilities, and motivations gained from the PM process. This psychometric subcategory of learning includes the extent to which ratings are free from errors and biases, are reliable and valid, and are accurate (Aguinis, 2013; Cardy & Dobbins, 1994).

It is important to differentiate learning from reactions in understanding PM effectiveness. Reactions capture the PM event or system as experienced by the employee or manager but are not direct measures of what one may have learned as a result of the PM experience (Kraiger et al., 1993). It is notable, and surprising to us, that prior discussions of PM effectiveness have not explicitly focused on these learning criteria (for employees or managers). Such criteria seem especially important given recent trends focused on more developmental approaches to PM (e.g., “feed-forward” interviews, Kluger & Nir, 2010; strengths-based evaluation, Bouskila-Yam & Kluger, 2011). Cappelli and Tavis (2016), for example, describe the recent PM revolution as a shift “from accountability to learning” (p. 2), and Buckingham and Goodall (2015) describe the focus of Deloitte’s new system as “constant learning” (p. 42). Without effectiveness measures focused on proximal PM-related learning, it may be unclear whether (and how) these new development-focused systems have achieved their goals. Thus, we include PM-related learning as an important evaluative criterion, positioned between reactions and transfer in our model.

Employee Transfer

The employee transfer component of our model includes employee attitudes, behaviors, and outcomes that may be affected by elements of PM but which extend beyond the PM context, in referent (i.e., they refer to the job or organization more broadly) and/or timing of measurement. This component would not include employees’ attitudes about PM specifically or behaviors that are confined to the PM context primarily (these would be classified as employee reactions or learning). Instead this component includes criteria that suggest that the effects of PM may “transfer” back to the job. In Kirkpatrick’s (1976, 1987) model, transfer was largely equated with behavior and performance and defined as “using learned principles and techniques on the job” (Alliger & Janak, 1989, p. 331). Because we are not talking about the effectiveness of just training but rather the outcomes of multifaceted PM systems, we use transfer in a broader sense, to include performance and other behaviors (e.g., withdrawal) but also attitudinal and motivational constructs (e.g., job attitudes, justice). Yet similar to Kirkpatrick’s initial meaning, this component represents the question of whether the effects of PM transfer beyond the immediate PM context (e.g., formal review meeting) back to the “job” to impact employee behaviors and attitudes more broadly. Unlike subsequent components, which are at the unit-level, Transfer criteria reside at the individual level (conceptually and empirically).⁴

There is a heavy focus on “transfer” criteria in the training literature (see, e.g., Baldwin & Ford, 1988; Ford & Weissbein, 1997), and the constructs in this category here are undoubtedly among the most frequently studied and important outcomes in organizational behavior and I/O psychology in general. Yet historically they have been less studied as explicit outcomes of PM. For example, in extant conceptual models (see Table 1), only task performance is referred to and in only a few examples (den Hartog et al., 2004; DeNisi & Murphy, 2017; DeNisi & Pritchard, 2006). In the empirical PM literature, however, examination of these criteria has more than doubled in recent, compared with older, research (i.e., there were 47 instances before 2000, compared with 121 post-2000). This is welcome empirical progress, as these

criteria play an important role theoretically in the various value chains of PM, as we develop later.

Manager Transfer

Like employee transfer, the manager transfer component includes criteria that extend beyond the PM context to the manager’s role in the organization more generally. Given the longstanding emphasis on interpersonal and decision-making activities in managerial work (Mintzberg, 1971), this component includes both relational and decision-making constructs. PM has been discussed as a critical tool that serves as a basis for making effective decisions about human resources (Cardy & Dobbins, 1994), making managers’ effectiveness in this regard an important evaluative criterion. The manager–employee relationship is also clearly relevant and has been noted as essential for increasing PM effectiveness (Pulakos & O’Leary, 2011). We agree wholeheartedly but argue here that these relationships can themselves be impacted by aspects of PM and thus should be studied as a DV in PM research, not just as an IV. In short, the manager transfer component concerns the extent to which PM changes how managers do their job (or at least employees’ perceptions of this, Kacmar, Wayne, & Wright, 1996), and it includes the quality of relationships formed with employees, the quality of decisions managers make about employees, and other indicators of general managerial effectiveness.

These transfer criteria would likely be affected by the learning managers amass as a result of aspects of PM (relational criteria specifically could also be impacted by employees’ reactions to PM). In turn, these improved aspects of managerial effectiveness impact employees’ attitudes and behaviors (see Figure 1). We also argue that manager transfer criteria exert an important influence on unit-level criteria (discussed in the following sections). Specifically, the quality of managers’ relationships with employees aggregate into several important emergence enablers such as climate and trust in management. And the quality of decisions managers make about employees aggregate into the quality of unit-level human capital decisions, which determines the unit’s ability to “leverage” the human capital available (see Lakshman, 2014).

Unit-Level Human Capital Resources

In our model, employee transfer constructs knowledge, skills, abilities, and other characteristics (KSAsOs), attitudes, and behaviors) aggregate to become unit-level human capital resources (HCRs; Ployhart & Moliterno, 2011; Ployhart et al., 2014), and it

⁴ In our discussion of unit-level criteria further below, we rely on Ployhart et al.’s (2014) recent theorizing about the construct of human capital resources. Our transfer criteria require some clarification vis-à-vis that theorizing. Ployhart et al. (2014) exclude constructs like attitudes, satisfaction, and motivation from their discussion of KSAsOs (the essential building blocks of human capital resources), because they view such characteristics as being situationally specific and induced. Setting aside evidence that such characteristics can in fact be stable (e.g., Staw & Ross, 1985), we argue that these other characteristics of employees (i.e., attitudes, motivation), especially when emergent at unit levels, do have economic relevance for organizations (see e.g., Barrick, Thurgood, Smith, & Courtright’s, 2015, and Harter, Schmidt, & Hayes’, 2002 work on employee engagement). For that reason, we include a comprehensive set of criteria under employee transfer (see Figure 1).

is these HCRs that can influence firm operational and financial performance (see Figure 1).⁵ Borrowing from the AMO framework popular within strategic HR, these unit-level HCRs are organized into the following two categories in our model: skills/abilities/potential, and motivational capabilities. Based in the view that employees' ability (A), motivation (M), and opportunity (O) to perform are key determinants of performance, the AMO model posits that HR systems relate to firm performance through their influence on these three elements (e.g., Becker & Huselid, 1998; Delery & Shaw, 2001; Jiang, Lepak, Hu, & Baer, 2012; Lepak, Liao, Chung, & Harden, 2006).⁶ For example, HR practices (including PM) might affect unit-level abilities or skills such as adaptability, creativity, or potential (our skills/abilities/potential category); and/or motivational capabilities, such as collective engagement (Barrick, Thurgood, Smith, & Courtright, 2015) and unit-level employee commitment and empowerment (Messersmith et al., 2011). These unit-level capabilities (or HCRs), in turn, lead to operational outcomes (see Figure 1).

Yet employee variables do not automatically become unit-level HCRs. As Bliese (2000) notes "the main difference between a lower-level and an aggregate-level variable . . . is that the aggregate variable contains higher-level contextual influences that are not captured by the lower-level construct" (p. 369). In other words, transfer variables and unit-level HCRs are only partially isomorphic, as they have different antecedents (Ployhart & Moliterno, 2011; and supported by our empirical review).⁷ Related, Ployhart, Nyberg, Reilly, and Maltarich (2014) distinguish between human capital and human capital *resources*, defining the latter as unit-level capacities that are *accessible for unit-relevant purposes*. Thus, in our model we depict unit-level HCRs as resulting from employee transfer variables yet moderated by accessibility-related contextual factors. As the next section describes, our emergence enablers category captures these key moderating influences.

Emergence Enablers

Central to the question of how unit-level HCRs are created from individual-level criteria is the process of "emergence" (Ployhart & Moliterno, 2011). Emergent phenomena "originate in the cognition, affect, behaviors, or other characteristics of individuals, [are] amplified by their interactions, and manifest as higher-level, collective phenomenon[a]" (Kozlowski & Klein, 2000, p. 55). Thus, the microfoundations of unit performance are not only employee KSAOs but also the social and psychological mechanisms that constitute this emergence enabling process (Li, Wang, van Jaarsveld, Lee, & Ma, 2018; Ployhart & Moliterno, 2011). Our model captures this important element, depicting emergence enablers as a key moderator between employee transfer and unit-level HCRs (as well as a direct determinant of HCRs and operational outcomes; see Figure 1). Thus, to the extent that PM alters these emergence enablers, it necessarily would result in the emergence of different kinds of HCRs (Ployhart & Moliterno, 2011).

Three categories of emergence enablers were identified by Ployhart and Moliterno (2011): behavioral processes (coordination, communication, and regulatory processes that affect the interdependence of employees, Kozlowski & Ilgen, 2006); cognitive mechanisms (unit climate, memory, and learning, Hinsz, Tindale, & Vollrath, 1997); and affective psychological states (the emotional bonds that tie unit members together, such as cohesion and

trust). Using this conceptual framework, along with the empirical PM literature, we identified the following unit-level outcomes of PM that could be classified as emergence enablers (see Figure 1): climate, culture, and leadership (per Rentsch, 1990, perceptions of unit leadership is part of climate); trust in management; unit learning and knowledge/information sharing; and team cohesion, trust, and collaboration. We add an additional category of emergence enablers, based on the role of managers in PM: the unit-level quality of human capital decisions made. This is an aggregate of the manager transfer criterion, quality of decisions made about employees, and at the unit level we argue that it serves an important enabling function for unit-level HCRs. As Ployhart et al. (2014) have noted, human capital has to be sufficiently *available* to the unit to be considered a resource; and the quality of human capital decisions made determines the extent to which the unit can actually leverage the potential HCRs (see Lakshman, 2014). Our model argues that the quality of decisions made at the unit level, through affecting the availability of human capital, is an important moderator of the link between employee transfer criteria and unit-level HCRs.

Unit-Level Operational and Financial Outcomes

Finally, our model includes organization-level performance and separates this into operational and financial outcomes (see Figure 1). This follows the lead from research in strategic HR, which has argued (although not always found) that operational outcomes are more closely aligned with the improved employee capabilities resulting from HR practices and therefore more strongly related to such practices than are financial outcomes (Combs, Liu, Hall, & Ketchen, 2006; Dyer & Reeves, 1995). Following researchers in strategic HR, we identified the following unit-level *operational* outcomes in the empirical PM literature (see Figure 1): labor productivity, product quality, innovation, safety performance, corporate social responsibility, turnover rates, absenteeism, and grievances.⁸ Per the strategic HR literature, these outcomes result in

⁵ Taking our lead from Ployhart and Moliterno (2011), we use the more generic "unit" terminology; as these authors note, "by defining the level of theory generically at the 'unit level,' [human capital] can exist at the group, department, store, or firm level of analysis, with the relevant aggregation of individual level KSAOs measured at the level that is theoretically and empirically relevant" (p. 144).

⁶ Following Jiang et al. (2012), we exclude opportunity capabilities from our model. As these authors note, ability and motivational capabilities are the two most important mediating paths. In addition, there were no empirical PM articles examining unit-level opportunity capabilities.

⁷ The various ways in which HCRs combine from individual constructs (e.g., composition vs. compilation models) is outside the scope of our model/article. This is discussed in Ployhart et al. (2014), and the interested reader is referred there.

⁸ Some strategic HR research has used a category of organization performance referred to as "HRM outcomes," which includes unit-level constructs such as employee commitment, competence, quality, and turnover (e.g., Beer, Spector, Lawrence, Mills, & Walton, 1984; Guest, 1987, 1997; Zheng et al., 2006). However, to us this seems to be a somewhat unclear mix of HCRs and operational outcomes. Ployhart et al. (2014) note that HCRs are "capacities for action, but they are not the action itself. Therefore, studies that define human capital in terms of employee performance behaviors are not studying HCRs but rather the results or outcomes of such resources" (p. 390). Thus, we classify human capital capacities under resources but human capital outcomes (such as unit-level performance, productivity, turnover, etc.) as operational outcomes.

part from unit-level HCRs (Daley, 1986; Kim, Atwater, Patel, & Smither, 2016; Zheng, Morrison, & O'Neill, 2006). Regarding financial outcomes, there are many ways to operationalize firm financial performance (see Batt, 2002; Goh & Anderson, 2007), but those examined in the PM literature have included return on investment (ROI), return on assets (ROA), sales growth, firm growth, and market competitiveness.⁹

Here we want to clarify the meaning of the horizontal arrangement of our model. That it ends with organizational outcomes does not signify that these are the “ultimate criteria.” Although some have argued that the overall purpose of PM is to improve firm performance (e.g., DeNisi & Smith, 2014; DeNisi & Sonesh, 2011), we argue that what is most relevant depends on the goals of the PM system and the specific effectiveness questions being asked (addressed in the final section of our article). Thus, the positioning of organizational performance at the end of our model should not be taken to imply its overarching importance. Rather, our model is generally organized from left to right in causal-logical sequence, from more micro criteria to more macro criteria, which is the generally established causal direction in training evaluation (Kirkpatrick, 1987) and multilevel research (Ostroff & Bowen, 2000), and allows us to map the emergence process (Ployhart & Moliterno, 2011). It is possible that, over time, there could be reciprocal relationships among components of the model; for example, improved financial performance might lead an organization to invest more into the PM system (see den Hartog et al., 2004). However, this is distinct from the causal sequence linking more proximal evaluative criteria to more distal evaluative criteria (the focus of our model) and is therefore not discussed here.

Synthesis of Empirical PM Research Vis-à-Vis the Model

This section summarizes conclusions from our systematic and comprehensive review of the empirical PM research from 1984–2018 vis-à-vis the components of our evaluative criteria model. Table 2 provides the frequencies of studies in each criterion category, organized by timeframe; Table 3 provides a description of specific variables examined, by criterion category. Rather than reviewing this research in detail criterion by criterion (which Appendix A does, provided for the interested reader), our discussion here is organized along several broader themes we identified in this empirical literature. The first section provides descriptive information on how frequently various criteria are studied in the PM literature and, based on our theoretical model, a discussion of what else we should be examining as a result. The second section summarizes what this empirical research suggests are the aspects of PM that most impact its effectiveness. The third section reviews empirical evidence for the criterion–criterion relationships implicated in our model. Finally, the fourth section identifies methodological trends and limitations in this research and associated recommendations for improvement. Each of these sections contains some suggestions for future research based on the explicit focus of the section. The final major section of the article goes beyond these research suggestions to develop specific research propositions tied to the longer value chains believed to underlie PM effectiveness.

Differential Empirical Emphasis Across PM Criteria and Time

An overall observation from our review is that there has been unequal empirical attention across criteria (and across time). Table 2 lists frequencies for each criterion category, organized by timeframe; several trends are apparent here. First, *employee reactions* (see Appendix A, section Ia) have become the most widely studied outcome in the PM literature (more frequent even than rating quality). Such research exploded following Murphy and Cleveland's (1995, p. 310) claim that reactions were “neglected criteria” in the PM literature and their inclusion in Cardy and Dobbins (1994) model of PA effectiveness, and our review suggests that this strong focus on reactions has continued post-2000. However, *managers' reactions* to PM (see Appendix A, section Ib) have been studied much less often (only 16% of all reactions variables), and this focus has in fact declined post-2000. Research suggests that managers' reactions to PM tend to differ substantially from employees' reactions (Manshor & Kamalanabhan, 2000; Taylor, Pettijohn, & Pettijohn, 1999), perhaps due to differences in knowledge of the PM system (Williams & Levy, 2000); and both play important and distinct roles in our theoretical model. Thus, future research should focus substantially more on manager reactions to PM.

Second, empirical focus on *employee transfer* criteria in PM (see Appendix A, section IV) has significantly increased post-2000 and in fact is essentially tied with employee reactions as the most commonly studied criterion in the more recent literature. Our review suggests transfer includes more than just task performance (indeed, job attitudes were actually studied as often as performance; see Table 2). Given that these constructs create the foundation for unit-level HCRs (Ployhart & Moliterno, 2011; Ployhart et al., 2014), this is a positive trend for understanding PM effectiveness. At the same time, there are criteria we conceptualized as part of employee transfer that have been studied infrequently, including counterproductive behavior (cf., Tziner, Fein, Sharoni, Bar-Hen, & Nord, 2010), employee creativity (cf., Jiang, Wang, & Zhao, 2012), organizational attraction (cf., Blume, Rubin, & Baldwin, 2013; Maas & Torres-González, 2011), and employee well-being (e.g., burnout, stress, self-esteem, safety behaviors; cf., Culig, Dickinson, Lindstrom-Hazel, & Austin, 2008; Gabris & Ihrke, 2001; Johnson & Helgeson, 2002; Milanowski, 2005). More research should be directed to each of these transfer criteria and also specific KSAOs, which are not typically examined as outcomes of PM but which, per our conceptual model, have clear relevance for unit-level HCRs.

Third, our review suggests a different story for learning criteria. Regarding *employee learning* specifically (see Appendix A, section II), there has been much less emphasis on this relative to

⁹ There are a number of moderators believed to affect the strength of the relationship between unit-level HCRs and various measures of organizational performance (some argue, for example, that HCRs must be firm-specific to result in improved organizational performance; Barney & Wright, 1998). In the interest of space and parsimony, because these have been reviewed in detail in other places (see e.g., Mahoney & Kor, 2015) and because we view the primary contribution of our model not in what is mapped out to the right of unit-level HCRs but rather how PM leads up to unit-level HCRs, these moderators are outside the scope of our model and review. Theoretically, they should not be unique to the PM context.

Table 2
Frequency of Criteria Across All PM Studies

Criterion category	Across all studies (<i>n</i> = 768)		1984–2000 (<i>n</i> = 334)		2001–2018 (<i>n</i> = 434)	
	Count	Percent	Count	Percent	Count	Percent
Employee	454	59.11	178	53.29	276	63.59
Reactions	230	29.95	106	31.74	124	28.57
Cognitive	106	13.80	52	15.57	54	12.44
Satisfaction	69	8.98	37	11.08	32	7.37
Utility	39	5.08	13	3.89	26	5.99
Affective	16	2.08	4	1.20	12	2.76
Learning	56	7.29	25	7.49	31	7.14
Cognitive	12	1.56	6	1.80	6	1.38
Skills-based	16	2.08	7	2.10	9	2.07
Attitudinal/motivational	28	3.65	12	3.59	16	3.69
Transfer	168	21.88	47	14.07	121	27.88
Job attitudes	57	7.42	19	5.69	38	8.76
Performance	57	7.42	17	5.09	40	9.22
Withdrawal	20	2.60	4	1.20	16	3.69
Fairness/justice	11	1.43	2	.60	9	2.07
Motivation	13	1.69	5	1.50	8	1.84
CWBs	1	.13	—	—	1	.23
Employee creativity	2	.26	—	—	2	.46
Organizational attraction	2	.26	—	—	2	.46
Employee well-being	5	.65	—	—	5	1.15
Manager	241	31.38	130	38.92	111	25.58
Reactions	45	5.86	24	7.19	21	4.84
Cognitive	17	2.21	10	2.99	7	1.61
Satisfaction	14	1.82	9	2.69	5	1.15
Utility	7	.91	—	—	7	1.61
Affective	7	.91	5	1.50	2	.46
Learning	167	21.74	90	26.95	77	17.74
Cognitive	9	1.17	4	1.20	5	1.15
Skills-based	32	4.17	19	5.69	13	3.00
Attitudinal/motivational	7	.91	3	.90	4	.92
Rating quality	119	15.49	64	19.16	55	12.67
Transfer	29	3.78	16	4.79	13	3.00
Quality of relationships with employees	20	2.60	12	3.59	8	1.84
Quality of decisions made for employees	8	1.04	3	.90	5	1.15
Managerial effectiveness	1	.13	1	.30	—	—
Emergence enablers	52	6.77	21	6.29	31	7.14
Climate and culture	31	4.04	10	2.99	21	4.84
Knowledge sharing	4	.52	2	.60	2	.46
Team cohesion/trust and collaboration	12	1.56	8	2.40	4	.92
Quality of human capital decisions	5	.65	1	.30	4	.92
Affect/mood	—	—	—	—	—	—
Unit-level	21	2.73	5	1.50	16	3.69
Human capital resources	2	.26	—	—	2	.46
Operational outcomes	5	.65	1	.30	4	.92
Financial outcomes	14	1.82	4	1.20	10	2.30

* *n* (and count) refers to the number of instances of each criterion, across studies. These numbers are more than the 544 studies included due to some studies measuring multiple performance management (PM) criteria. Percentages reflect column totals for each of the three time periods.

employee reactions or transfer (although the emphasis on employee learning has at least not declined post-2000). The sparse empirical focus is at odds with the theoretical importance of employee learning for subsequent attitudes, motivation and performance (per our model). Indeed, such learning criteria have been found to completely mediate the relationship between reactions to performance feedback and one's behavioral responses to it (Kinicki, Prussia, Wu, & McKee-Ryan, 2004). Regarding *manager learning* specifically (see Appendix A, section III), although this criterion appears to be frequently studied (see Table 2), that is

almost entirely a function of a continued disproportionate emphasis on rating quality specifically (which has remained post-2000). As a field we know significantly less about other aspects of managers' learning from PM. For example, rater self-efficacy has emerged as an important construct in the literature, and in our model it is categorized as a manager learning criterion. Yet most of the extant research in this area has considered it primarily as an individual difference that predicts other aspects of PM. We suggest the need for more research—such as Tziner and Kopelman (2002) and Wood and Marshall (2008)—that examines the PM system

Table 3
Summary of Empirical PM Research by Component

Model components and subcategories	Variables and sample research
PM reactions	
Manager	
Cognitive	Fairness/justice (Williams & Levy, 2000)
Satisfaction	Appraisal satisfaction (Williams & Levy, 2000)
Utility	Utility of feedback (Erdemli, Sümer, & Bilgiç, 2007)
Affective	Discomfort with PA (Saffie-Robertson & Brutus, 2014)
Employee	
Cognitive	Perceived fairness/justice (Taylor, Tracy, Renard, Harrison, & Carroll, 1995)
	Perceived accuracy (Kinicki, Prussia, Wu, & McKee-Ryan, 2004)
	Acceptance of PM (Hedge & Teachout, 2000)
	Perceived quality of feedback (Anseel, Lievens, & Schollaert, 2009)
Satisfaction	Satisfaction with PM (Nathan, Mohrman, & Milliman, 1991)
Utility	Perceived utility of feedback (Elicker, Levy, & Hall, 2006)
	Utility of PA (Payne, Horner, Boswell, Schroeder, & Stine-Cheyne, 2009)
Affective	Discomfort with PA (Spence & Wood, 2007)
	Negative and positive emotions (David, 2013)
PM learning	
Manager	
Cognitive	Idiosyncratic performance standards (Schleicher & Day, 1998)
	Performance schema accuracy (Gorman & Rentsch, 2009)
	Understanding employee strength/weakness (Selden, Sherrier, & Wooters, 2012)
	Managerial knowledge of PA (Davis & Mount, 1984)
	Memory strength (Martell & Leavitt, 2002)
Skills-based	Understanding one's contribution to unit objectives (Mabey, 2001)
	Effectiveness in completing PA forms (Davis & Mount, 1984)
	Taking better notes (Mero, Motowidlo, & Anna, 2003)
	Behavioral specificity in evaluation comments (Macan et al., 2011)
	Performance information recall ability (DeNisi & Peters, 1996)
Attitudinal/motivational	Effectiveness of supervisor appraisal behavior (Eberhardt & Pooyan, 1988)
	Agreement with org. performance theories (Schleicher & Day, 1998)
	PA self-efficacy (Tziner, Murphy, Cleveland, & Roberts-Thompson, 2001; Wood & Marshall, 2008)
	Self-serving motives (Goerke, Möller, Schulz-Hardt, Napiersky, & Frey, 2004)
Rating quality	Error, biases, and accuracy (Cardy & Dobbins, 1994)
	Reliability and validity criteria (Aguinis, 2013)
Employee	
Cognitive	Awareness of development opportunities (Boswell & Boudreau, 2002)
	Task thoughts (Harackiewicz, Abrahams, & Wageman, 1987)
	Self-awareness (Morgan, Cannan, & Cullinane, 2005)
	Role clarity (Prince & Lawler, 1986)
Skills-based	Perceived benefits of development (Linderbaum & Levy, 2010)
	Way in which employees do their work (Morgan et al., 2005)
	Feedback sharing between peers (Wang, 2007)
	Feedback seeking (Linderbaum & Levy, 2010)
Attitudinal/motivational	Desire to participate in PA (Langan-Fox, Waycott, Morizzi, & McDonald, 1998)
	View of how the PM system aids in performance (Harris, 1988)
	Motivation to improve (Harackiewicz et al., 1987)
	Intended future use of development (Boswell & Boudreau, 2002)
	Goal clarity, acceptance, and commitment (Tziner & Kopelman, 1988)
	Self-efficacy (Bartol, Durham, & Poon, 2001)
	Intentions to change behavior (Johnson & Helgeson, 2002)
Transfer	
Manager	
Quality of relationship with employees	Trust in manager (Korsgaard, Roberson, & Rymph, 1998)
	Supervisor liking/satisfaction (Kacmar, Wayne, & Wright, 1996)
	LMX (Dahling, Chau, & O'Malley, 2012)
	Quality of the coaching relationship (Gregory & Levy, 2012)
	Perceived supervisor support (Armstrong-Stassen & Schlosser, 2010)
	Employee-supervisor working relationship (McBriarty, 1988)
	Confidence in collaborating with manager (Tjosvold & Halco, 1992)
	Cooperation with supervisor (Taylor & Pierce, 1999)
Quality of decisions made about employees	Quality of decisions on job assignment/resource utilization (McBriarty, 1988)
	Accuracy of personnel decisions (Jawahar, 2001)
Managerial effectiveness	Perceptions of supervisor effectiveness (Burke, 1996)
Employee	
Job attitudes	Job satisfaction (Lam, Schaubroeck, & Aryee, 2002; Nathan et al., 1991)
	Organizational commitment (Lam et al., 2002; Pearce & Porter, 1986)

(table continues)

Table 3 (continued)

Model components and subcategories		Variables and sample research
Performance		Perceived organizational support (Masterson, Lewis, Goldman, & Taylor, 2000)
		Job embeddedness (Bambacas & Kulik, 2013)
		Role ambiguity (Youngcourt, Leiva, & Jones, 2007)
		Overall performance (Klein & Snell, 1994)
Withdrawal		Task performance (Nathan et al., 1991; Prince & Lawler, 1986)
		OCB (Findley, Giles, & Mossholder, 2000; Masterson et al., 2000; Norris-Watts & Levy, 2004)
		Intention to turnover (Brown, Hyatt, & Benson, 2010)
		Intention to remain (Taylor et al., 1995)
Fairness/justice		Turnover (Milanowski, 2005)
		Procedural justice (Lam et al., 2002; Masterson et al., 2000)
		Distributive justice (Cheng, 2014; Lam et al., 2002)
		Interactional justice (Linna et al., 2012; Masterson et al., 2000)
Motivation		Intrinsic/extrinsic motivation (Sundgren, Selart, Ingelgård, & Bengtson, 2005)
		Employee engagement (Gruman & Saks, 2011)
		Motivation to work hard (Tjosvold & Halco, 1992)
		Motivation to improve (Taylor et al., 1995)
CWBs		Effort on the job (Taylor & Pierce, 1999)
		Deviant behavior (Tziner, Fein, Sharoni, Bar-Hen, & Nord, 2010)
		Employee creativity (Jiang, Wang, & Zhao, 2012)
		Organizational attractiveness (Blume, Rubin, & Baldwin, 2013)
Employee creativity		Burnout (Gabris & Ihrke, 2001)
		Stress (Milanowski, 2005)
		Self-esteem (Johnson & Helgeson, 2002)
		Safety behaviors (Culig, Dickinson, Lindstrom-Hazel, & Austin, 2008)
Organizational attraction		
Employee well-being		
Emergence enablers		
Climate and culture		Office morale (Burke, 1996)
		Unit-level satisfaction (Daley, 1986; Mullin & Sherman, 1993)
		Support culture (Mamatoglu, 2008)
		Perceived psychological contract fulfillment (Raeder, Knorr, & Hilb, 2012)
Knowledge and information sharing		Ethical climate (Guerci, Radaelli, Siletti, Cirella, & Rami Shani, 2015)
		Creative climate (Sundgren et al., 2005)
		Communication atmosphere of the unit (Mamatoglu, 2008)
		Knowledge sharing of R&D employees (Liu & Liu, 2011)
Team cohesion trust, and collaboration		Knowledge management effectiveness (Tan & Nasuridin, 2011)
		Organizational learning (Wang, Tseng, Yen, & Huang, 2011)
		Team cohesion (McBriarty, 1988; Rowland, 2013)
		Trust for top management (Mayer & Davis, 1999)
Quality of human capital decisions		Effectiveness for influencing performance (Lawler, 2003)
		Effectiveness for differentiating top/poor performer (Lawler, 2003)
Human capital (Unit-level)		
Employee skill/abilities/potential capabilities		Adaptability/flexibility (Mullin & Sherman, 1993)
		Performance potential of workforce (Scullen, Bergey, & Aiman-Smith, 2005)
		Workforce quality (Giumetti, Schroeder, & Switzer, 2015)
		Employee's knowledge about how work and strategy aligns (Ayers, 2013)
Employee motivation		Employee motivation (Roberts, 1995)
		Staff commitment (Rao, 2007)
Capabilities		
Operational outcomes		
Labor productivity		Labor productivity (Roberts, 1995; Kim, Atwater, Patel, & Smither, 2016)
		Attainment of quality (Waite, Newman, & Krzystofiak, 1994)
		Production (Zheng, Morrison, & O'Neill, 2006)
		Production quality (Lee, Lee, & Wu, 2010)
Productive quality or quantity		Administrative/process/product innovation (Tan & Nasuridin, 2011)
		Administrative/technological innovation (Jiang et al., 2012)
		Safety behavior (Laitinen & Ruohomäki, 1996)
		Number and rate of occupational injuries/accidents (Reber & Wallin, 1994)
Organizational innovation		
Safety performance		Perceived CSR (Daley, 1986)
		Turnover rate (Batt, 2002)
		Absenteeism (Roberts, 1995)
		Perceived organizational performance (Daley, 1986; Rodwell & Teo, 2008)
CSR		
Collective turnover		
Absenteeism		
Others		
Financial outcomes		
ROI		ROI (Goh & Anderson, 2007)
		Sales growth (Batt, 2002)
		Market competitiveness (Zheng et al., 2006)

Note. PM = performance management; PA = performance appraisal; OCB = organizational citizenship behavior; LMX = leader-member exchange; ROI = return-on-investment; CSR = corporate social responsibility.

antecedents of rater self-efficacy, to understand how aspects of PM can actually build such PM-related self-efficacy in managers. More generally, the importance of these types of learning outcomes will only increase over time, given the move toward more development-focused PM.

Fourth, the empirical emphasis on *manager transfer* (see Appendix A, section V)—which is not widespread overall—has unfortunately declined somewhat post-2000 (see Table 2). Especially needed is additional work on the quality of managers' employee-related decisions and how this is impacted by aspects of PM. This serves as an important building block for unit-level HCRs, yet we found only a few studies examining this criterion. Especially important will be longitudinal designs that capture the implications of PM changes (e.g., eliminating ratings, Adler et al., 2016) on managers' decisions. The assumption is that improved PM processes increase the quality of HR decisions, but this is largely an untested assumption in the literature.

Fifth, *unit-level criteria* (see Appendix A, sections VI., VII., and VIII) have not been as frequently studied overall as the other criteria, but fortunately this focus has increased post-2000 (see Table 2). This trend is especially notable for operational and financial measures of firm performance, and it may largely be due to strategic HR researchers beginning to focus on PM. Regardless, it is a positive trend for understanding the overall effectiveness of PM. At the same time, future research needs to examine other relevant aspects of organizational performance that have received less attention (see Appendix A, section VIII). For example, grievances are listed in our model of operational outcomes (and likely are significantly impacted by the type of PM system; see Payne & Mendoza, 2017), but we could find no empirical research in this area. Future unit-level PM research also needs to focus on additional emergence enablers (see Appendix A, section VII). Our model specifies the *quality of human capital decisions made* as critical in this regard (as it affects the unit's ability to leverage human capital and thus both should determine the amount of HCRs available as well as moderate the link between individual-level human capital and unit-level human capital), yet we found only one study in this area.

The Most Impactful Aspects of PM

The previous section reviewed the prevalence of the criteria themselves in the empirical literature. This section concerns the question of what aspects of PM (i.e., the IVs) are most impactful for PM effectiveness based on this literature, and our overall observation is that the answer appears to vary across the types of evaluative criteria (see Appendix A for details). As noted in the introduction, to synthesize these findings we rely on Schleicher et al.'s (2018) systems-based taxonomy of the IVs of PM, which identifies the following six main components of PM systems: *tasks* (the activities involved in PM, including setting performance expectations, observing performance, integrating performance information, rendering a formal performance evaluation, generating and delivering performance feedback, the formal performance review meeting, and performance coaching); *inputs* (e.g., environmental context, strategy); *individuals* (characteristics of the people involved in the PM tasks, especially employees and managers); *formal processes* (formal procedures for how the PM tasks are conducted; the PM methods and approaches); *informal processes*

(unwritten or implicit elements that emerge over time, e.g., informal feedback norms); and *outputs* (e.g., performance ratings, feedback generated, creation of a development plan, career planning, administrative recommendations).

Our empirical review shows that *employee reactions* (see Appendix A, section Ia) are most influenced by *informal processes*, with research suggesting pretty clearly that more positive cognitive and utility reactions (as well as greater satisfaction) result when employees participate in the PM process, when they have knowledge about how the process works, and when they believe their supervisors are unbiased and fair. In fact, it appears that perceptions of fairness and accuracy in PM may depend as much on trust in the supervisor as on characteristics of the PM process itself (e.g., Fulk, Brief, & Barr, 1985). On the other hand, our review suggests that *manager reactions* (see Appendix A, section Ib) are more influenced by *formal processes*, include rating approach (e.g., Dale et al., 2013; Schleicher, Bull, & Green, 2009), as well as by managers' *individual* factors (e.g., previous PM experience, personality).

For *employee learning* criteria (see Appendix A, section II), both *informal processes* (e.g., delivery of feedback) and *formal processes* (e.g., type of evaluation) are important for motivational and skills-based learning. For *manager learning* criteria (see Appendix A, section III), *formal processes* appear most impactful, especially rater training. In fact, this focus characterizes the bulk of research in this area, and more work is needed on other aspects of PM likely to result in significant learning for managers, such as the experience of rating or the feedback session with employees. There is also some evidence that *individual* factors play a role here, but that is generally confined to effects on rating quality criteria.

For *employee transfer* (see Appendix A, section IV), overall the research suggests that the effects of PM can in fact transfer beyond the immediate PM context, to affect more general employee attitudes and behaviors. Yet what PM aspects are most impactful in this regard varies a bit across specific transfer criteria. For example, *turnover intentions* (see Appendix A., section IVc) are particularly impacted by due process elements of PM (implicating both *formal and informal processes*) and the reactions that accompany them. Both *formal and informal processes* are also important for *fairness/justice perceptions* (see Appendix A, section IVe). Yet for *employee motivation* (see Appendix A, section IVd), it appears to be the *task components of goal-setting and feedback* (components of more developmentally oriented PM systems) that appear most impactful, not specific processes (formal or informal) within these tasks. For *manager transfer* (especially the quality of relationship with employees, see Appendix A., section Va), our review again shows that both *informal and formal PM processes* can impact these relationships, either positively or negatively.

Finally, for *unit-level criteria*, the IV is usually the *PM system* (as opposed to components or processes of PM), especially given that such investigations are often conducted by strategic HR researchers. For example, the research on HCRs (see Appendix A, section VI) shows that PM systems (and FDRS systems specifically) can impact both ability-based and motivation-based HCRs (sometimes positively and sometimes negatively). Research has shown that emergence enablers too (see Appendix A, section VII) can result from the implementation of new PM systems and the general type of PM system ("high quality" PM, Searle & Ball, 2003). There is a clear need for future research to try to link more

specific processes or other components within PM systems to unit-level criteria. Related, we have observed in our review that the strategic HR research, especially that examining PM as part of HR bundles, tends to take a very simplistic approach to the measurement of PM systems (e.g., the percentage of employees who receive a formal PA). We believe strongly that a fuller understanding of the relationship between PM and firm performance requires a much more nuanced approach to measuring the PM construct. For example, the fact that research has been equivocal on the link between PM and creativity (e.g., Zhou & Shalley, 2003) can be explained by the strong likelihood that this relationship is determined not by the *existence* of PM, but rather by the *type* of PM.

Relationships Among Evaluative Criteria

Unfortunately, our review revealed insufficient empirical research exploring any of the longer mediational relationships in our model of evaluative criteria (per Figure 1); consequently, these are discussed at primarily a theoretical level in the final section of our article, which lays out specific propositions for some key value chains implicated in our model. However, there is research that has reported bivariate relationships among our evaluative criteria (often as incidental, as opposed to focal, results). We coded such relationships as part of our comprehensive review and then computed the average sample-weighted correlation for any criterion–criterion relationships with at least two samples. Figure 2 reports these correlations linking the evaluative criteria in our model, and Appendix B describes these findings in more detail. For the most part these results show sizable positive relationships between adjacent criteria and support the theoretical linkages between model components previously discussed. Although some of these estimates are based on a small number of samples, and many of

them are likely inflated from same-source/-method data, we still feel reporting these criterion–criterion estimates is useful for this review, particularly in terms of providing some preliminary evidence to serve as a foundation for the value chains discussed in the final section.

How We Study PM Criteria

In this final section of observations from our empirical review, we discuss several needed improvements in *how* we study PM effectiveness, including the measurement and conceptualization of criteria; the use of stronger designs and field contexts; and the simultaneous examination of employees and managers. Future research will need to improve in each of these areas in order to advance cumulative knowledge of PM effectiveness.

Measuring and conceptualizing criteria. From our empirical review we conclude that greater care must be taken in both the conceptualization and measurement of specific evaluative criteria. We highlight two examples here. First, our review revealed that the distinction between reactions and learning is not always clearly articulated in the PM literature. For example, a closer examination of an article purporting to measure feedback reactions (Johnson & Helgeson, 2002) shows that three distinct “reactions” variables were measured: agreement with the feedback, changes in self-esteem, and intentions to change behavior. Whereas the first is categorized as reactions in our model, the second and third would be considered learning. Similarly, Tziner, Latham, Price, and Haccoun (1996) examined a “usefulness for employee development” criterion by measuring employee satisfaction (a reactions criterion) as well as goal perception and the quality of goals set (both learning criteria in our model). These distinct types of criteria are likely to behave differently, a possibility supported by the different

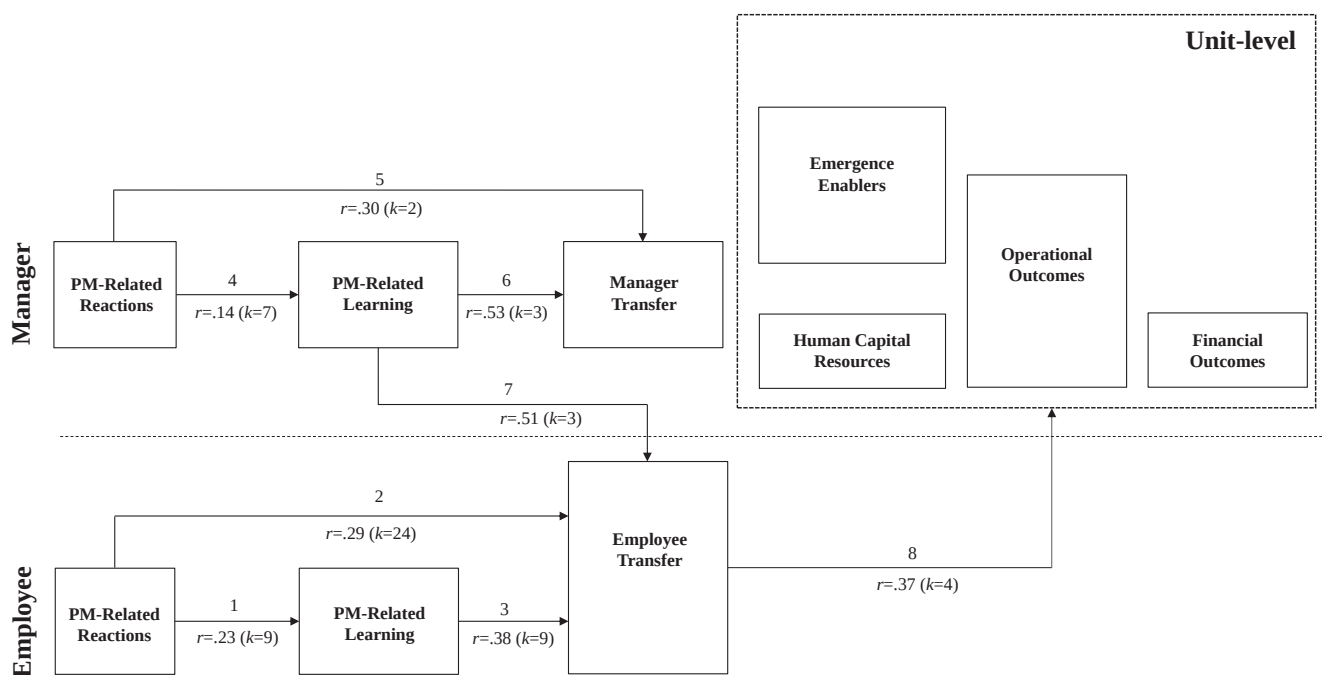


Figure 2. Average bivariate correlations between criterion categories.

results across variables in these two studies. We would encourage future researchers to avoid using more generic “reactions” labels for variables conceptually closer to learning criteria (including “buy-in” or commitment to the PM system). This advice is particularly important in light of the criterion–criterion relationships reported in Figure 2. These show relationships between reactions and learning of $r = .23$ and $.14$ only (for employees and managers, respectively), suggesting these are clearly distinct (albeit related) constructs. These results also show employee learning has a somewhat stronger relationship with employee transfer than does employee reactions, again affirming the need to accurately indicate whether one is measuring reactions or learning.

Second, our review also revealed a need to more clearly conceptualize and operationalize types of firm performance as an outcome of PM. We found a number of articles that examined aspects of organizational performance that could not be clearly categorized (see Appendix A, section VIII). Although almost all of these studies show a positive impact of PM on organizational performance, these types of measures and the related methods seem problematic for drawing clear conclusions. In particular, *perceptions* of organizational effectiveness (e.g., from HR executives or managers) are suspect as measures of actual effectiveness.

Using stronger research designs and contexts. There were a number of PM studies using relatively weak methods. This may explain why some criteria showed more equivocal results, such as with employee transfer (see Appendix A, section IV). Much of the research on transfer criteria was characterized by weaker designs (including common method issues, especially with job attitudes and fairness as transfer outcomes). We would also argue that transfer criteria, by definition, should probably not be studied in the lab (e.g., Holbrook, 1999).

Other criteria with an overreliance on lab research were learning and managerial transfer. A lot of the research on cognitive learning outcomes for employees has been conducted in the lab; such findings should be replicated in field studies, as our review suggests that effects are often smaller in these settings (e.g., Boswell & Boudreau, 2002; Morgan, Cannan, & Cullinane, 2005; Tjosvold & Halco, 1992). For managerial learning, over 50% of cognitive learning research (and almost 50% of skills-based learning research) has been conducted in the lab with students; thus, many of the results discussed in Appendix A (section III) need to be replicated with managers in organizational settings who may experience greater cognitive load and additional constraints and thus different learning processes. Even in the lab, increased attentional demands meant to emulate actual work settings have been found to affect results (e.g., Martell, 1991); this is likely even more pronounced with managers in the field. Finally, nearly one third of the quality of relationship studies (under managerial transfer) were conducted in the lab, which is concerning because this then typically represents a hypothetical (i.e., “paper people”) or extremely short-term (formed within hours or minutes) relationship. As such, it is unclear whether these findings would generalize to complex workplace relationships.

Learning criteria were also prone to common-method issues (cross-sectional, single-source designs), especially cognitive learning and its relation to employee transfer (attitudinal/motivational learning research was much more likely to employ time-lagged or experimental designs than other learning). In addition, the vast majority of skills-based learning research employs self-report for

this outcome; this is contrasted with the approach in the training literature, where skills-based learning often relies on observation by others (Kraiger et al., 1993). Incorporating others’ reports of employees’ PM learning will be important for future research in this area. With regard to manager learning, we found it interesting that research on rating quality (often as a result of rater training) was actually more likely than other forms of PM-related learning to use time-lagged, longitudinal, or quasi-experimental designs. Unfortunately, most of the field research on other aspects of manager learning relied on single-source, cross sectional surveys. These methodological issues for learning criteria need to be strengthened in future research. However, the most glaring issues of weak methodologies involved the unit-level criteria, where many of the studies use cross-sectional, single-source surveys.

Simultaneously examining employees and managers. A final need for future research is to measure both employee and manager criteria within the same study. Failure to do so was a primary limitation observed in the reactions literature in particular. For the quality of relationships as a PM criterion, although both employees (e.g., Dahling, Chau, & O’Malley, 2012; Taylor & Pierce, 1999) and managers (e.g., Taylor et al., 1995) have been used across studies as sources for measuring such quality, no single study has collected these relational criteria from both perspectives. This is problematic because the same PM factor (e.g., more frequent negative feedback) may differentially impact the manager–employee relationship, depending on perspective. Similar concerns have been cited in other manager–employee dyadic research (e.g., Matta, Scott, Koopman, & Conlon, 2015). PM research should also broaden to examine the reactions and behaviors of managers who are also ratees/employees of their own supervisors (Langan-Fox, Bell, McDonald, & Morizzi, 1996). We know that managers’ experiences as recipients of PM can affect their reactions and behavior while executing PM (see Latham, Budworth, Yanar, & Whyte, 2008), and we need additional research on such role duality.

An Agenda for PM Effectiveness Research and Practice: Understanding Key Value Chains

The previous sections suggest that although a lot of research has examined the impact of PM on separate evaluative criteria, there has been very little explicit focus on how the multiple criteria are interrelated and link together to form the “value chains” of PM. Thus, in this final section of our integrative review, we explicitly consider the longer value chains underlying PM effectiveness. This is not meant to be exhaustive with regard to all possible linkages in our model. Rather we organize our discussion of specific linkages around three questions with particular import for theory and practice: (a) *How do individual-level outcomes of PM emerge to become unit-level outcomes?* (b) *How essential are positive reactions to the overall effectiveness of PM?* and (c) *What is the value of a performance rating?* For each question we identify several propositions (listed in Table 4) that could and should be tested in future research and discuss the implications for the practice of PM in organizations (summarized in Table 5). As such, this section serves to illustrate how our model might be used productively by both researchers and practitioners to make grounded hypotheses about the PM value chains.

Table 4
Specific Propositions Underlying the PM Value Chains

How do individual-level outcomes of PM emerge to become unit-level outcomes?
Proposition 1: Both employee transfer and employee learning criteria can become unit-level HCRs. Effects of learning criteria on unit-level human capital may be partially mediated by transfer criteria but are unlikely to be fully mediated.
Proposition 2: PM is likely to lead to organizational performance outcomes via both ability-based unit-level HCRs and motivation-based unit-level HCRs.
Proposition 3: Emergence enablers are an important moderator between employee learning and transfer criteria, and unit-level HCRs. When emergence enablers are nonexistent or weak, these individual criteria are less likely to emerge as unit-level HCRs.
Proposition 4: The relationship between PM and organizational performance is mediated by (a) an “individual path,” whereby employee criteria (learning and transfer) mediate this relationship, and this link is moderated by emergence enablers; and by (b) an “emergence path,” whereby emergence enablers mediate this relationship.
Proposition 5: Manager transfer criteria are related to unit-level HCRs in multiple ways, including (a) via the impact on employee transfer (for both relationship and decision effectiveness criteria); (b) via the emergence enabler of unit-level quality of human capital decisions made (for manager decision effectiveness criteria); and (c) via other emergence enablers such as climate and trust in management (for manager relationship effectiveness criteria).
How essential are positive reactions to the overall effectiveness of PM?
Proposition 6: The relationship between employee reactions and unit-level criteria is mediated by (a) employee learning, (b) employee transfer, and (c) manager transfer criteria (especially quality of relationship with employee).
Proposition 7: Employee reactions should predict employee learning, with moderately-sized positive relationships.
Proposition 8: Employee learning outcomes are likely to at least partially mediate the relationship between employee reactions and employee transfer criteria.
Proposition 9: It is unlikely that positive employee reactions to PM are essential (i.e., a necessary condition) for learning.
Proposition 10: Employee reactions should be moderately and positively related to employee transfer. These relationships are likely to be stronger for job attitudes and employee well-being than they are for performance and other behaviors.
Proposition 11: It is unlikely that employee reactions are essential (i.e., a necessary condition) for employee transfer.
Proposition 12: Relationships between employee reactions and other criteria are likely to vary based on the nature of the reactions (affective vs. cognitive vs. utility). Relationships with transfer are likely to be strongest for utility and fairness (cognitive) reactions. Fairness may be more important than utility reactions for turnover intentions specifically.
Proposition 13: Managers’ PM-related learning is likely to be more strongly related to manager transfer and unit-level criteria than are managers’ reactions. The effects of managers’ reactions on transfer and unit-level criteria are likely to be mediated by managers’ PM-related learning.
What is the value of a performance rating?
Proposition 14: Engaging in the rating process can have a positive impact on managers’ PM-related learning.
Proposition 15: Eliminating performance ratings will increase the strength of the relationship between managers’ PM-related learning and other criteria
Proposition 16: The quality of human capital decisions (an emergence enabler) is likely to be lower in the absence of performance ratings, thus weakening the relationship between employee transfer criteria and unit-level HCRs and negatively impacting unit-level HCRs directly.

Note. PM = performance management; HCR = human capital resources.

How Do Individual-Level Outcomes of PM Emerge to Become Unit-Level Outcomes?

The link between PM systems and firm performance has been severely underspecified (empirically and conceptually) in the research literature (DeNisi & Smith, 2014). Our model can help shed light on this link and in turn also meaningfully contribute to the broader strategic HR literature, which has long been interested in better understanding the mediational processes (the so-called “black box”) linking HR practices with organizational performance (Becker & Huselid, 2006; Messersmith et al., 2011). We discuss two aspects of these mediational processes here: the relationship between individuals and unit-level HCRs, and the role of emergence enablers.

Individuals and unit-level HCRs. Like others (e.g., Daley, 1986; Kim et al., 2016; Zheng et al., 2006), we suggest that organizational performance is a function of unit-level HCRs, in this case HCRs that eventually emerge from aspects of PM processes. Ployhart and colleagues (Ployhart & Moliterno, 2011; Ployhart et al., 2014) argue that unit-level human capital has its origins in the *full range* of individual KSAOs. Unique to our model is the observation and acknowledgment that this full range of individual KSAOs includes both PM-specific KSAOs (those criteria in our learning category, such as increased self-awareness and

increased feedback-seeking) as well as broader job and organizationally relevant constructs (captured in our transfer category, including job attitudes, motivation, and performance). We argue that the impact of PM on unit-level HCRs operates through both types of criteria. Thus, our first proposition (see Table 4) recognizes the dual nature of the individual source underlying unit-level HCRs. This line of thinking suggests that PM may positively impact unit-level outcomes even if it does not improve employee performance. This is interesting given the frequent assertion that such improvement is the ultimate goal (e.g., DeNisi & Pritchard, 2006) and “has been the major focus” of PM (DeNisi & Smith, 2014, p. 133). Future research should empirically address this possibility.

Future research should also examine whether performance improvement (individual and/or unit levels) with PM is due to increases in competencies and skills or primarily to attitudinal or motivational effects. AMO researchers have suggested that the effects of PA on performance are primarily through motivation, and have consistently categorized PA as a motivation-enhancing HR practice as opposed to an ability/skills-enhancing HR practice (e.g., Chuang & Liao, 2010; Delery & Doty, 1996; Gong, Law, Chang, & Xin, 2009). Yet our review actually shows greater research support for the impact of PM on ability-based unit-level

Table 5

Key Practical Implications of Our Review

Organizations must first identify the relevant criteria for evaluating PM based on their objectives.

- There is no “ultimate criterion” for evaluating PM.
- The most relevant criteria depend on the objectives of the PM system and the specific effectiveness questions being asked. Objectives should be based on an organization’s strategy and key stakeholders and tied to the various purposes of PM (e.g., administrative, developmental).
- For example:
 - an emphasis on employees as stakeholders would suggest a focus on well-being outcomes (employee transfer) as criteria, whereas an emphasis on top management as stakeholders (or a push to establish clear financial returns from investment in PM) would suggest a focus on organizational performance outcomes; and
 - more developmental approaches to and purposes for PM would suggest a greater emphasis on learning criteria.

Organizations must measure more than one criterion category in evaluating their PM systems.

- Regardless of objectives for the PM system, it is unlikely that a single component in our evaluative model can provide a complete evaluation picture for organizations.
- This is especially true given the equifinality inherent in overall PM effectiveness.
- With interest in more distal criteria (e.g., organizational performance), measuring more proximal (intermediary) criteria becomes essential for both understanding how the distal criteria did (or did not) manifest and ruling out alternative explanations for changes in such distal outcomes.

Organizations must measure more than one dimension within a criterion category in evaluating PM.

- Similar to the above, it is very unlikely that a single reactions, or learning, or transfer measure, for example, can provide a complete picture for organizations of the impact of PM on that criterion category.
- With reactions, for example, measures need to go beyond basic satisfaction; fairness (cognitive) and utility-based reactions in particular appear quite impactful and should be included.
- The dimensions measured should also include a focus on specific referents (e.g., the feedback meeting, the manager’s role in this), not just the PM system overall, as data on specific referents are more helpful for improving the system.

Organizations must collect evaluation data from multiple sources.

- In particular, both employees and managers should be included in evaluation of PM (and, because their perspectives vary, data should be coded for source). This is especially important in measuring reactions, where managers have been significantly ignored relative to employees.
- Data from multiple sources is also key for learning criteria, where too often the emphasis has been on self-report. A complete view of learning must include observation by others.

Organizational interventions aimed at improving PM should focus on different levers based on the relevant criteria of interest.

- Our review shows that the aspects of PM that exert the biggest influence differ across the evaluative criteria. For example:
 - employee reactions are particularly influenced by informal processes (e.g., employee participation in PA, trust in supervisor);
 - manager reactions are particularly influenced by formal processes (e.g., rating approach);
 - employee turnover intentions are particularly impacted by due process elements (perceived fairness), not perceived value of PM; and
 - employee motivation is particularly impacted by provision of goal-setting and feedback.
- Organizations should use this information to choose the levers likely to be most impactful in improving criteria of interest.

Organizations should focus substantially more on learning criteria in evaluating PM systems.

- Learning criteria represent the greatest untapped potential in evaluating PM. This criterion category (for both employees and managers) involves more than just rating quality, and it serves as a key mediator in our model, linking to more distal criteria. For example:
 - what employees learn from PM (especially in terms of attitudinal and motivational learning) can transfer into improved attitudes and performance back on the job; and
 - managers’ PM-related learning can impact both employees’ attitudes and performance as well as the quality of decisions managers make. In fact, our review suggests it may be more important that managers “do PM well” (learning) than that they react positively to PM.
- The importance of learning outcomes in general will only increase over time, given the move towards more development-focused PM.

Organizations should rethink what reactions criteria mean and how they should be managed.

- Organizations appear singularly focused on employee reactions to PM, which are often negative (and cited as a reason for change). Yet our review shows that positive reactions, although relevant, are just one of many criteria of interest and are not, in fact, essential for PM effectiveness.
- Rather than focusing heavily on maximizing positive reactions, organizations might focus on helping employees work through negative reactions to PM. Companies and employees might learn to reframe negative reactions as okay, as long as learning occurs.

Organizations need to focus more on manager transfer variables, especially quality of decisions.

- Another criterion category underutilized in PM evaluation is manager transfer, especially the quality of decisions made.
- These manager transfer variables play multiple important roles in the relationship between PM and unit-level outcomes (e.g., firm performance) and therefore should be measured by organizations interested in such outcomes.
- Organizations should measure how PM design choices (especially changes in design) help (or hinder) managers in making better quality decisions about employees (e.g., who should be promoted). The assumption is that improved PM processes increase the quality of HR decisions, but this is largely untested because of the neglect of this criterion.

Organizations need to focus more on emergence enablers in evaluating their PM systems.

- Emergence enablers (especially culture, climate, and trust in management) should be included in organizations’ evaluation of their PM systems (especially for those organizations concerned about the PM-organizational performance link), for two reasons:
 - when these emergence enablers are nonexistent or weak, individual criteria are less likely to emerge as unit-level HCRs and thus unlikely to translate into organizational performance; and
 - PM processes themselves can directly affect these emergence enablers, either positively or negatively, which in turn can directly impact organizational performance (as well as weaken or strengthen the above link).

Note. PM = performance management.

HCRs than on motivation-based HCRs (see [Appendix A](#), section VI). More importantly, we believe this motivation-enhancing categorization is too simplistic for understanding the complex ways that PM may impact both individuals and organizations, and we would call for future research to empirically test some of these various paths. In particular it would be interesting to examine the relative importance of ability- versus motivation-based HCRs in explaining the PM-organizational performance link (see Proposition 2 in [Table 4](#)).

Emergence enablers. Emergence enablers are a critical moderator of the link between employee constructs and unit-level human capital; quite simply, they enable individual-level constructs to become unit-level phenomena. When emergence enablers are nonexistent or weak, any individual criteria would be less likely to emerge as unit-level HCRs (see Proposition 3, [Table 4](#)). We argue here that emergence enablers actually play multiple important roles in understanding the longer chains of how PM impacts organizational performance.

First, PM can affect organizational performance through its impact on employee learning and transfer (as shown in our review). The effect of these criteria on unit-level HCRs would then be moderated by emergence enablers. Second, our review also suggests that PM can affect emergence enablers directly (e.g., climate, coordination, trust in leadership), positively or negatively. Taken together these two points mean that PM essentially influences both the IV (employee learning or transfer) and the moderator (emergence enablers) in the overall employee to unit-level HCR relationship. Third, there is evidence that emergence enablers can themselves directly affect unit-level HCRs and related operational outcomes. For example, without sufficient cohesion (an emergence enabler), members can begin to question their involvement in the unit and withdraw from it ([Ployhart & Moliterno, 2011](#)), thus negatively affecting motivation-related HCRs and outcomes such as absenteeism and turnover. In addition, [Evans and Davis \(2005\)](#) have noted that positive changes in social structure (an emergence enabler) increase organizational flexibility and efficiency, which are key to operational outcomes. Thus, as [Ployhart and Moliterno \(2011\)](#) have suggested, to the extent that PM alters the way that unit members interact behaviorally, cognitively, and affectively, this necessarily would result in the emergence of different kinds of HCRs. We argue this is true in the context of PM both because these emergence enablers moderate the impact of employee criteria on unit-level HCRs and because they are directly linked to other unit-level phenomena (see Proposition 4 in [Table 4](#)).

Also unique to our article is a consideration of the role of managers in this emergence process. In our model, manager transfer criteria (which reflect the effectiveness of managers as managers) impact unit-level HCRs in multiple ways. First, improved managerial effectiveness can positively impact employees' attitudes and behaviors (employee transfer), the effect of which would then proceed via the relationships outlined above. Second, the quality of decisions that managers make about employees (which is a manager transfer criterion) should significantly impact the unit-level HCRs. This is because quality of decisions made by managers would aggregate to the quality of unit-level human capital decisions made (depicted as an emergence enabler in our model), which in turn determines the unit's ability to "leverage" the human capital available and turn it into a resource ([Lakshman,](#)

2014). Third, the quality of managers' relationships with employees (another manager transfer criterion), which our empirical review suggests is significantly impacted by PM processes, would aggregate at the unit-level into important emergence enablers such as climate and trust in management. In turn, these emergence enablers both moderate the relationship between employee criteria and unit-level HCRs and directly impact other unit-level outcomes, as specified above (see [Figure 1](#)). Thus, we argue that manager transfer criteria (in terms of both the quality of relationships with employees and the quality of decisions made about employees) play multiple important roles in the relationship between PM and unit-level outcomes (see Proposition 5 in [Table 4](#)). This is something that should be explicitly tested in future research.

We are aware of no other research that has attempted to articulate the impact of specific HR practices (as we do here for PM) on the emergence enabling process (see [Ployhart & Moliterno, 2011](#) on the general need for this), and we believe the specific propositions here offer important directions for future research on PM effectiveness. These arguments also have meaningful implications for practice (see [Table 5](#)). Organizations are likely to be particularly interested in the impact of PM on organizational performance, yet "80–90% of HR professionals consider that their PM system does not improve organizational performance" ([Haines & St-Onge, 2012](#), p. 1158). We suspect it is unlikely that most HR professionals have the data to support this link one way or the other, as establishing it can be very complex. The difficulty of linking aspects of operational PM systems to very distal outcomes while ruling out other possible explanations for the effects (i.e., threats to validity, [Cook & Campbell, 1979](#)) requires thoughtful consideration of the underlying intermediary processes. We believe that our articulation of some of these processes can be useful in practice, suggesting that organizations should measure the following outcomes of PM in order to understand whether and how PM is resulting in improved organizational performance: employee PM-related learning, employee transfer (at least some ability/skill- and some motivational-related criteria), managerial decision-making and relationship quality, and culture and climate constructs.

How Essential Are Positive Reactions to the Overall Effectiveness of PM?

Reports from the popular press suggest that employees and managers alike downright detest their PM systems. As [Levy, Tseng, Rosen, and Lueke \(2017, p. 156\)](#) recently noted "... you can do a simple Google search and tap into the uproar." For their part, organizations appear very sensitive to these negative reactions, often citing them as reasons for modifications to their PM systems (see [Corporate Leadership Council, 2012](#)). Our review of the research (see [Appendix A](#), section i) suggests that negative reactions are, in fact, quite prevalent. For example, reports of procedural injustice are frequent, and affective reactions to PM are quite negative, especially as a result of negative feedback, and even when perceived importance is quite high. Given the extent of negative reactions, it is critical to understand how such reactions relate to other PM effectiveness criteria. In this section we integrate the research findings with our overall model to break down what has been a widespread assumption in the research literature (and perhaps in practice as well, judging from organizations'

actions): that positive reactions are essential to the effectiveness of PM. For example, [Murphy and Cleveland \(1995\)](#) stated in 1995 that “reaction criteria are almost always relevant, and an unfavorable reaction may doom the most carefully constructed appraisal system” (p. 314), and this was later reiterated by [Keeping and Levy \(2000\)](#) in their review of PA reactions. We examine this assumption here in a more detailed manner, explicating through what paths reactions might exert their influence on more distal criteria and also describing paths that likely accrue value without necessitating positive reactions. We consider this first for employee reactions and then for manager reactions.

Employee reactions. Some strong statements have been made about the importance of employee reactions in the effectiveness of PM. For example:

The effectiveness of appraisal and feedback *depends substantially* on the extent to which ratees accept the appraisal system. ([Tziner et al., 1996](#), p. 177, emphasis added)

With dissatisfaction and feelings of unfairness in process and inequity in evaluations, any appraisal system *will be doomed to failure*. ([Cardy & Dobbins, 1994](#), p. 54, emphasis added)

The empirical evidence and our model suggest that employee reactions are likely to relate to unit-level effectiveness in multiple ways (i.e., via multiple paths in our model, including through employee learning, employee transfer, and manager transfer; see Proposition 6 in [Table 4](#) and below). At the same time, the equifinality inherent in our model also suggests little reason to believe that any PM system is “doomed to failure” without positive employee reactions.

Within employee-level criteria, there is likely to be a positive relationship between employee reactions to PM and both employee learning from aspects of PM and employee transfer criteria. Such relationships have certainly been suggested and found in the training evaluation literature ([Alliger et al., 1997](#); [Kirkpatrick, 1987](#)), and our empirical results confirm this. First, as [Figure 2](#) shows, PM research suggests a moderately positive relationship ($r = .23$) between employee reactions and learning (especially motivation to improve as a result of the PM). It is also likely that these learning criteria mediate the relationship between reactions and behavioral responses to PM. Such a prediction is in-line both with arguments in the training evaluation literature and with PM-specific empirical findings by [Kinicki, Prussia, Wu, and McKee-Ryan \(2004\)](#), who found that such constructs completely mediated the relationship between reactions to feedback and behavioral responses to it. At the same time, our review suggests that the employee reactions–learning relationship is not so strong as to suggest that positive reactions are a necessity for learning. In addition, it appears that some positive reactions (i.e., believing that PM is distributively fair) can actually reduce some learning outcomes such as self-efficacy for improvement (e.g., [Taylor, Masterson, Renard, & Tracy, 1998](#)). [Table 4](#) summarizes the above arguments into propositions (Propositions 7, 8, and 9) that should be further tested with empirical work.

Second, our review also shows a moderate positive relationship between employee reactions and transfer ($r = .29$, see [Figure 2](#)). This estimate is mainly based on relationships between reactions and job attitudes such as job satisfaction and organizational commitment (subject to method bias). It is less clear from the empirical

PM research how employee reactions impact employee performance and other behavioral outcomes. It is also unknown whether the magnitude of these reactions–transfer relationships would hold if one accounted for employee learning, given its likely mediating role (see above). We do believe, however, that employee reactions would likely be relevant for transfer constructs related to employee well-being, a category we proposed as part of our model but one with relatively little empirical research in the PM literature. We would encourage future research to examine each of these probable relationships (see Propositions 10 and 11 in [Table 4](#)).

Finally, based on both our review as well as work in other areas, we argue that the relationship between employee reactions and other criteria is likely to depend on the nature of the reactions. For example, our model draws a distinction between affective, cognitive, and utility-based reactions; these various types of reactions have been shown in different contexts to have differential effects on other criteria ([Alliger et al., 1997](#)). In particular, utility reactions may exhibit stronger relationships with performance and other behavioral outcomes than other types of reactions ([Alliger et al., 1997](#), “What we think is useful may correlate with what we use,” p. 352). Yet our review also suggests there are likely to be particularly strong effects for fairness reactions (a cognitive reaction in our model), including due process perceptions, especially in terms of relationships with transfer criteria such as turnover intentions. In fact, several pieces of evidence converge to suggest it is in fact the “due process” and perceived fairness aspects of PM, as opposed to perceived value created by the PM system, that drive relationships with employee turnover intentions specifically (see [Burke, 1996](#); [Poon, 2004](#); [Si & Li, 2012](#), all in [Appendix B](#)). These arguments are summarized in Proposition 12 in [Table 4](#); they suggest that future research (as well as practitioners interested in evaluating PM) should jointly examine multiple types of reactions to better understand their relevance. The same could be said for referent of reactions: system, specific events or aspects of PM, or the person (e.g., the manager implementing PM). Our model is meant to apply to all such referents of reactions, but we know little about whether the referent matters for the relationship with downstream criteria and would encourage future work on this. From a practical perspective we would strongly encourage organizations to measure reactions to specific aspects of PM, rather than the overall system, as the former provides better information for making improvements (see [Table 5](#)).

Manager reactions. Both the practice and scholarly literatures have suggested that the effectiveness of PM depends greatly on managers. For example, [den Hartog, Boselie, and Paaauwe \(2004\)](#) noted that “Most PM practices . . . are facilitated and implemented by direct supervisors or front-line managers. Therefore, the behavior of line managers will mediate the effect of (most) practices on employee perception (and behavior)” (p. 565), and “Without managers’ support and cooperation, it is unlikely that employees can experience fairness in organizational HR systems” (p. 568). We agree that the role of managers in PM is paramount and we strongly encourage additional research that highlights this. But does this mean that *positive reactions* by managers are *necessary* for PM to have value and/or to affect other criteria? Although our model and empirical review suggest there are multiple ways in which manager reactions can relate to other criteria (see [Figures 1 and 2](#)), the empirical results suggest quite modest relationships (with the exception of manager reactions to manager transfer, but this was based on only two studies published together in

a single article and subject to common method variance, see [Appendix B](#)).

We believe, based on both our review and theoretical model, that it may be more important that managers “do PM well” (captured by manager learning criteria) than that they react positively to PM (see Proposition 13 in [Table 4](#)). Although learning criteria have not been previously explicitly discussed in PM and there needs to be substantially more empirical attention paid to this, some indirect evidence exists for this link via a validation study of the Performance Management Behavior Questionnaire ([Kinicki, Jacobson, Peterson, & Prussia, 2013](#)). Five dimensions of this scale represent manager PM-related learning, and across multiple samples all five were linked to reports of managers’ general effectiveness, a key aspect of manager transfer. This suggests that managers’ PM-related learning is likely a critical criterion in the value chain of PM, as does our empirical review of criterion–criterion relationships (see [Figure 2](#)), which shows that managers’ PM-related learning actually demonstrates stronger relationships with employee transfer outcomes (a cross-level effect) than does employee learning (a within-level effect).

The arguments presented here about employee and manager reactions suggest first that organizations should acknowledge and understand the equifinality present in the value chain(s) leading from PM to organizational performance and not overemphasize the role of reactions in an effective PM system (by, e.g., scrapping a PM system because of negative reactions to it, as has been reported). Reactions should of course be measured, but they are simply one of many relevant criteria. In addition, rather than focusing so heavily on maximizing positive reactions, organizations might focus on helping employees work through negative reactions (especially short-term negative affective reactions that appear unavoidable in the face of negative feedback). Companies and employees might learn to even reframe negative reactions as okay, as long as learning occurs. Moreover, when reactions are measured, it should be from the perspective of both employees and managers within the PM system, and organizations are well-advised to code survey responses for this status, given that the two sets of reactions tend to differ and differentially affect other criteria in the value chain. Also, a variety of types of reactions measures (i.e., affective, cognitive, utility-based) should be included; fairness (cognitive) and utility-based reactions in particular appear quite impactful and might receive extra attention.

What Is the Value of a Performance Rating?

PM practice has seen major changes in recent years, and one of the most salient has been eliminating annual performance ratings. Companies ranging from the technology sector to professional service firms to manufacturing are eliminating their formal performance ratings ([Cappelli & Tavis, 2016](#)), and this trend has resulted in a heated debate that spans practice and research (“The pros and cons of retaining performance ratings were the subject of a lively, standing-room-only debate at the 2015 Society for Industrial and Organizational Psychology conference in Philadelphia,” [Adler et al., 2016](#), p. 222). This is an area where practice has far outpaced research, and there is unfortunately very little empirical work examining what the impact of such a practice might be. Yet because the discourse surrounding the benefits and disadvantages of eliminating ratings implicates many of the components in our criterion model, it would seem that our model and corresponding review may have something to say about this debate.

Justifications for the elimination of performance ratings within PM have included negative reactions from employees and managers to this component (the necessity of which was addressed in the prior section), as well as extensive evidence that performance ratings are never as reliable or valid as we would like them to be (see [Adler et al., 2016](#)), criteria categorized as manager learning in our model. [Adler et al. \(2016\)](#) go on to conclude that no previous review “leads to the conclusion that performance rating is particularly successful either as a tool for accurately measuring employee performance or as a component of a broader program of [PM]” (p. 223). After our own very comprehensive review of the literature, we simply do not see any sufficient empirical basis for deciding whether or not performance rating adds value to PM systems, especially given the limited ways in which value has been operationalized in prior PM research. We can, however, through our model identify a few paths through which performance ratings might add value (see Propositions 14–16 in [Table 4](#)), possibilities that should be empirically studied in future research and in practice, in order to better inform this debate.

First is the likely possibility that managers can learn from the process of rating performance. Engaging in this could build managers’ PM-related skills and knowledge (as suggested indirectly by [Kinicki et al., 2013](#); [Longenecker, Liverpool, & Wilson, 1988](#); [Spence & Keeping, 2011](#)). We have previously established this criterion as quite important (if empirically understudied) in the value chain of PM, as it impacts both managerial transfer (the quality of decisions made about employees and relationships with employees) and employee transfer. Interestingly, if performance rating is removed from a PM system, it may make the other criteria under manager learning even more important. This is because it is “naïve to think that, relieved of the burden of ratings and without the ‘crutch’ of a structured feedback tool, managers will somehow overcome this weakness and consistently engage in positive and impactful conversations. Indeed, one important mechanism for assuring that quality ongoing performance conversations occur . . . is by setting goals as well as evaluating and rewarding managers for how effectively they manage the performance of their own subordinates” ([Adler et al., 2016](#), p. 239).

Second, the existence of performance ratings seems important for making decisions about human capital (an important manager transfer criterion and unit-level criterion as well). [Adler et al. \(2016\)](#) themselves admit the heavy reliance of decisions on performance ratings (e.g., “It is fair to say that tens if not hundreds of billions of dollars in compensation and rewards are riding on the backs of performance ratings,” p. 223). Thus, in the absence of performance ratings, it is unclear how (well) such decisions might be made, thus diminishing an important emergence enabler (and moderator) in our model (see [Figure 1](#)). Future research should specifically test what happens to the quality of human capital decisions in the absence of formal ratings.

Finally, the relevance of the psychometric quality of ratings (an important focus in this debate) can also be framed in terms of our model. As suggested throughout this review, there is a great deal of “mediational equifinality” in the value chains of PM (i.e., multiple mediational paths for the more proximal evaluative criteria). This is contrasted with claims sometimes made in the PM literature. For example, [DeNisi and Smith \(2014\)](#) note that the *only* way rating accuracy matters is if it affects employee motivation for improvement, via perceived fairness. But the causal chain is likely more complicated than this; accuracy could also positively impact things like developmental or training assignments, job tasks, or relocation based on poor fit, all of which could impact performance improvement, via the

quality of decisions made about employees. Our model and review serve to make this equifinality explicit, and we would encourage future research that examines competing mediational mechanisms such as these.

Conclusions

This review sets forth a theoretically grounded, comprehensive, and integrative model for understanding and measuring PM effectiveness. In using this model as a framework for reviewing and synthesizing the empirical research in PM, we find that although there has been a great deal of empirical work on the relationship between aspects of PM and each evaluative criterion considered separately, very little work has examined the longer “value chains” of PM. This represents an important opportunity for future work. We believe that this model and review (including the propositions we develop) can be very helpful for advancing both research and practice in PM, moving the field from more simplistic questions like “Is PM effective?” and “What is the ultimate criterion for PM?” to more nuanced and fruitful inquiries regarding *how* PM creates value and *for whom*.

References

- Adler, S., Campion, M., Colquitt, A., Grubb, A., Murphy, K., Ollander-Krane, R., & Pulakos, E. D. (2016). Getting rid of performance ratings: Genius or folly? A debate. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 9, 219–252. <http://dx.doi.org/10.1017/iop.2015.106>
- Aguinis, H. (2013). *Performance management* (3rd ed.). Upper Saddle River, NJ: Pearson.
- Alliger, G. M., & Janak, E. A. (1989). Kirkpatrick's levels of training criteria: Thirty years later. *Personnel Psychology*, 42, 331–342. <http://dx.doi.org/10.1111/j.1744-6570.1989.tb00661.x>
- Alliger, G. M., Tannenbaum, S. I., Bennett, W., Jr., Traver, H., & Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology*, 50, 341–358. <http://dx.doi.org/10.1111/j.1744-6570.1997.tb00911.x>
- Anseel, F., Lievens, F., & Schollaert, E. (2009). Reflection as a strategy to enhance task performance after feedback. *Organizational Behavior and Human Decision Processes*, 110, 23–35. <http://dx.doi.org/10.1016/j.obhdp.2009.05.003>
- Armstrong-Stassen, M., & Schlosser, F. (2010). When hospitals provide HR practices tailored to older nurses, will older nurses stay? It may depend on their supervisor. *Human Resource Management Journal*, 20, 375–390. <http://dx.doi.org/10.1111/j.1748-8583.2010.00143.x>
- Atwater, L. E., & Brett, J. F. (2006). 360-degree feedback to leaders: Does it relate to changes in employee attitudes? *Group & Organization Management*, 31, 578–600. <http://dx.doi.org/10.1177/1059601106286887>
- Austin, J. T., & Villanova, P. (1992). The criterion problem: 1917–1992. *Journal of Applied Psychology*, 77, 836–874. <http://dx.doi.org/10.1037/0021-9010.77.6.836>
- Ayers, R. S. (2013). Building goal alignment in federal agencies' performance appraisal programs. *Public Personnel Management*, 42, 495–520. <http://dx.doi.org/10.1177/0091026013496077>
- Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, 41, 63–105. <http://dx.doi.org/10.1111/j.1744-6570.1988.tb00632.x>
- Balfour, D. L. (1992). Impact of agency investment in the implementation of performance appraisal. *Public Personnel Management*, 21, 1–15. <http://dx.doi.org/10.1177/009102609202100101>
- Bambacas, M., & Kulik, T. C. (2013). Job embeddedness in China: How HR practices impact turnover intentions. *International Journal of Human Resource Management*, 24, 1933–1952. <http://dx.doi.org/10.1080/09585192.2012.725074>
- Banks, C. G., & Murphy, K. R. (1985). Toward narrowing the research-practice gap in performance appraisal. *Personnel Psychology*, 38, 335–345. <http://dx.doi.org/10.1111/j.1744-6570.1985.tb00551.x>
- Bannister, B. D. (1986). Performance outcome feedback and attributional feedback: Interactive effects on recipient responses. *Journal of Applied Psychology*, 71, 203–210. <http://dx.doi.org/10.1037/0021-9010.71.2.203>
- Barney, J. B., & Wright, P. M. (1998). On becoming a strategic partner: The role of human resources in gaining competitive advantage. *Human Resource Management*, 37, 31–46. [http://dx.doi.org/10.1002/\(SICI\)1099-050X\(199821\)37:1<31::AID-HRM4>3.0.CO;2-W](http://dx.doi.org/10.1002/(SICI)1099-050X(199821)37:1<31::AID-HRM4>3.0.CO;2-W)
- Barrick, M. R., Thurgood, G. R., Smith, T. A., & Courtright, S. H. (2015). Collective organizational engagement: Linking motivational antecedents, strategic implementation, and firm performance. *Academy of Management Journal*, 58, 111–135. <http://dx.doi.org/10.5465/amj.2013.0227>
- Bartol, K. M., Durham, C. C., & Poon, J. M. (2001). Influence of performance evaluation rating segmentation on motivation and fairness perceptions. *Journal of Applied Psychology*, 86, 1106–1119. <http://dx.doi.org/10.1037/0021-9010.86.6.1106>
- Batt, R. (2002). Managing customer services: Human resource practices, quit rates, and sales growth. *Academy of Management Journal*, 45, 587–597.
- Becker, B. E., & Huselid, M. A. (1998). High performance work systems and firm performance: A synthesis of research and managerial implications. *Research in Personnel and Human Resources Management*, 16, 53–101.
- Becker, B. E., & Huselid, M. A. (2006). Strategic human resources management: Where do we go from here? *Journal of Management*, 32, 898–925. <http://dx.doi.org/10.1177/0149206306293668>
- Beer, M., Spector, B., Lawrence, P. R., Mills, D. Q., & Walton, R. E. (1984). *Managing human assets: The groundbreaking Harvard Business School program*. New York, NY: Free Press.
- Belschak, F. D., & Den Hartog, D. N. (2009). Consequences of positive and negative feedback: The impact on emotions and extra-role behaviors. *Applied Psychology: An International Review*, 58, 274–303. <http://dx.doi.org/10.1111/j.1464-0597.2008.00336.x>
- Bhave, D. P. (2014). The invisible eye? Electronic performance monitoring and employee job performance. *Personnel Psychology*, 67, 605–635.
- Bhave, D. P., & Brutus, S. (2011). A macro perspective to micro issues. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 4, 165–168. <http://dx.doi.org/10.1111/j.1754-9434.2011.01316.x>
- Bipp, T., & Kleingeld, A. (2011). Goal-setting in practice: The effects of personality and perceptions of the goal-setting process on job satisfaction and goal commitment. *Personnel Review*, 40, 306–323. <http://dx.doi.org/10.1108/0048348111118630>
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). San Francisco, CA: Jossey-Bass.
- Blume, B. D., Rubin, R. S., & Baldwin, T. T. (2013). Who is attracted to an organization using a forced distribution performance management system? *Human Resource Management Journal*, 23, 360–378. <http://dx.doi.org/10.1111/1748-8583.12016>
- Boswell, W. R., & Boudreau, J. W. (2002). Separating the developmental and evaluative performance appraisal uses. *Journal of Business and Psychology*, 16, 391–412. <http://dx.doi.org/10.1023/A:1012872907525>
- Bouskila-Yam, O., & Kluger, A. N. (2011). Strength-based performance appraisal and goal setting. *Human Resource Management Review*, 21, 137–147. <http://dx.doi.org/10.1016/j.hrmr.2010.09.001>

- Brett, J. F., & Atwater, L. E. (2001). 360° feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology*, 86, 930–942. <http://dx.doi.org/10.1037/0021-9010.86.5.930>
- Bretz, R. D., Jr., Milkovich, G. T., & Read, W. (1992). The current state of performance appraisal research and practice: Concerns, directions, and implications. *Journal of Management*, 18, 321–352. <http://dx.doi.org/10.1177/014920639201800206>
- Brown, M., Hyatt, D., & Benson, J. (2010). Consequences of the performance appraisal experience. *Personnel Review*, 39, 375–396. <http://dx.doi.org/10.1108/00483481011030557>
- Brutus, S., Fletcher, C., & Baldry, C. (2009). The influence of independent self-construal on rater self-efficacy in performance appraisal. *International Journal of Human Resource Management*, 20, 1999–2011. <http://dx.doi.org/10.1080/09585190903142431>
- Buckingham, M., & Goodall, A. (2015). Reinventing performance management. *Harvard Business Review*, 93, 40–50.
- Burke, R. J. (1996). Performance evaluation and counselling in a professional services firm. *Leadership and Organization Development Journal*, 17, 21–26. <http://dx.doi.org/10.1108/01437739610116957>
- Cappelli, P., & Tavis, A. (2016). The performance management revolution. *Harvard Business Review*, 94, 58–67.
- Cardy, R. L., & Dobbins, G. H. (1994). *Performance appraisal: Alternative perspectives*. Cincinnati, OH: South-Western.
- Catano, V. M., Darr, W., & Campbell, C. A. (2007). Performance appraisal of behavior-based competencies: A reliable and valid procedure. *Personnel Psychology*, 60, 201–230. <http://dx.doi.org/10.1111/j.1744-6570.2007.00070.x>
- Cawley, B. D., Keeping, L. M., & Levy, P. E. (1998). Participation in the performance appraisal process and employee reactions: A meta-analytic review of field investigations. *Journal of Applied Psychology*, 83, 615–633. <http://dx.doi.org/10.1037/0021-9010.83.4.615>
- Chalykoff, J., & Kochan, T. A. (1989). Computer-aided monitoring: Its influence on employee job satisfaction and turnover. *Personnel Psychology*, 42, 807–834. <http://dx.doi.org/10.1111/j.1744-6570.1989.tb00676.x>
- Cheng, S. (2014). The mediating role of organizational justice on the relationship between administrative performance appraisal practices and organizational commitment. *International Journal of Human Resource Management*, 25, 1131–1148. <http://dx.doi.org/10.1080/09585192.2013.816864>
- Chuang, C.-H., & Liao, H. (2010). Strategic human resource management in service context: Taking care of business by taking care of employees and customers. *Personnel Psychology*, 63, 153–196. <http://dx.doi.org/10.1111/j.1744-6570.2009.01165.x>
- Clarke, C., Harcourt, M., & Flynn, M. (2013). Clinical governance, performance appraisal and interactional and procedural fairness at a New Zealand public hospital. *Journal of Business Ethics*, 117, 667–678. <http://dx.doi.org/10.1007/s10551-012-1550-9>
- Coff, R., & Kriscynski, D. (2011). Invited editorial: Drilling for micro-foundations of human capital-based competitive advantages. *Journal of Management*, 37, 1429–1443. <http://dx.doi.org/10.1177/0149206310397772>
- Combs, J., Liu, Y., Hall, A., & Ketchen, D. (2006). How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance. *Personnel Psychology*, 59, 501–528. <http://dx.doi.org/10.1111/j.1744-6570.2006.00045.x>
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Boston, MA: Houghton Mifflin.
- Corporate Leadership Council. (2012). *Improving employee performance in the economic downturn. Four cost-effective strategies to improve employee outcomes*. Arlington, VA: Corporate Executive Board.
- Culig, K. M., Dickinson, A. M., Lindstrom-Hazel, D., & Austin, J. (2008). Combining workstation design and performance management to increase ergonomically correct computer typing postures. *Journal of Organizational Behavior Management*, 28, 146–175. <http://dx.doi.org/10.1080/01608060802251064>
- Dahling, J. J., Chau, S. L., & O'Malley, A. (2012). Correlates and consequences of feedback orientation in organizations. *Journal of Management*, 38, 531–546. <http://dx.doi.org/10.1177/0149206310375467>
- Dale, J. C., Drews, B., Dimmitt, P., Hildebrandt, E., Hittle, K., & Tielsch-Goddard, A. (2013). Novice to expert: The evolution of an advanced practice evaluation tool. *Journal of Pediatric Health Care*, 27, 195–201. <http://dx.doi.org/10.1016/j.pedhc.2011.12.004>
- Daley, D. M. (1986). Humanistic management and organizational success: The effect of job and work environment characteristics on organizational effectiveness, public responsiveness, and job satisfaction. *Public Personnel Management*, 15, 131–142. <http://dx.doi.org/10.1177/009102608601500204>
- David, E. M. (2013). Examining the role of narrative performance appraisal comments on performance. *Human Performance*, 26, 430–450. <http://dx.doi.org/10.1080/08959285.2013.836197>
- Davis, B. L., & Mount, M. K. (1984). Effectiveness of performance appraisal training using computer assisted instruction and behavior modeling. *Personnel Psychology*, 37, 439–452. <http://dx.doi.org/10.1111/j.1744-6570.1984.tb00521.x>
- Day, J. W., Holladay, C. L., Johnson, S. K., & Barron, L. G. (2014). Organizational rewards: Considering employee need in allocation. *Personnel Review*, 43, 74–95. <http://dx.doi.org/10.1108/PR-09-2012-0156>
- Delery, J. E., & Doty, D. H. (1996). Modes of theorizing in strategic human resource management: Tests of universalistic, contingency, and configurational performance predictions. *Academy of Management Journal*, 39, 802–835.
- Delery, J. E., & Shaw, J. D. (2001). The strategic management of people in work organizations: Review, synthesis, and extension. In G. R. Ferris (Ed.), *Research in personnel and human resources management* (Vol. 20, pp. 165–197). Stamford, CT: JAI Press.
- den Hartog, D. N., Boselie, P., & Paauwe, J. (2004). Performance management: A model and research agenda. *Applied Psychology*, 53, 556–569. <http://dx.doi.org/10.1111/j.1464-0597.2004.00188.x>
- DeNisi, A. S., & Murphy, K. R. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology*, 102, 421–433. <http://dx.doi.org/10.1037/apl0000085>
- DeNisi, A. S., & Peters, L. H. (1996). Organization of information in memory and the performance appraisal process: Evidence from the field. *Journal of Applied Psychology*, 81, 717–737. <http://dx.doi.org/10.1037/0021-9010.81.6.717>
- DeNisi, A. S., & Pritchard, R. D. (2006). Performance appraisal, performance management and improving individual performance: A motivational framework. *Management and Organization Review*, 2, 253–277. <http://dx.doi.org/10.1111/j.1740-8784.2006.00042.x>
- DeNisi, A. S., & Smith, C. E. (2014). Performance appraisal, performance management, and firm-level performance: A review, a proposed model, and new directions for future research. *The Academy of Management Annals*, 8, 127–179. <http://dx.doi.org/10.5465/19416520.2014.873178>
- DeNisi, A. S., & Sonesh, S. (2011). The appraisal and management of performance at work. In S. Zedeck (Ed.), *Handbook of industrial and organizational psychology* (pp. 255–280). Washington, DC: APA Press. <http://dx.doi.org/10.1037/12170-009>
- Dusterhoff, C., Cunningham, J., & MacGregor, J. (2014). The effects of performance rating, leader-member exchange, perceived utility, and organizational justice on performance appraisal satisfaction: Applying a moral judgment perspective. *Journal of Business Ethics*, 119, 265–273. <http://dx.doi.org/10.1007/s10551-013-1634-1>
- Dyer, L., & Reeves, T. (1995). Human resource strategies and firm performance: What do we know and where do we need to go? *International Journal of Human Resource Management*, 6, 656–670. <http://dx.doi.org/10.1080/095851995000000041>

- Eberhardt, B. J., & Pooyan, A. (1988). The effects of appraisal system redesign on perceptions of and satisfaction with performance appraisal: A quasi-experiment. *Journal of Business and Psychology*, 3, 230–241. <http://dx.doi.org/10.1007/BF01014491>
- Effron, M., & Ort, M. (2010). *One page talent management: Eliminating complexity, adding value*. Boston, MA: Harvard Business School Publishing.
- Elicker, J. D., Levy, P. E., & Hall, R. J. (2006). The role of leader-member exchange in the performance appraisal process. *Journal of Management*, 32, 531–551. <http://dx.doi.org/10.1177/0149206306286622>
- Erdemli, Ç., Sümer, H. C., & Bilgiç, R. (2007). A comparison of written feedback and written plus verbal feedback methods in performance management. *Turkish Journal of Psychology*, 22, 71–85.
- Erdogan, B. (2002). Antecedents and consequences of justice perceptions in performance appraisals. *Human Resource Management Review*, 12, 555–578. [http://dx.doi.org/10.1016/S1053-4822\(02\)00070-0](http://dx.doi.org/10.1016/S1053-4822(02)00070-0)
- Erdogan, B., Kraimer, M. L., & Liden, R. C. (2001). Procedural justice as a two-dimensional construct: An examination in the performance appraisal context. *The Journal of Applied Behavioral Science*, 37, 205–222. <http://dx.doi.org/10.1177/0021886301372004>
- Evans, E. M., & McShane, S. L. (1988). Employee perceptions of performance appraisal fairness in two organizations. *Canadian Journal of Behavioural Science*, 20, 177–191. <http://dx.doi.org/10.1037/h0079926>
- Evans, W. R., & Davis, W. D. (2005). High-performance work systems and organizational performance: The mediating role of internal social structure. *Journal of Management*, 31, 758–775. <http://dx.doi.org/10.1177/0149206305279370>
- Famdale, E., Hope-Hailey, V., & Kelliher, C. (2011). High commitment performance management: The roles of justice and trust. *Personnel Review*, 40, 5–23. <http://dx.doi.org/10.1108/00483481111095492>
- Famdale, E., & Kelliher, C. (2013). Implementing performance appraisal: Exploring the employee experience. *Human Resource Management*, 52, 879–897. <http://dx.doi.org/10.1002/hrm.21575>
- Favero, J. L., & Ilgen, D. R. (1989). The effects of rater prototypicality on rater observation and accuracy. *Journal of Applied Social Psychology*, 19, 932–946. <http://dx.doi.org/10.1111/j.1559-1816.1989.tb01230.x>
- Fedor, D. B., Davis, W. D., Maslyn, J. M., & Mathieson, K. (2001). Performance improvement efforts in response to negative feedback: The roles of source power and recipient self-esteem. *Journal of Management*, 27, 79–97. <http://dx.doi.org/10.1177/014920630102700105>
- Findley, H. M., Giles, W. F., & Mossholder, K. W. (2000). Performance appraisal process and system facets: Relationship with contextual performance. *Journal of Applied Psychology*, 85, 634–640. <http://dx.doi.org/10.1037/0021-9010.85.4.634>
- Ford, J. K., & Weissbein, D. A. (1997). Transfer of training: An updated review and analysis. *Performance Improvement Quarterly*, 10, 22–41. <http://dx.doi.org/10.1111/j.1937-8327.1997.tb00047.x>
- Fulk, J., Brief, A. P., & Barr, S. H. (1985). Trust-in-supervisor and perceived fairness and accuracy of performance evaluations. *Journal of Business Research*, 13, 301–313. [http://dx.doi.org/10.1016/0148-2963\(85\)90003-7](http://dx.doi.org/10.1016/0148-2963(85)90003-7)
- Gabris, G. T., & Ihrke, D. M. (2001). Does performance appraisal contribute to heightened levels of employee burnout? The results of one study. *Public Personnel Management*, 30, 157–172. <http://dx.doi.org/10.1177/009102600103000203>
- Galang, M. C. (2004). The transferability question: Comparing HRM practices in the Philippines with the U.S. and Canada. *International Journal of Human Resource Management*, 15, 1207–1233. <http://dx.doi.org/10.1080/0958519042000238419>
- Gavino, M. C., Wayne, S. J., & Erdogan, B. (2012). Discretionary and transactional human resource practices and employee outcomes: The role of perceived organizational support. *Human Resource Management*, 51, 665–686. <http://dx.doi.org/10.1002/hrm.21493>
- Girometti, G. W., Schroeder, A. N., & Switzer, F. S. (2015). Forced distribution rating systems: When does “rank and yank” lead to adverse impact? *Journal of Applied Psychology*, 100, 180–193. <http://dx.doi.org/10.1037/a0037191>
- Goerke, M., Möller, J., Schulz-Hardt, S., Napiersky, U., & Frey, D. (2004). “It’s not my fault—But only I can change it”: Counterfactual and prefactual thoughts of managers. *Journal of Applied Psychology*, 89, 279–292. <http://dx.doi.org/10.1037/0021-9010.89.2.279>
- Goh, F. A., & Anderson, M. C. (2007). Driving business value from performance management at Caterpillar. *Organization Development Journal*, 25, 219–226.
- Gomez-Mejia, L. R. (1988). The role of human resources strategy in export performance: A longitudinal study. *Strategic Management Journal*, 9, 493–505. Retrieved from <http://www.jstor.org/stable/2485958>
- Gong, Y., Law, K. S., Chang, S., & Xin, K. R. (2009). Human resources management and firm performance: The differential role of managerial affective and continuance commitment. *Journal of Applied Psychology*, 94, 263–275. <http://dx.doi.org/10.1037/a0013116>
- Gorman, C. A., & Rentsch, J. R. (2009). Evaluating frame-of-reference rater training effectiveness using performance schema accuracy. *Journal of Applied Psychology*, 94, 1336–1344. <http://dx.doi.org/10.1037/a0016476>
- Gregory, J. B., & Levy, P. E. (2011). It’s not me, it’s you: A multilevel examination of variables that impact employee coaching relationships. *Consulting Psychology Journal*, 63, 67–88. <http://dx.doi.org/10.1037/a0024152>
- Gregory, J. B., & Levy, P. E. (2012). Employee feedback orientation: Implications for effective coaching relationships. *Coaching*, 5, 86–99. <http://dx.doi.org/10.1080/17521882.2012.691888>
- Greller, M. M., & Jackson, J. H. (1997). A subordinate’s experience and prior feedback as determinants of participation in performance appraisal reviews. *Psychological Reports*, 80, 547–561. <http://dx.doi.org/10.2466/pr0.1997.80.2.547>
- Gruman, J. A., & Saks, A. M. (2011). Performance management and employee engagement. *Human Resource Management Review*, 21, 123–136. <http://dx.doi.org/10.1016/j.hrmr.2010.09.004>
- Guerci, M., Radaelli, G., Siletti, E., Cirella, S., & Rami Shani, A. (2015). The impact of human resource management practices and corporate sustainability on organizational ethical climates: An employee perspective. *Journal of Business Ethics*, 126, 325–342. <http://dx.doi.org/10.1007/s10551-013-1946-1>
- Guest, D. E. (1987). Human resource management and industrial relations. *Journal of Management Studies*, 24, 503–521. <http://dx.doi.org/10.1111/j.1467-6486.1987.tb00460.x>
- Guest, D. E. (1997). Human resource management and performance: A review and research agenda. *International Journal of Human Resource Management*, 8, 263–276. <http://dx.doi.org/10.1080/095851997341630>
- Haines, V. Y., III, & St-Onge, S. (2012). Performance management effectiveness: Practices or context? *International Journal of Human Resource Management*, 23, 1158–1175. <http://dx.doi.org/10.1080/09585192.2011.561230>
- Harackiewicz, J. M., Abrahams, S., & Wageman, R. (1987). Performance evaluation and intrinsic motivation: The effects of evaluative focus, rewards, and achievement orientation. *Journal of Personality and Social Psychology*, 53, 1015–1023. <http://dx.doi.org/10.1037/0022-3514.53.6.1015>
- Harris, C. (1988). A comparison of employee attitudes toward two performance appraisal systems. *Public Personnel Management*, 17, 443–456. <http://dx.doi.org/10.1177/009102608801700408>
- Harter, J. K., Schmidt, F. L., & Hayes, T. L. (2002). Business-unit-level relationship between employee satisfaction, employee engagement, and business outcomes: A meta-analysis. *Journal of Applied Psychology*, 87, 268–279. <http://dx.doi.org/10.1037/0021-9010.87.2.268>

- Hedge, J. W., & Teachout, M. S. (2000). Exploring the concept of acceptability as a criterion for evaluating performance measures. *Group & Organization Management*, 25, 22–44. <http://dx.doi.org/10.1177/1059601100251003>
- Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin*, 121, 43–64. <http://dx.doi.org/10.1037/0033-2909.121.1.43>
- Holbrook, R. L., Jr. (1999). Managing reactions to performance appraisal: The influence of multiple justice mechanisms. *Social Justice Research*, 12, 205–220. <http://dx.doi.org/10.1023/A:1022196301372>
- Ilgen, D. R., Barnes-Farrell, J. L., & McKellin, D. B. (1993). Performance appraisal process research in the 1980s: What has it contributed to appraisals in use? *Organizational Behavior and Human Decision Processes*, 54, 321–368. <http://dx.doi.org/10.1006/obhd.1993.1015>
- Inderrieden, E. J., Allen, R. E., & Keaveny, T. J. (2004). Managerial discretion in the use of self-ratings in an appraisal system: The antecedents and consequences. *Journal of Managerial Issues*, 16, 460–482.
- Irs, R., & Türk, K. (2012). Implementation of the performance-related pay in the general educational schools of Estonia: Perspectives and possibilities. *Employee Relations*, 34, 360–393. <http://dx.doi.org/10.1108/01425451211236823>
- Jacobs, G., Belschak, F. D., & den Hartog, D. N. (2014). (Un)ethical behavior and performance appraisal: The role of affect, support, and organizational justice. *Journal of Business Ethics*, 121, 63–76.
- Jawahar, I. M. (2010). The mediating role of appraisal feedback reactions on the relationship between rater feedback-related behaviors and rater performance. *Group & Organization Management*, 35, 494–526. <http://dx.doi.org/10.1177/1059601110378294>
- Jawahar, M. (2001). Attitudes, self-monitoring, and appraisal behaviors. *Journal of Applied Psychology*, 86, 875–883. <http://dx.doi.org/10.1037/0021-9010.86.5.875>
- Jiang, J., Wang, S., & Zhao, S. (2012). Does HRM facilitate employee creativity and organizational innovation? A study of Chinese firms. *International Journal of Human Resource Management*, 23, 4025–4047. <http://dx.doi.org/10.1080/09585192.2012.690567>
- Jiang, K., Lepak, D., Hu, J., & Baer, J. (2012). How does human resource management influence organizational outcomes? A meta-analytic investigation of mediating mechanisms. *Academy of Management Journal*, 55, 1264–1294. <http://dx.doi.org/10.5465/amj.2011.0088>
- Jiang, K., Takeuchi, R., & Lepak, D. P. (2013). Where do we go from here? New perspectives on the black box in strategic human resource management research. *Journal of Management Studies*, 50, 1448–1480.
- Johnson, M., & Helgeson, V. S. (2002). Sex differences in response to evaluative feedback: A field study. *Psychology of Women Quarterly*, 26, 242–251. <http://dx.doi.org/10.1111/1471-6402.00063>
- Juhdi, N., Pa'wan, F., & Hansaram, R. K. (2013). HR practices and turnover intention: The mediating roles of organizational commitment and organizational engagement in a selected region in Malaysia. *International Journal of Human Resource Management*, 24, 3002–3019. <http://dx.doi.org/10.1080/09585192.2013.763841>
- Kacmar, K. M., Wayne, S. J., & Wright, P. M. (1996). Subordinate reactions to the use of impression management tactics and feedback by the supervisor. *Journal of Managerial Issues*, 8, 35–53.
- Kamer, B., & Annen, H. (2010). The role of core self-evaluations in predicting performance appraisal reactions. *Swiss Journal of Psychology*, 69, 95–104. <http://dx.doi.org/10.1024/1421-0185/a000011>
- Kaya, N., Koc, E., & Topcu, D. (2010). An exploratory analysis of the influence of human resource management activities and organizational climate on job satisfaction in Turkish banks. *International Journal of Human Resource Management*, 21, 2031–2051. <http://dx.doi.org/10.1080/09585192.2010.505104>
- Keaveny, T. J., Inderrieden, E. J., & Allen, R. E. (1987). An integrated perspective of performance appraisal interviews. *Psychological Reports*, 61, 639–646. <http://dx.doi.org/10.2466/pr0.1987.61.2.639>
- Keeping, L. M., & Levy, P. E. (2000). Performance appraisal reactions: Measurement, modeling, and method bias. *Journal of Applied Psychology*, 85, 708–723. <http://dx.doi.org/10.1037/0021-9010.85.5.708>
- Kim, K. Y., Atwater, L., Patel, P. C., & Smither, J. W. (2016). Multisource feedback, human capital, and the financial performance of organizations. *Journal of Applied Psychology*, 101, 1569–1584. <http://dx.doi.org/10.1037/apl0000125>
- Kinicki, A. J., Jacobson, K. L., Peterson, S. J., & Prussia, G. E. (2013). Development and validation of the performance management behavior questionnaire. *Personnel Psychology*, 66, 1–45. <http://dx.doi.org/10.1111/peps.12013>
- Kinicki, A. J., Prussia, G. E., Wu, B. J., & McKee-Ryan, F. M. (2004). A covariance structure analysis of employees' response to performance feedback. *Journal of Applied Psychology*, 89, 1057–1069. <http://dx.doi.org/10.1037/0021-9010.89.6.1057>
- Kirkpatrick, D. L. (1976). Evaluation of training. In R. L. Craig (Ed.), *Training and development handbook: A guide to human resource development* (2nd ed., pp. 18–27). New York, NY: McGraw-Hill.
- Kirkpatrick, D. L. (1987). Evaluation of training. In R. L. Craig (Ed.), *Training and development handbook: A guide to human resource development* (3rd ed., pp. 301–319). New York, NY: McGraw-Hill.
- Klein, H. J., & Snell, S. A. (1994). The impact of interview process and context on performance appraisal interview effectiveness. *Journal of Managerial Issues*, 6, 160–175.
- Kluger, A. N., & Nir, D. (2010). The feedforward interview. *Human Resource Management Review*, 20, 135–246. <http://dx.doi.org/10.1016/j.hrmr.2009.08.002>
- Korsgaard, M. A. (1996). The impact of self-appraisals on reactions to feedback from others: The role of self-enhancement and self-consistency concerns. *Journal of Organizational Behavior*, 17, 301–311. [http://dx.doi.org/10.1002/\(SICI\)1099-1379\(199607\)17:4<301::AID-JOB749>3.0.CO;2-N](http://dx.doi.org/10.1002/(SICI)1099-1379(199607)17:4<301::AID-JOB749>3.0.CO;2-N)
- Korsgaard, M. A., & Roberson, L. (1995). Procedural justice in performance evaluation: The role of instrumental and non-instrumental voice in performance appraisal discussions. *Journal of Management*, 21, 657–669. <http://dx.doi.org/10.1177/014920639502100404>
- Korsgaard, M. A., Roberson, L., & Rymph, R. D. (1998). What motivates fairness? The role of subordinate assertive behavior on manager's interactional fairness. *Journal of Applied Psychology*, 83, 731–744. <http://dx.doi.org/10.1037/0021-9010.83.5.731>
- Kozlowski, S. W., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, 7, 77–124. <http://dx.doi.org/10.1111/j.1529-1006.2006.00030.x>
- Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 3–90). San Francisco, CA: Jossey-Bass.
- Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology*, 78, 311–328. <http://dx.doi.org/10.1037/0021-9010.78.2.311>
- Kuvaas, B. (2006). Performance appraisal satisfaction and employee outcomes: Mediating and moderating roles of work motivation. *International Journal of Human Resource Management*, 17, 504–522. <http://dx.doi.org/10.1080/09585190500521581>
- Kuvaas, B. (2007). Different relationships between perceptions of developmental performance appraisal and work performance. *Personnel Review*, 36, 378–397. <http://dx.doi.org/10.1108/00483480710731338>
- Kuvaas, B. (2011). The interactive role of performance appraisal reactions and regular feedback. *Journal of Managerial Psychology*, 26, 123–137. <http://dx.doi.org/10.1108/02683941111102164>
- Ladyshewsky, R. K. (2010). The manager as coach as a driver of organizational development. *Leadership and Organization Development Journal*, 31, 292–306. <http://dx.doi.org/10.1108/01437731011043320>

- Laitinen, H., & Ruohomäki, L. (1996). The effects of feedback and goal setting on safety performance at two construction sites. *Safety Science*, 24, 61–73. [http://dx.doi.org/10.1016/S0925-7535\(96\)00070-7](http://dx.doi.org/10.1016/S0925-7535(96)00070-7)
- Lakshman, C. (2014). Leveraging human capital through performance management process: The role of leadership in the USA, France and India. *International Journal of Human Resource Management*, 25, 1351–1372. <http://dx.doi.org/10.1080/09585192.2013.870310>
- Lam, S. S. K., & Schaubroeck, J. (1999). Total quality management and performance appraisal: An experimental study of process versus results and group versus individual approaches. *Journal of Organizational Behavior*, 20, 445–457. [http://dx.doi.org/10.1002/\(SICI\)1099-1379\(199907\)20:4<445::AID-JOB938>3.0.CO;2-J](http://dx.doi.org/10.1002/(SICI)1099-1379(199907)20:4<445::AID-JOB938>3.0.CO;2-J)
- Lam, S. S. K., Schaubroeck, J., & Aryee, S. (2002). Relationship between organizational justice and employee work outcomes: A cross-national study. *Journal of Organizational Behavior*, 23, 1–18. <http://dx.doi.org/10.1002/job.131>
- Lam, S. S. K., Yik, M. S. M., & Schaubroeck, J. (2002). Responses to formal performance appraisal feedback: The role of negative affectivity. *Journal of Applied Psychology*, 87, 192–201. <http://dx.doi.org/10.1037/0021-9010.87.1.192>
- Langan-Fox, J., Bell, R., McDonald, L., & Morizzi, M. (1996). The dimensionality of ratings of performance appraisal systems. *Australian Psychologist*, 31, 194–203. <http://dx.doi.org/10.1080/00050069608260206>
- Langan-Fox, J., Waycott, J., Morizzi, M., & McDonald, L. (1998). Predictors of participation in performance appraisal: A voluntary system in a blue-collar work environment. *International Journal of Selection and Assessment*, 6, 249–260. <http://dx.doi.org/10.1111/1468-2389.00096>
- Latham, G. P., Budworth, M.-H., Yanar, B., & Whyte, G. (2008). The influence of a manager's own performance appraisal on the evaluation of others. *International Journal of Selection and Assessment*, 16, 220–228. <http://dx.doi.org/10.1111/j.1468-2389.2008.00428.x>
- Lawler, E. E., III. (2003). Reward practices and performance management system effectiveness. *Organizational Dynamics*, 32, 396–404. <http://dx.doi.org/10.1016/j.orgdyn.2003.08.007>
- Lawler, E. E., III, & McDermott, M. (2003). Current performance management practices-Examining the varying impacts. *World at Work Journal*, 12, 49–60.
- Lee, F. H., Lee, T. Z., & Wu, W. Y. (2010). The relationship between human resource management practices, business strategy and firm performance: Evidence from steel industry in Taiwan. *International Journal of Human Resource Management*, 21, 1351–1372. <http://dx.doi.org/10.1080/09585192.2010.488428>
- Lepak, D. P., Liao, H., Chung, Y., & Harden, E. E. (2006). A conceptual review of human resource management systems in strategic human resource management research. *Research in Personnel and Human Resources Management*, 25, 217–271.
- Levy, P. E., Tseng, S. T., Rosen, C. C., & Lueke, S. B. (2017). Performance management: A marriage between practice and science—just say “I do”. *Research in Personnel and Human Resources Management*, 35, 155–213. <http://dx.doi.org/10.1108/S0742-730120170000035005>
- Levy, P. E., & Williams, J. R. (2004). The social context of performance appraisal: A review and framework for the future. *Journal of Management*, 30, 881–905. <http://dx.doi.org/10.1016/j.jm.2004.06.005>
- Li, Y., Wang, M., van Jaarsveld, D., Lee, G. K., & Ma, D. (2018). From employee-experienced high-involvement work system to innovation: An emergence-based human resource management framework. *Academy of Management Journal*, 61, 2000–2019. <http://dx.doi.org/10.5465/amj.2015.1101>
- Linderbaum, B. A., & Levy, P. E. (2010). The development and validation of the feedback orientation scale (FOS). *Journal of Management*, 36, 1372–1405. <http://dx.doi.org/10.1177/0149206310373145>
- Linna, A., Elovainio, M., Van den Bos, K., Kivimäki, M., Pentti, J., & Vahtera, J. (2012). Can usefulness of performance appraisal interviews change organizational justice perceptions? A 4-year longitudinal study among public sector employees. *International Journal of Human Resource Management*, 23, 1360–1375. <http://dx.doi.org/10.1080/09585192.2011.579915>
- Liu, N. C., & Liu, M. S. (2011). Human resource practices and individual knowledge-sharing behavior - An empirical study for Taiwanese R&D professionals. *International Journal of Human Resource Management*, 22, 981–997. <http://dx.doi.org/10.1080/09585192.2011.555138>
- Longenecker, C. O., Liverpool, P. R., & Wilson, K. Y. (1988). An assessment of manager/subordinate perceptions of performance appraisal effectiveness. *Journal of Business and Psychology*, 2, 311–320. <http://dx.doi.org/10.1007/BF01013763>
- Maas, V., & Torres-González, R. (2011). Subjective performance evaluation and gender discrimination. *Journal of Business Ethics*, 101, 667–681. <http://dx.doi.org/10.1007/s10551-011-0763-7>
- Mabey, C. (2001). Closing the circle: Participant views of a 360 degree feedback programme. *Human Resource Management Journal*, 11, 41–53. <http://dx.doi.org/10.1111/j.1748-8583.2001.tb00031.x>
- Macan, T., Mehner, K., Havill, L., Meriac, J. P., Roberts, L., & Heft, L. (2011). Two for the price of one: Assessment center training to focus on behaviors can transfer to performance appraisals. *Human Performance*, 24, 443–457. <http://dx.doi.org/10.1080/08959285.2011.614664>
- Mahoney, J. T., & Kor, Y. Y. (2015). Advancing the human capital perspective on value creation by joining capabilities and governance approaches. *The Academy of Management Perspectives*, 29, 296–308. <http://dx.doi.org/10.5465/amp.2014.0151>
- Mamatoglu, N. (2008). Effects on organizational context (culture and climate) from implementing a 360-degree feedback system: The case of Arcelik. *European Journal of Work and Organizational Psychology*, 17, 426–449. <http://dx.doi.org/10.1080/13594320802281094>
- Manshor, A. T., & Kamalanabhan, T. J. (2000). An examination of raters' and ratees' preferences in process and feedback in performance appraisal. *Psychological Reports*, 86, 203–214. <http://dx.doi.org/10.2466/pr0.2000.86.1.203>
- Martell, R. F. (1991). Sex bias at work: The effects of attentional and memory demands on performance ratings of men and women. *Journal of Applied Social Psychology*, 21, 1939–1960. <http://dx.doi.org/10.1111/j.1559-1816.1991.tb00515.x>
- Martell, R. F., & Leavitt, K. N. (2002). Reducing the performance-cue bias in work behavior ratings: Can groups help? *Journal of Applied Psychology*, 87, 1032–1041. <http://dx.doi.org/10.1037/0021-9010.87.6.1032>
- Masterson, S. S., Lewis, K., Goldman, B. M., & Taylor, M. S. (2000). Integrating justice and social exchange: The differing effects of fair procedures and treatment on work relationships. *Academy of Management Journal*, 43, 738–748.
- Matta, F. K., Scott, B. A., Koopman, J., & Conlon, D. E. (2015). Does seeing “eye to eye” affect work engagement and organizational citizenship behavior? A role theory perspective on LMX agreement. *Academy of Management Journal*, 58, 1686–1708. <http://dx.doi.org/10.5465/amj.2014.0106>
- Mayer, R. C., & Davis, J. H. (1999). The effect of the performance appraisal system on trust for management: A field quasi-experiment. *Journal of Applied Psychology*, 84, 123–136. <http://dx.doi.org/10.1037/0021-9010.84.1.123>
- McBriarty, M. A. (1988). Performance appraisal: Some unintended consequences. *Public Personnel Management*, 17, 421–434. <http://dx.doi.org/10.1177/009102608801700406>
- Mero, N. P., Guidice, R. M., & Brownlee, A. L. (2007). Accountability in a performance appraisal context: The effect of audience and form of accounting on rater response and behavior. *Journal of Management*, 33, 223–252. <http://dx.doi.org/10.1177/0149206306297633>
- Mero, N. P., Motowidlo, S. J., & Anna, A. L. (2003). Effects of accountability on rating behavior and rater accuracy. *Journal of Applied Social*

- Psychology*, 33, 2493–2514. <http://dx.doi.org/10.1111/j.1559-1816.2003.tb02777.x>
- Messersmith, J. G., Patel, P. C., Lepak, D. P., & Gould-Williams, J. (2011). Unlocking the black box: Exploring the link between high-performance work systems and performance. *Journal of Applied Psychology*, 96, 1105–1118. <http://dx.doi.org/10.1037/a0024710>
- Milanowski, A. T. (2005). Split roles in performance evaluation: A field study involving new teachers. *Journal of Personnel Evaluation in Education*, 18, 153–169. <http://dx.doi.org/10.1007/s11092-006-9017-0>
- Mintzberg, H. (1971). Managerial work: Analysis from observation. *Management Science*, 18, B-97–B-110. <http://dx.doi.org/10.1287/mnsc.18.2.B97>
- Morgan, A., Cannan, K., & Cullinane, J. (2005). 360° feedback: A critical enquiry. *Personnel Review*, 34, 663–680. <http://dx.doi.org/10.1108/00483480510623457>
- Moss, S. E., Valenzi, E. R., & Taggart, W. (2003). Are you hiding from your boss?: The development of a taxonomy and instrument to assess the feedback management behaviors of good and bad performers. *Journal of Management*, 29, 487–510. [http://dx.doi.org/10.1016/S0149-2063\(03\)00022-9](http://dx.doi.org/10.1016/S0149-2063(03)00022-9)
- Mulligan, J. R., & Bull Schaefer, R. A. (2011). A new hope for rank and yank. *Journal of Leadership & Organizational Studies*, 18, 385–396. <http://dx.doi.org/10.1177/1548051811405206>
- Mullin, R. F., & Sherman, R. (1993). Creativity and performance appraisal: Shall never the twain meet? *Creativity Research Journal*, 6, 425–434. <http://dx.doi.org/10.1080/10400419309534497>
- Murphy, K. R., & Cleveland, J. (1995). *Understanding performance appraisal: Social, organizational, and goal-based perspectives*. Thousand Oaks, CA: Sage.
- Narcisse, S., & Harcourt, M. (2008). Employee fairness perceptions of performance appraisal: A Saint Lucian case study. *International Journal of Human Resource Management*, 19, 1152–1169. <http://dx.doi.org/10.1080/09585190802051451>
- Nathan, B. R., Mohrman, A. M., Jr., & Milliman, J. (1991). Interpersonal relations as a context for the effects of appraisal interview on performance and satisfaction: A longitudinal study. *Academy of Management Journal*, 34, 352–369.
- Norris-Watts, C., & Levy, P. E. (2004). The mediating role of affective commitment in the relation of the feedback environment to work outcomes. *Journal of Vocational Behavior*, 65, 351–365. <http://dx.doi.org/10.1016/j.jvb.2003.08.003>
- Northcraft, G. B., Schmidt, A. M., & Ashford, S. J. (2011). Feedback and the rationing of time and effort among competing tasks. *Journal of Applied Psychology*, 96, 1076–1086. <http://dx.doi.org/10.1037/a0023221>
- Nurse, L. (2005). Performance appraisal, employee development and organizational justice: Exploring the linkages. *International Journal of Human Resource Management*, 16, 1176–1194. <http://dx.doi.org/10.1080/09585190500144012>
- Ostroff, C., & Bowen, D. E. (2000). Moving HR to a higher level: Human resource practices and organizational effectiveness. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 211–266). San Francisco CA: Jossey-Bass.
- Pampino, R. N., Jr., MacDonald, J. E., Mullin, J. B., & Wilder, D. A. (2003). Weekly feedback vs. daily feedback: An application in retail. *Journal of Organizational Behavior Management*, 23, 21–43. http://dx.doi.org/10.1300/J075v23n02_03
- Payne, S. C., Horner, M. T., Boswell, W. R., Schroeder, A. N., & Stine-Cheyne, K. J. (2009). Comparison of online and traditional performance appraisal systems. *Journal of Managerial Psychology*, 24, 526–544. <http://dx.doi.org/10.1108/02683940910974116>
- Payne, S. C., & Mendoza, S. (2017, April). *Performance ratings: Does case law reveal more harm than good?* Paper presented at the Society for Industrial and Organizational Psychology Conference, Orlando, FL.
- Pearce, J. L., & Porter, L. W. (1986). Employee responses to formal performance appraisal feedback. *Journal of Applied Psychology*, 71, 211–218. <http://dx.doi.org/10.1037/0021-9010.71.2.211>
- Peretz, H., & Fried, Y. (2012). National cultures, performance appraisal practices, and organizational absenteeism and turnover: A study across 21 countries. *Journal of Applied Psychology*, 97, 448–459. <http://dx.doi.org/10.1037/a0026011>
- Peterson, S. L. (2004). Toward a theoretical model of employee turnover: A human resource development perspective. *Human Resource Development Review*, 3, 209–227. <http://dx.doi.org/10.1177/1534484304267832>
- Pettijohn, C. E., Pettijohn, L. S., & d'Amico, M. (2001). Characteristics of performance appraisals and their impact on sales force satisfaction. *Human Resource Development Quarterly*, 12, 127–146. <http://dx.doi.org/10.1002/hrd.4>
- Pfeffer, J. (1998). *The human equation: Building profits by putting people first*. Boston, MA: Harvard Business Press.
- Pichler, S. (2012). The social context of performance appraisal and appraisal reactions: A meta-analysis. *Human Resource Management*, 51, 709–732. <http://dx.doi.org/10.1002/hrm.21499>
- Ployhart, R. E., & Moliterno, T. P. (2011). Emergence of the human capital resource: A multilevel model. *Academy of Management Review*, 36, 127–150. <http://dx.doi.org/10.5465/amr.2009.0318>
- Ployhart, R. E., Nyberg, A. J., Reilly, G., & Maltarich, M. A. (2014). Human capital is dead; Long live human capital resources! *Journal of Management*, 40, 371–398. <http://dx.doi.org/10.1177/0149206313512152>
- Podsakoff, P. M., & Farh, J. L. (1989). Effects of feedback sign and credibility on goal setting and task performance. *Organizational Behavior and Human Decision Processes*, 44, 45–67. [http://dx.doi.org/10.1016/0749-5978\(89\)90034-4](http://dx.doi.org/10.1016/0749-5978(89)90034-4)
- Pollack, M. J., Fleming, R. K., & Sulzer-Azaroff, B. (1994). Enhancing professional performance through organizational change. *Behavioral Interventions*, 9, 27–42. <http://dx.doi.org/10.1002/bin.2360090104>
- Poon, J. M. L. (2004). Effects of performance appraisal politics on job satisfaction and turnover intention. *Personnel Review*, 33, 322–334. <http://dx.doi.org/10.1108/00483480410528850>
- Prince, J. B., & Lawler, E. E., III. (1986). Does salary discussion hurt the developmental performance appraisal? *Organizational Behavior and Human Decision Processes*, 37, 357–375. [http://dx.doi.org/10.1016/0749-5978\(86\)90035-X](http://dx.doi.org/10.1016/0749-5978(86)90035-X)
- Pulakos, E. D., & O'Leary, R. S. (2011). Why is performance management broken? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 4, 146–164. <http://dx.doi.org/10.1111/j.1754-9434.2011.01315.x>
- Raeder, S., Knorr, U., & Hilb, M. (2012). Human resource management practices and psychological contracts in Swiss firms: An employer perspective. *International Journal of Human Resource Management*, 23, 3178–3195. <http://dx.doi.org/10.1080/09585192.2011.637066>
- Rahman, S. A. (2006). Attitudes of Malaysian teachers toward a performance-appraisal system. *Journal of Applied Social Psychology*, 36, 3031–3042. <http://dx.doi.org/10.1111/j.0021-9029.2006.00141.x>
- Randall, R., & Sharples, D. (2012). The impact of rater agreeableness and rating context on the evaluation of poor performance. *Journal of Occupational and Organizational Psychology*, 85, 42–59. <http://dx.doi.org/10.1348/2044-8325.002002>
- Rao, A. S. (2007). Effectiveness of performance management systems: An empirical study in Indian companies. *International Journal of Human Resource Management*, 18, 1812–1840. <http://dx.doi.org/10.1080/09585190701570973>
- Reber, R. A., & Wallin, J. A. (1994). Utilizing performance management to improve offshore oilfield diving safety. *The International Journal of Organizational Analysis*, 2, 88–98. <http://dx.doi.org/10.1108/eb028803>
- Redman, T., & Mathews, B. P. (1995). Do corporate turkeys vote for Christmas? Managers' attitudes towards upward appraisal. *Personnel Review*, 24, 13–24. <http://dx.doi.org/10.1108/00483489510095761>

- Rentsch, J. R. (1990). Climate and culture: Interaction and qualitative differences in organizational meanings. *Journal of Applied Psychology*, 75, 668–681. <http://dx.doi.org/10.1037/0021-9010.75.6.668>
- Roberts, G. E. (1995). Municipal government performance appraisal system practices: Is the whole less than the sum of its parts? *Public Personnel Management*, 24, 197–221. <http://dx.doi.org/10.1177/009102609502400209>
- Rodwell, J. J., & Teo, S. T. T. (2008). The influence of strategic HRM and sector on perceived performance in health services organizations. *International Journal of Human Resource Management*, 19, 1825–1841. <http://dx.doi.org/10.1080/09585190802323934>
- Rowland, C. (2013). Managing team performance: Saying and paying. *The International Journal of Organizational Analysis*, 21, 38–52. <http://dx.doi.org/10.1108/19348831311322524>
- Russell, J. S., & Goode, D. L. (1988). An analysis of managers' reactions to their own performance appraisal feedback. *Journal of Applied Psychology*, 73, 63–67. <http://dx.doi.org/10.1037/0021-9010.73.1.63>
- Saffie-Robertson, M. C., & Brutus, S. (2014). The impact of interdependence on performance evaluations: The mediating role of discomfort with performance appraisal. *International Journal of Human Resource Management*, 25, 459–473. <http://dx.doi.org/10.1080/09585192.2013.792864>
- Schleicher, D. J., Baumann, H. M., Sullivan, D. W., Levy, P. E., Hargrove, D. C., & Barros-Rivera, B. A. (2018). Putting the system into performance management systems: A review and agenda for performance management research. *Journal of Management*, 44, 2209–2245. <http://dx.doi.org/10.1177/0149206318755303>
- Schleicher, D. J., Bull, R. A., & Green, S. G. (2009). Rater reactions to forced distribution rating systems. *Journal of Management*, 35, 899–927. <http://dx.doi.org/10.1177/0149206307312514>
- Schleicher, D. J., & Day, D. V. (1998). A cognitive evaluation of frame-of-reference rater training: Content and process issues. *Organizational Behavior and Human Decision Processes*, 73, 76–101. <http://dx.doi.org/10.1006/obhd.1998.2751>
- Schleicher, D. J., Smith, T. A., Casper, W. J., Watt, J. D., & Greguras, G. J. (2015). It's all in the attitude: The role of job attitude strength in job attitude-outcome relationships. *Journal of Applied Psychology*, 100, 1259–1274. <http://dx.doi.org/10.1037/a0038664>
- Schleicher, D. J., Watt, J. D., & Greguras, G. J. (2004). Reexamining the job satisfaction-performance relationship: The complexity of attitudes. *Journal of Applied Psychology*, 89, 165–177. <http://dx.doi.org/10.1037/0021-9010.89.1.165>
- Scullen, S. E., Bergey, P. K., & Aiman-Smith, L. (2005). Forced distribution rating systems and the improvement of workforce potential: A baseline simulation. *Personnel Psychology*, 58, 1–31. <http://dx.doi.org/10.1111/j.1744-6570.2005.00361.x>
- Searle, R. H., & Ball, K. S. (2003). Supporting innovation through HR policy: Evidence from the U. K. *Creativity and Innovation Management*, 12, 50–62. <http://dx.doi.org/10.1111/1467-8691.00268>
- Seiden, S., & Sowa, J. E. (2011). Performance management and appraisal in human service organizations: Management and staff perspectives. *Public Personnel Management*, 40, 251–264. <http://dx.doi.org/10.1177/009102601104000305>
- Selden, S., Sherrier, T., & Wooters, R. (2012). Experimental study comparing a traditional approach to performance appraisal training to a whole-brain training method at C. B. Fleet Laboratories. *Human Resource Development Quarterly*, 23, 9–34. <http://dx.doi.org/10.1002/hrdq.21123>
- Selvarajan, T. T., & Cloninger, P. A. (2012). Can performance appraisals motivate employees to improve performance? A Mexican study. *International Journal of Human Resource Management*, 23, 3063–3084. <http://dx.doi.org/10.1080/09585192.2011.637069>
- Shen, J., D'Netto, B., & Tang, J. (2010). Effects of human resource diversity management on organizational citizen behaviour in the Chinese context. *International Journal of Human Resource Management*, 21, 2156–2172. <http://dx.doi.org/10.1080/09585192.2010.509622>
- Si, S., & Li, Y. (2012). Human resource management practices on exit, voice, loyalty, and neglect: Organizational commitment as a mediator. *International Journal of Human Resource Management*, 23, 1705–1716. <http://dx.doi.org/10.1080/09585192.2011.580099>
- Smith, W. J., Harrington, K. V., & Houghton, J. D. (2000). Predictors of performance appraisal discomfort: A preliminary examination. *Public Personnel Management*, 29, 21–32. <http://dx.doi.org/10.1177/009102600002900102>
- Song, C., Sommer, S. M., & Hartman, A. E. (1998). The impact of adding an external rater on interdepartmental cooperative behaviors of workers. *International Journal of Conflict Management*, 9, 117–138. <http://dx.doi.org/10.1108/eb022806>
- Spence, D. G., & Wood, E. E. (2007). Registered nurse participation in performance appraisal interviews. *Journal of Professional Nursing*, 23, 55–59. <http://dx.doi.org/10.1016/j.profnurs.2005.11.003>
- Spence, J. R., & Keeping, L. (2011). Conscious rating distortion in performance appraisal: A review, commentary, and proposed framework for research. *Human Resource Management Review*, 21, 85–95. <http://dx.doi.org/10.1016/j.hrmr.2010.09.013>
- Staw, B. M., & Ross, J. (1985). Stability in the midst of change: A dispositional approach to job attitudes. *Journal of Applied Psychology*, 70, 469–480. <http://dx.doi.org/10.1037/0021-9010.70.3.469>
- Stumpf, S. A., Doh, J. P., & Tymon, W. G., Jr. (2010). The strength of HR practices in India and their effects on employee career success, performance, and potential. *Human Resource Management*, 49, 353–375. <http://dx.doi.org/10.1002/hrm.20361>
- Sundgren, M., Selart, M., Ingelgård, A., & Bengtson, C. (2005). Dialogue-based evaluation as a creative climate indicator: Evidence from the pharmaceutical industry. *Creativity and Innovation Management*, 14, 84–98. <http://dx.doi.org/10.1111/j.1476-8691.2005.00328.x>
- Szell, S., & Henderson, R. (1997). The impact of self-supervisor/subordinate performance rating agreement on subordinates' job satisfaction and organizational commitment. *Journal of Applied Social Behaviour*, 3, 25–37.
- Tan, C. L., & Nasurdin, A. M. (2011). Human resource management practices and organizational innovation: Assessing the mediating role of knowledge management effectiveness. *Electronic Journal of Knowledge Management*, 9, 155–167.
- Tang, T. L. P., Baldwin, L. J., & Frost, A. G. (1997). Locus of control as a moderator of the self-reported performance feedback-personal sacrifice relationship. *Personality and Individual Differences*, 22, 201–211. [http://dx.doi.org/10.1016/S0191-8869\(96\)00195-X](http://dx.doi.org/10.1016/S0191-8869(96)00195-X)
- Taylor, A. J., Pettijohn, L. S., & Pettijohn, C. E. (1999). Salespersons and sales managers: A descriptive study of topics and perceptions of retail sales performance appraisals. *Human Resource Development Quarterly*, 10, 271–291. <http://dx.doi.org/10.1002/hrdq.3920100306>
- Taylor, M. S., Masterson, S. S., Renard, M. K., & Tracy, K. B. (1998). Managers' reactions to procedurally just performance management systems. *Academy of Management Journal*, 41, 568–579.
- Taylor, M. S., Tracy, K. B., Renard, M., Harrison, J. K., & Carroll, S. J. (1995). Due process in performance appraisal: A quasi-experiment in procedural justice. *Administrative Science Quarterly*, 40, 495–523. <http://dx.doi.org/10.2307/2393795>
- Taylor, P. J., & Pierce, J. L. (1999). Effects of introducing a performance management system on employees' subsequent attitudes and effort. *Public Personnel Management*, 28, 423–452. <http://dx.doi.org/10.1177/009102609902800308>
- Tharenou, P. (1995). The impact of a developmental performance appraisal program on employee perceptions in an Australian federal agency. *Group & Organization Management*, 20, 245–271. <http://dx.doi.org/10.1177/1059601195203002>

- Thorsteinson, T. J., Breier, J., Atwell, A., Hamilton, C., & Privette, M. (2008). Anchoring effects on performance judgments. *Organizational Behavior and Human Decision Processes*, 107, 29–40. <http://dx.doi.org/10.1016/j.obhdp.2008.01.003>
- Thurston, P. W., Jr., & McNall, L. (2010). Justice perceptions of performance appraisal practices. *Journal of Managerial Psychology*, 25, 201–228. <http://dx.doi.org/10.1108/02683941011023712>
- Tjosvold, D., & Halco, J. A. (1992). Performance appraisal of managers: Goal interdependence, ratings, and outcomes. *The Journal of Social Psychology*, 132, 629–639. <http://dx.doi.org/10.1080/00224545.1992.9713902>
- Toegel, G., & Conger, J. A. (2003). 360-degree assessment: Time for reinvention. *Academy of Management Learning & Education*, 2, 297–311. <http://dx.doi.org/10.5465/amle.2003.10932156>
- Tuytens, M., & Devos, G. (2012). The effect of procedural justice in the relationship between charismatic leadership and feedback reactions in performance appraisal. *International Journal of Human Resource Management*, 23, 3047–3062. <http://dx.doi.org/10.1080/09585192.2011.639535>
- Tymon, W. G., Jr., Stumpf, S. A., & Doh, J. P. (2010). Exploring talent management in India: The neglected role of intrinsic rewards. *Journal of World Business*, 45, 109–121. <http://dx.doi.org/10.1016/j.jwb.2009.09.016>
- Tziner, A. (1999). The relationship between distal and proximal factors and the use of political considerations in performance appraisal. *Journal of Business and Psychology*, 14, 217–231. <http://dx.doi.org/10.1023/A:1022931106379>
- Tziner, A., Fein, E. C., Sharoni, G., Bar-Hen, P., & Nord, T. (2010). Constructive deviance, leader-member exchange, and confidence in appraisal: How do they interrelate, if at all? *Revista de Psicología del Trabajo y de las Organizaciones*, 26, 95–100. <http://dx.doi.org/10.5093/tr2010v26n2a1>
- Tziner, A., & Kopelman, R. E. (1988). Effects of rating format on goal-setting dimensions: A field experiment. *Journal of Applied Psychology*, 73, 323–326. <http://dx.doi.org/10.1037/0021-9010.73.2.323>
- Tziner, A., & Kopelman, R. E. (2002). Is there a preferred performance rating format?: A non-psychometric perspective. *Applied Psychology*, 51, 479–503. <http://dx.doi.org/10.1111/1464-0597.00104>
- Tziner, A., Latham, G. P., Price, B. S., & Haccoun, R. (1996). Development and validation of a questionnaire for measuring perceived political considerations in performance appraisal. *Journal of Organizational Behavior*, 17, 179–190. [http://dx.doi.org/10.1002/\(SICI\)1099-1379\(199603\)17:2<179::AID-JOB740>3.0.CO;2-Z](http://dx.doi.org/10.1002/(SICI)1099-1379(199603)17:2<179::AID-JOB740>3.0.CO;2-Z)
- Tziner, A., Murphy, K. R., Cleveland, J. N., & Roberts-Thompson, G. P. (2001). Relationships between attitudes toward organizations and performance appraisal systems and rating behavior. *International Journal of Selection and Assessment*, 9, 226–239. <http://dx.doi.org/10.1111/1468-2389.00176>
- Tziner, A., Prince, J. B., & Murphy, K. R. (1997). PCPAQ - The questionnaire for measuring perceived political considerations in performance appraisal: Some new evidence regarding its psychometric qualities. *Journal of Social Behavior and Personality*, 12, 189–199.
- Villanova, P., Bernardin, H. J., Dahmus, S. A., & Sims, R. L. (1993). Rater leniency and performance appraisal discomfort. *Educational and Psychological Measurement*, 53, 789–799. <http://dx.doi.org/10.1177/0013164493053003023>
- Waite, M. L., Newman, J. M., & Krzystofiak, F. J. (1994). Associations among performance appraisal, compensation, and total quality programs. *Psychological Reports*, 75, 524–526. <http://dx.doi.org/10.2466/pr0.1994.75.1.524>
- Waldman, D. A., & Atwater, L. E. (2001). Attitudinal and behavioral outcomes of an upward feedback process. *Group & Organization Management*, 26, 189–205. <http://dx.doi.org/10.1177/1059601101262004>
- Wang, H. K., Tseng, J. F., Yen, Y. F., & Huang, I. C. (2011). University staff performance evaluation systems, organizational learning, and organizational identification in Taiwan. *Social Behavior and Personality*, 39, 43–54. <http://dx.doi.org/10.2224/sbp.2011.39.1.43>
- Wang, W. (2007). Evaluation of 2+2 alternative teacher performance appraisal program in Shanxi, People's Republic of China. *Teaching and Teacher Education*, 23, 1012–1023. <http://dx.doi.org/10.1016/j.tate.2006.05.005>
- Waung, M., & Jones, D. R. (2005). The effect of feedback packaging on rater reactions. *Journal of Applied Social Psychology*, 35, 1630–1652. <http://dx.doi.org/10.1111/j.1559-1816.2005.tb02188.x>
- Wexley, K. N., & Snell, S. A. (1987). Managerial power: A neglected aspect of the performance appraisal interview. *Journal of Business Research*, 15, 45–54. [http://dx.doi.org/10.1016/0148-2963\(87\)90017-8](http://dx.doi.org/10.1016/0148-2963(87)90017-8)
- Wexley, K. N., & Youtz, M. A. (1985). Rater beliefs about others: Their effects on rating errors and rater accuracy. *Journal of Occupational and Organizational Psychology*, 58, 265–275. <http://dx.doi.org/10.1111/j.2044-8325.1985.tb00200.x>
- Whiting, H. J., & Kline, T. J. B. (2007). Testing a model of performance appraisal fit on attitudinal outcomes. *The Psychologist Manager Journal*, 10, 127–148. <http://dx.doi.org/10.1080/10887150701451288>
- Whiting, H. J., Kline, T. J. B., & Sulsky, L. M. (2008). The performance appraisal congruency scale: An assessment of person-environment fit. *International Journal of Productivity and Performance Management*, 57, 223–236. <http://dx.doi.org/10.1108/17410400810857239>
- Williams, J. R., & Levy, P. E. (2000). Investigating some neglected criteria: The influence of organizational level and perceived system knowledge on appraisal reactions. *Journal of Business and Psychology*, 14, 501–513. <http://dx.doi.org/10.1023/A:1022988402319>
- Williams, M. J. (1997). Performance appraisal is dead. Long live performance management. *Harvard Management Update*, 2, 1–6.
- Wood, R. E., & Marshall, V. (2008). Accuracy and effectiveness in appraisal outcomes: The influence of self-efficacy, personal factors and organisational variables. *Human Resource Management Journal*, 18, 295–313. <http://dx.doi.org/10.1111/j.1748-8583.2008.00067.x>
- Yang, H., & Klaas, B. S. (2011). Pay dispersion and the financial performance of the firm: Evidence from Korea. *International Journal of Human Resource Management*, 22, 2147–2166. <http://dx.doi.org/10.1080/09585192.2011.580182>
- Youndt, M. A., Snell, S. A., Dean, J. W., & Lepak, D. P. (1996). Human resource management, manufacturing strategy, and firm performance. *Academy of Management Journal*, 39, 836–865.
- Youngcourt, S. S., Leiva, P. I., & Jones, R. G. (2007). Perceived purposes of performance appraisal: Correlates of individual- and position-focused purposes on attitudinal outcomes. *Human Resource Development Quarterly*, 18, 315–343. <http://dx.doi.org/10.1002/hrdq.1207>
- Zheng, C., Morrison, M., & O'Neill, G. (2006). An empirical study of high performance HRM practices in Chinese SMEs. *International Journal of Human Resource Management*, 17, 1772–1803. <http://dx.doi.org/10.1080/09585190600965282>
- Zheng, W., Zhang, M., & Li, H. (2012). Performance appraisal process and organizational citizenship behavior. *Journal of Managerial Psychology*, 27, 732–752. <http://dx.doi.org/10.1108/02683941211259548>
- Zhou, J., & Shalley, C. E. (2003). Research on employee creativity: A critical review and directions for future research. *Research in Personnel and Human Resources Management*, 22, 165–217.

Appendix A

Detailed Summary of Empirical Research on Each Criterion Category

I. PM-Related Reactions

a. Employee reactions. The empirical PM literature has examined 230 variables classified as employee reactions across 166 studies. Cognitive reactions are most prevalent (46%), followed by satisfaction (30%), utility (17%), and affective (7%) employee reactions. Of the 94 studies examining *cognitive reactions*, perceived fairness/justice (especially procedural justice) has been most prevalent (41%; see Clarke, Harcourt, & Flynn, 2013, for a particularly good discussion of justice and PA), followed by perceived accuracy (15%; e.g., Kinicki et al., 2004; Selvarajan & Cloninger, 2012) and acceptance of PM (9%; e.g., Hedge & Teachout, 2000; Morgan et al., 2005). Perceived quality of feedback has also been studied (e.g., Anseel, Lievens, & Schollaert, 2009; Payne, Horner, Boswell, Schroeder, & Stine-Cheyne, 2009). Typically, fairness reactions have been studied with regard to the overall PM system (e.g., Whiting, Kline, & Sulsky, 2008) but have also been examined with regard to more specific aspects of PM, such as ratings (Inderrieden, Allen, & Keaveny, 2004), participation in the PA (Evans & McShane, 1988), and formal processes (Taylor et al., 1998). Research also suggests that interactional injustice in PM is much less common than procedural injustice, which is frequently reported (Narcisse & Harcourt, 2008). More generally, positive cognitive employee reactions and greater satisfaction result when employees participate in the PM process (e.g., Keaveny et al., 1987; Nathan et al., 1991; Prince & Lawler, 1986), when they have knowledge about how the process works, and when they believe their supervisors are unbiased. In fact, as Fulk et al. (1985) have suggested, perceptions of fairness and accuracy in PM may depend as heavily on the level of trust in the supervisor–employee relationship as on characteristics of the PM process itself (see also Dusterhoff, Cunningham, & MacGregor, 2014; Russell & Goode, 1988).

Of the 40 studies examining employee *utility reactions* to PM, the majority focused on evaluating the effectiveness of feedback (Catano, Darr, & Campbell, 2007; Elicker, Levy, & Hall, 2006; Tuytens & Devos, 2012) and overall usefulness of PA (Balfour, 1992; Payne et al., 2009; Seiden & Sowa, 2011). Research in this area shows that employees report greater utility of PM when they participate in the PM process (Keaveny et al., 1987; Prince & Lawler, 1986) and when 360 evaluations are used (Mamatoglu, 2008); there is some evidence that negative feedback is not viewed as useful by employees (Brett & Atwater, 2001). *Affective reactions* of employees have been examined in 26 studies and generally fall into one of two categories: discomfort with the PA (also common among raters), and emotional (positive or negative) responses to aspects of PM. Some research has suggested that negative affective reactions to PM are common across employees, even when perceived importance might be high (Spence & Wood,

2007). The most common finding is that negative feedback is associated with negative affective reactions (emotions such as anger, frustration, discouragement; Atwater & Brett, 2006; Belchak & Den Hartog, 2009; David, 2013; Podsakoff & Farh, 1989). Interestingly, however, positive feedback does not appear to result in positive affective reactions, but rather in an absence of negative affective reactions (Brett & Atwater, 2001). Thus, it remains unclear whether there are any aspects of PM that could actually result in positive affective reactions, or whether the absence of negative reactions is the best that one can hope for.

b. Manager reactions. Our review suggests that the manager reactions literature has not developed as extensively as employee reactions, as there has been much less empirical focus on the former. Specifically, we found 45 variables (across 32 studies) that could be classified as manager reactions (just 16% of all the empirical reactions articles). Similar to employee reactions, the most frequently researched manager reactions have been cognitive reactions (38%, especially fairness/justice), followed by satisfaction reactions (31%). Unique from employee reactions, however, was a sizable number of studies on discomfort with PA (an affective reaction, which accounted in general for 16% of manager reactions studies). Utility reactions comprised 16%. This research indicates that the following aspects are important determinants of managers' PM reactions: type of feedback (e.g., Erdemli, Sümer, & Bilgiç, 2007; Mabey, 2001; Redman & Mathews, 1995), rating approach (FDRS, Schleicher et al., 2009; graphic rating scales, Dale et al., 2013), general comfort with PM processes or the system (Villanova, Bernardin, Dahmus, & Sims, 1993), previous PM experience (Brutus, Fletcher, & Baldry, 2009; Smith et al., 2000), and personality and leadership qualities of the manager (Waldman & Atwater, 2001; Wexley & Youtz, 1985).

II. PM-Related Learning (Employees)

Our review revealed 56 studies examining variables that could be categorized as employee learning, with attitudinal/motivational learning being most frequent (50%), followed by skills-based (29%) and cognitive (21%) learning. Unlike the reactions category, there are several idiosyncratic operationalizations comprising this criterion category, as detailed below.

a. Cognitive learning. Per Kraiger et al. (1993), we delineate cognitive PM-related learning as knowledge (declarative, procedural, and tacit; e.g., awareness of development opportunities, Boswell & Boudreau, 2002), knowledge organization (e.g., task thoughts, Harackiewicz, Abrahams, & Wageman, 1987), and cognitive strategies (e.g., self-awareness, Morgan et al., 2005) resulting from participation in PM processes. We found 13 such variables (in 11 studies) in the PM literature. This research suggests

(Appendices continue)

that cognitive learning outcomes result from the type of evaluation (amount of task-based thoughts increased with a task-focused, as compared with normative-focused, evaluation; Harackiewicz et al., 1987; and value-focused thinking increased with dialogue-based, as compared with control-based, evaluation; Sundgren, Selart, Ingelgård, & Bengtson, 2005). They also result from involvement with and participation in the appraisal meeting (which increased perceived learning and role clarity; Greller & Jackson, 1997; Prince & Lawler, 1986).

b. Skills-based learning. Employees can also learn PM-related skills in response to the PM system, and many of the studies in this category (13 variables across nine studies) focused on skills-based learning surrounding feedback. Outcomes examined in this regard include how employees do their work, which can change as a result of 360 feedback (Morgan et al., 2005); feedback sharing between peers, which increased with a new PM system entailing regular observation and feedback (Wang, 2007), and feedback seeking, which increased as a result of PM in employees higher on feedback orientation (Dahling et al., 2012; Linderbaum & Levy, 2010), self-esteem, and fear of negative evaluation (Moss et al., 2003), and decreased when receiving an evaluation inconsistent with previous feedback (Greller & Jackson, 1997).

c. Attitudinal/motivational learning. The vast majority of the research on employee attitudinal/motivational learning outcomes (28 variables across 26 studies) has examined motivational learning, with only a couple studies focusing on attitudinal learning as an outcome of PM. This research suggests that employees' PM-related motivation (i.e., concern about one's performance level and motivation and effort to improve) can improve with (a) the use of performance-contingent rewards (Harackiewicz et al., 1987); (b) managers who set cooperative (vs. competitive) goals (Tjosvold & Halco, 1992), have power (Fedor, Davis, Maslyn, & Mathieson, 2001; Wexley & Snell, 1987), and have credibility as a feedback source (Kinicki et al., 2004); (c) perceived voice in the PA session (Elicker et al., 2006); and (d) a feedback-rich (Kinicki et al., 2004) and PA-supportive (Langan-Fox, Waycott, Morizzi, & McDonald, 1998) organizational environment. Research has also found greater intentions to use feedback and to engage in related development opportunities when it comes from a credible source (Bannister, 1986) and for employees with a higher feedback orientation (Linderbaum & Levy, 2010). Other research has focused on motivational learning outcomes specifically related to aspects of goals or goal-setting and self-efficacy or other expectancy beliefs. For example, the implementation of PM can increase the number of goals employees plan to achieve (Pollack, Fleming, & Sulzer-Azaroff, 1994); and a behavioral observation scale (BOS) rating format can result in higher levels of goal clarity, goal acceptance, and goal commitment as compared with a Graphic Rating Scale format (Tziner, 1999; Tziner, Prince, & Murphy, 1997) and a behaviorally anchored rating scale (BARS) format (Tziner et al., 1996). Self-efficacy and related expectancy-based

beliefs also can be impacted by rating format, with greater efficacy/expectancy with a process-focused as opposed to results only focused performance evaluation (Lam & Schaubroeck, 1999) and with a five-category as opposed to a three-category rating system. The presence of higher quality feedback also impacts these efficacy/expectancy beliefs (Northcraft, Schmidt, & Ashford, 2011). A final theme in this research is the examination of individual factors as determinants (or especially as moderators) of motivational learning outcomes of PM. Research has found, for example, that higher core self-evaluations were associated with greater goal commitment following the PA discussion (Kamer & Annen, 2010); that high achievers were more concerned with their performance improvement following feedback (Harackiewicz et al., 1987); that women reported greater intentions to change behavior based on evaluation (Johnson & Helgeson, 2002); and that women, under subjective but not objective evaluation, expect more positive evaluation outcomes as the probability of evaluation by a female manager increases (Maas & Torres-González, 2011).

III. PM-Related Learning (Manager)

Of the four types of manager PM-related learning criteria found in our review, rating quality has received the most attention (71%), followed by skills-based (19%), cognitive (5%), and attitudinal/motivational (4%) learning. Similar to employee learning, the manager learning variables are quite idiosyncratic.

a. Cognitive learning. In the vast majority of studies examining manager PM-related cognitive learning, such learning was examined as an outcome of formal PM processes, particularly rater training. Specifically, frame of reference (FOR) training has been found to relate to the holding of less idiosyncratic performance standards (Schleicher & Day, 1998) as well as to greater declarative knowledge and performance schema accuracy (Gorman & Rentsch, 2009). Managers in "whole brain training," a newer form of rater training, showed better understanding of their employees' strengths and weaknesses (as reported by employees, Seiden & Sowa, 2011). Another study found computer training to be effective in increasing managerial knowledge of PA (Davis & Mount, 1984). Outside of training, it has also been found that using groups (as opposed to individuals) to rate can increase rater memory strength and use of neutral decision criteria (Martell & Leavitt, 2002); and managers' participation in a 360-feedback program can result in greater understanding of one's contribution to unit objectives (Mabey, 2001).

b. Skills-based learning. After rating quality, this learning category has received the most attention in empirical PM research. Skills-based learning variables assessed in the literature include effectiveness in completing PA forms (which improved with training, Davis & Mount, 1984); effective note-taking and attention to relevant subordinate performance (which resulted from rater accountability and ultimately improved decision accuracy; Mero, Motowidlo, & Anna, 2003); behavioral specificity in PA evaluation comments (which was higher for raters who had previously

participated in an assessment center; Macan et al., 2011); performance information recall ability (which was improved from a structured diary intervention and ultimately led to higher quality ratings; DeNisi & Peters, 1996); engaging in less intentional rating distortion (after implementation of a new due process PA system; Taylor et al., 1995); and effectiveness of supervisors' appraisal behaviors (as rated by employees, after implementation of a new behaviorally based, pay-for-performance system; Eberhardt & Pooyan, 1988). Finally, in addition to being influenced by a variety of PM system factors, manager skills-based learning demonstrated important relationships with rating quality (e.g., DeNisi & Peters, 1996; Favero & Ilgen, 1989; Mero et al., 2007).

c. Attitudinal/motivational learning. Attitudinal and motivational manager learning variables, like cognitive learning, are often linked to rater training. For example, frame of reference training results in higher levels of agreement with organizational performance theories (Schleicher & Day, 1998), and amount of training relates positively to raters' PA self-efficacy (Wood & Marshall, 2008). Participation in a 360-feedback program is associated with perceived opportunity and satisfaction with PM changes (Mabey, 2001).

d. Rating quality. Our review shows that the emphasis on rating quality as an outcome in PM research, although still substantial, has declined somewhat over time (see Table 2). Rating quality criteria examined include errors (36%), accuracy (35%), bias (14%), validity (9%), and reliability (6%). This research has examined rating quality as an outcome of both formal (e.g., rating scale format, Gomez-Mejia, 1988; training, Wood & Marshall, 2008) and informal (e.g., exposure to anchoring information, Thorsteinson, Breier, Atwell, Hamilton, & Privette, 2008) processes, as well as individual differences (e.g., agreeableness, Randall & Sharples, 2012). Other learning variables also predict rating quality, including cognitive (e.g., less idiosyncratic performance standards, Schleicher & Day, 1998; performance schema, Gorman & Rentsch, 2009) and skills-based (e.g., note-taking, Mero et al., 2007) learning criteria.

IV. Employee Transfer

The empirical PM literature has examined 168 variables classified as employee transfer criteria. The most numerous of these have been job attitudes (34%) and performance (34%), followed by withdrawal (12%), motivation (8%), and fairness/justice (7%).

a. Job attitudes. The most frequently studied job attitude outcomes of PM are job satisfaction and organizational commitment, examined in 26 and 27 studies, respectively. Other job attitudes studied much more infrequently include perceived organizational support (Armstrong-Stassen & Schlosser, 2010; Gavino, Wayne, & Erdogan, 2012; Jacobs, Belschak, & den Hartog, 2014; Masterson et al., 2000), job embeddedness (Bambacas & Kulik, 2013), and role ambiguity (Youngcourt, Leiva, & Jones, 2007). This research suggests that the following aspects of the PM process (formal and informal) can influence employee job satisfac-

tion: the content of the review meeting (e.g., Nathan et al., 1991); the type of criteria used in the PA (e.g., Pettijohn, Pettijohn, & d'Amico, 2001); whether goal setting (e.g., Bipp & Kleingeld, 2011), feedback (e.g., Lam, Yik, & Schaubroeck, 2002), and an explanation for the PA (e.g., Rahman, 2006) are included as part of PM; the existence of political motives in PA (e.g., Poon, 2004); and the extent to which the supervisor and subordinate agree on the ratings (e.g., Szell & Henderson, 1997). Research suggests that *organizational commitment* is similarly influenced by the provision of goal setting (e.g., Taylor & Pierce, 1999); feedback (Lam, Schaubroeck, & Aryee, 2002; Pearce & Porter, 1986; Tang, Baldwin, & Frost, 1997); developmental PA more generally (Youngcourt et al., 2007) and "high-commitment" PM practices (Farndale & Kelliher, 2013); fair treatment by one's supervisor in the PA (Farndale, Hope-Hailey, & Kelliher, 2011); and supervisor/subordinate agreement on ratings (Szell & Henderson, 1997).

The research summarized above would seem to suggest a clear link between elements of PM and employees' job attitudes. Yet much of this research is plagued by common method concerns (single source cross-sectional surveys), and research using stronger designs has often concluded a lack of effect of PM on these more distal job attitudes. For example, using quasi-experimental designs, Eberhardt and Pooyan (1988), Korsgaard, Roberson, and Rymph (1998), and Taylor et al. (1998) all failed to find an effect of PM on job attitudes. Taylor and Pierce (1999), which utilized a stronger longitudinal design, found an effect of PM practices on organizational commitment but not on job satisfaction. Thus, it appears that the stronger the design, the less the evidence for a link between PM and these job attitudes. An exception to this is Mabey (2001), which utilized a matched sample of nonparticipants and found that participating in a 360-degree program leads to more positive attitudes about the organization.

b. Performance. The other most frequently studied employee transfer construct is performance (34% of employee transfer articles), operationalized most often as overall performance (20 articles), organizational citizenship behavior (OCB; 14 articles), and task performance (11 articles); counterproductive behavior (three articles) and career success (two articles) have also been examined. There is quite a bit of evidence that employee performance is related to PM, including the implementation of new or different PM systems (e.g., Pampino, MacDonald, Mullin, & Wilder, 2003; Stumpf, Doh, & Tymon, 2010) as well as specific elements of PM. For example, research suggests that overall and task performance are positively related to the implementation of goal setting (e.g., Klein & Snell, 1994; Pollack et al., 1994), feedback (e.g., Pollack et al., 1994; Wang, 2007), and more developmentally oriented PA programs (e.g., Tharenou, 1995). More discussion and participation in the PA interview is also associated with higher employee performance (e.g., Nathan et al., 1991; Prince & Lawler, 1986). Regarding different approaches to PA, higher employee performance has been found to result from BOS rating formats compared with graphic rating scale formats (Tziner, 1999), with greater as

(Appendices continue)

compared with less rating segmentation (Bartol, Durham, & Poon, 2001), and with process- as opposed to results-oriented feedback (Lam & Schaubroeck, 1999). The manager's role in PM is also important, with research showing that the types of goals set with one's supervisor (cooperative vs. competitive, Tjosvold & Halco, 1992), the type of power employed by one's supervisor in the PA context (Wexley & Snell, 1987), and perceived interactional justice within the PA setting (Masterson et al., 2000; Thurston & McNall, 2010) are all positively related to employee performance. Increased OCBs are similarly positively related to these same elements of PA process (e.g., Findley, Giles, & Mossholder, 2000; Masterson et al., 2000; Norris-Watts & Levy, 2004; Si & Li, 2012; Zheng, Zhang, & Li, 2012). The most common moderators examined involve the feedback component of PM, showing that performance is most likely to improve when feedback is more frequent (Bhave, 2014; Kuvaas, 2011; Pampino et al., 2003), timely and specific (e.g., Northcraft et al., 2011), when reflection accompanies it (e.g., Anseel et al., 2009), and with greater self-awareness (e.g., Korsgaard & Roberson, 1995). Unlike with job attitudes, much (but certainly not all) of the research examining the link between PM and employee performance has avoided common method issues, with 15 of these studies measuring performance using managers as the source. Despite the generally consistent evidence that performance can and does improve as a result of multiple aspects of PM, there also is a handful of studies failing to find such effects (see Erdemli et al., 2007; Milanowski, 2005; Shen, D'Netto, & Tang, 2010).

c. Withdrawal. Of the 20 withdrawal-related variables studied in PM, the vast majority (17) have been operationalized as intention to turnover versus remain; actual turnover has been studied in one article (Milanowski, 2005, no effect found), and another article measured withdrawal as neglect, or putting in less effort (Si & Li, 2012). It has been suggested that "a high-quality PA system deters turnover (Peterson, 2004; Brown, Hyatt, & Benson, 2010) [and] a low-quality PA system increases intentions to leave (Brown et al., 2010)" (Bambacas & Kulik, 2013, p. 1936). There does appear to be some evidence of this, especially if "high-quality" is defined in terms of due process. The empirical research shows, specifically, that employee turnover intentions are positively related to the use of computer aided performance monitoring (Chalykoff & Kochan, 1989; although for some employees this can be mitigated by attention to feedback and other aspects of the PA process) and to employees receiving negative as opposed to positive feedback (Lam et al., 2002, although these effects generally did not last beyond 3 months). Conversely, employees are more likely to express intentions to remain with the organization under due-process PA systems (Bambacas & Kulik, 2013; Taylor et al., 1998) and, related, when they have high procedurally just perceptions of PA (Juhdi, Pa'wan, & Hansaram, 2013; Masterson et al., 2000) and greater satisfaction with PA (Kuvaas, 2006;

Tymon, Stumpf, & Doh, 2010; Whiting & Kline, 2007), although much of this research is plagued by common method issues.

d. Motivation. Motivational criteria are the last category of employee transfer that has been studied with some regularity (8% of employee transfer research). Most commonly studied has been intrinsic motivation (four articles), including engagement (Gruman & Saks, 2011 provide a conceptual model that identifies key drivers of employee engagement at each stage of PM, but there has been little empirical testing of these ideas). The empirical PM research offers the following conclusions. First, the goal-setting and feedback components of PM, not surprisingly, are particularly important for employee motivation; research has found that cooperative (vs. competitive) goals lead to greater motivation to work hard (Tjosvold & Halco, 1992), that having high quality goals (i.e., specific and observable) varies across rating scale format (BOS vs. graphic rating scale vs. BARS; Tziner et al., 1996), and that higher quality (more timely and more specific) feedback on tasks leads to greater resource allocation on those same tasks (Northcraft et al., 2011). Second, greater intrinsic motivation results from PM that utilizes a more dialogue-based evaluation from the manager (Sundgren et al., 2005) and is more developmental in nature (Kuvaas, 2007), and when employees assess the aspects of PM (e.g., goal setting, evaluation, feedback) more positively (Juhdi et al., 2013; Kuvaas, 2006; Tymon et al., 2010). Some research has also failed to find a link between PM and motivation criteria: Taylor et al. (1998) found no relationship between due process PA and motivation to improve, and Taylor and Pierce (1999) found that the introduction of a new PM system did not increase effort. Together, this research suggests that there is likely not a straightforward relationship between PM and motivation; rather, the type of PM processes in place are likely to determine the type of motivation resulting (e.g., intrinsic vs. extrinsic, Sundgren et al., 2005).

e. Fairness/justice. Although fairness perceptions regarding PM itself are frequently studied (as discussed under Reactions, see section I in this Appendix), less frequently examined has been the impact of PM on more generalized fairness/justice perceptions (just 7% of the empirical work on employee transfer). Variables examined include overall justice (three articles), procedural justice (four articles), distributive justice (two articles), and interactional justice (three articles). This research has found more positive perceptions of justice/fairness are associated with the provision of feedback (but only among high performers, Lam et al., 2002); perceived usefulness of the PA process (Linna et al., 2012); and the implementation of PA for administrative purposes (Cheng, 2014), including reward allocation (Day, Holladay, Johnson, & Barron, 2014). In terms of moderators, Linna et al. (2012) discovered an interesting and potentially important finding: during negative changes in work life, employees' experienced usefulness of the PA feedback interview was especially important in helping prevent the deterioration of justice perceptions.

(Appendices continue)

V. Manager Transfer

Twenty-nine manager transfer variables have been examined across 26 studies. The focus has mostly been on quality of relationships (69%), with less research (28%) on quality of decisions made about employees. One study examined perceived supervisor effectiveness more generally (Burke, 1996), finding a positive relationship with employee beliefs that they had received an accurate and fair evaluation.

a. Quality of relationships with employees. Aspects of the PM system can alter the manager–employee relationship in various ways, and research has operationalized managers’ relationships with employees in terms of trust in the manager (four studies), supervisor liking/satisfaction (one study), leader-member exchange (LMX; two studies), perceived supervisor support (two studies), general supervisor-subordinate working relationship (five studies), and the quality of the coaching relationship (two studies). This research shows that managers report better working relationships with employees after using a due process PA system (Taylor et al., 1995) and that employee feedback orientation positively relates to both LMX (Dahling et al., 2012) and the quality of the employee–supervisor coaching relationship (Gregory & Levy, 2012). The quality of the employee–supervisor coaching relationship results from effective communication and the facilitation of development by the supervisor (Gregory & Levy, 2011). Conversely, the supervisor–subordinate relationship can degrade under certain PM conditions, including forced distribution PM systems (McBriarty, 1988). Research has also examined more specific aspects of a manager’s relationship with employees, including trust in the manager (which is related to employees expressing noninstrumental voice and being assertive in the PA interview; Korsgaard, 1996; Korsgaard et al., 1998); employee confidence in future collaboration with his or her manager (related to the manager establishing cooperative as opposed to competitive goals with the employee, Tjosvold & Halco, 1992); liking of one’s supervisor (related to supervisors’ impression management and provision of feedback; Kacmar et al., 1996); and satisfaction and cooperation with one’s supervisor (which increased with the implementation of a new PM system including merit-based pay, but only for low performers, Taylor & Pierce, 1999).

b. Quality of decisions made about employees. Despite the important implications for downstream criteria such as the emergence of unit-level HCRs, very little empirical research (four studies) has examined the quality of decisions made about employees as an outcome of PM. Research has found that the implementation of a forced distribution PM system actually decreased the quality of managers’ decisions relating to job assignments and resource utilization (McBriarty, 1988), and that the accuracy of employee-related decisions by managers was not related to their attitudes about appraisal but was related to their self-monitoring personality (Jawahar, 2001). More research should examine the

degree to which PM actually provides useful information for making other HR decisions (see Roberts, 1995).

VI. Unit-Level Human Capital Resources

Our review uncovered only nine studies that have empirically examined unit-level HCRs as an outcome of PM. There is a distinction in multilevel scholarship between the level of theory and the level of measurement (Kozlowski & Klein, 2000), and it is important to note that our categorizations here are based on the level of theory, not measurement. Nonetheless, all but one of these nine studies (Mulligan & Bull Schaefer, 2011, a simulation at the employee level) measured this variable at the organization level.

Six studies examined skills/abilities/potential capabilities, including adaptability/flexibility (Mullin & Sherman, 1993); performance potential of the workforce (“the average potential of an organization’s workforce to perform on the job,” Mulligan & Bull Schaefer, 2011; Scullen, Bergey, & Aiman-Smith, 2005); workforce quality (Giumetti, Schroeder, & Switzer, 2015) and staff competency (Zheng et al., 2006); and employees’ knowledge about how their work relates to the organization’s strategy (Ayers, 2013). This research shows that each of these “ability” unit-level HCRs can be impacted by PM. Interestingly, three of the studies in this category are about FDRS specifically (and are simulations; Giumetti et al., 2015; Mulligan & Bull Schaefer, 2011; Scullen et al., 2005). These results suggest that improvement in workforce potential and quality as a result of FDRS should be most noticeable over the first few years (Giumetti et al., 2015; Scullen et al., 2005), except for the findings of Mulligan and Bull Schaefer (2011), which suggested that temporary use of FDRS may do more harm than good in terms of workforce performance potential.

Three articles examined motivational capabilities. In the context of municipal PM systems, Roberts (1995) found that most respondents agreed that the PM system had a positive effect on employee motivation. Zheng et al. (2006) found positive effects of PA on staff commitment (per Pfeffer, 1998 and Youndt, Snell, Dean, & Lepak, 1996, commitment is classified as motivational), which mediated the relationship with firm performance. On the other hand, McBriarty (1988) found that a forced distribution system in the Air Force had a negative effect on motivation at the organization level, arguing that such systems focused “an inordinate amount of attention on the basic human concerns about survival, security, and ego maintenance at the expense of the higher order ‘motivators’ of more productive organizational behavior” (p. 428). We uncovered no empirical studies examining the relationship between PM and unit-level opportunity capabilities. Some (e.g., Combs et al., 2006) have suggested that climate is part of the opportunity part of the AMO framework, and there are studies looking at the impact of PM on climate. However, following Ployhart and Moliterno (2011), we categorize climate as an emergence enabler as opposed to a HCR and therefore review this in the next section.

(Appendices continue)

VII. Emergence Enablers

There is evidence that PM can affect an organization's climate, culture, and perceptions of leadership, all of which are important emergence enablers in our model. Eleven studies have examined aspects of climate as outcomes of PM, operationalized as office morale (Burke, 1996), group- or organization-level satisfaction (Daley, 1986; Mullin & Sherman, 1993), a "support" dimension of organizational culture (Mamatoglu, 2008), perceived psychological contract fulfillment (Raeder, Knorr, & Hilb, 2012), six dimensions of organizational climate (Kaya, Koc, & Topcu, 2010), and ethical organizational climate (Guerci, Radaelli, Siletti, Cirella, & Rami Shani, 2015). One study examined how PM affects the creativity culture of an organization (Sundgren et al., 2005), and one study examined the relationship between PM and strong positive perceptions of leadership (Lakshman, 2014), also considered an aspect of climate (Rentsch, 1990). This research suggests that climate, culture, and leadership can be influenced by the implementation of new systems as well as different types of PM systems. For example, in a longitudinal study, Mamatoglu (2008) found that a new 360-feedback system positively impacted employees' perceptions of a support and achievement culture; and Raeder et al. (2012) found that PA related to perceived psychological contract fulfillment, but only in the presence of performance-based pay (a tangible consequence). Sundgren et al. (2005) found that dialogue- versus control-based PA systems had a stronger impact on the organization's creativity culture; and Guerci et al. (2015) found that the use of performance goals and behavior-based evaluations is linked to egoistic, rather than ethical, climates. Another emergence enabler—trust in management—was examined in one study. Mayer and Davis (1999) found that the implementation of a more acceptable PA system increased trust for top management, and that this relationship was mediated by the three factors of trustworthiness: ability, benevolence, and integrity. The unit's ability to learn via the sharing of knowledge and information is part of cognitive emergence enabling states (reflecting the unit's ability to acquire, absorb, and transfer information) and has been examined as an outcome of aspects of PM in four studies. Operationalizations of this include the communication atmosphere of the unit (Mamatoglu, 2008), the knowledge sharing of R&D employees (Liu & Liu, 2011), knowledge management effectiveness (Tan & Nasurdin, 2011, which served as a mediator between PA and organization-level innovation), and organizational learning (Wang, Tseng, Yen, & Huang, 2011). This research has suggested that these emergence enabling states can be impacted by high quality PA practices in general (Liu & Liu, 2011; Tan & Nasurdin, 2011; Wang et al., 2011) and the implementation of a 360-feedback system specifically (Mamatoglu, 2008).

A fourth category of emergence enablers concerns team cohesion, trust, and collaboration. Findings across four articles suggest that these elements can definitely be affected by the type of PM system, either positively or negatively. For example, in three laboratory experiments, Song, Sommer, and Hartman (1998)

showed that modifying PA to include intergroup behavior explicitly (and an external supervisor as evaluator) led to more helping behavior and more positive attitudes toward cooperating; and Wang (2007) found that a new approach to evaluating teachers that relied on additional interaction, classroom observation, and feedback significantly increased the frequency of teacher collaboration and peer feedback. Conversely, some forms of PM can have negative effects on team cohesion and team effectiveness, including forced distribution systems (McBriarty, 1988) and those based on individual contributions and rewards and otherwise incongruent with a teamwork culture (Rowland, 2013). Finally, we conceptualize the unit-level quality of human capital decisions as another important emergence enabler. We found only one article examining this. Lawler (2003) evaluated the organization-level effectiveness of the PM systems of 55 Fortune 500 companies on two factors: effectiveness for influencing performance (the right kind of performance) and effectiveness for differentiating between top and poor performers/talent. Their results show that PM systems were more effective on these two criteria when there is a connection between the results of PM and the reward system of the organization.

VIII. Firm Performance

The articles we review here are those where the effects of PM specifically (not just bundled HR practices) on organization performance could be isolated. We found 16 such studies that examined operational outcomes (almost all of which were conducted at the organization level) and four articles that examined financial outcomes. Overall this research shows that PM can indeed have a positive impact on organizational outcomes, both operational (especially turnover) and financial performance.

a. Operational outcomes. Operational outcomes examined include labor productivity (Roberts, 1995 found that the majority of respondents believed their PA system had a positive effect on employee productivity) and production quality or quantity (which has been found to be positively related to the existence of PA, Zheng et al., 2006, and Lee, Lee, & Wu, 2010; as well as more related to "progressive" than "traditional" PA approaches, where the former includes more informal and multiple rating source approaches, Waite, Newman, & Krzystofiak, 1994). Organizational innovation is another operational outcome, and two studies have examined its relationship to PM. Tan and Nasurdin (2011) found PA practices had a positive effect on administrative innovation but not on product or process innovation (and knowledge management effectiveness mediated this relationship). Jiang et al. (2012) found evidence for the link between PA and both administrative and technological innovation, but this was not mediated by creativity as it was for other HR practices. Two studies examining improved safety performance as an outcome suggested PM can be very effective in this regard. Laitinen and Ruohomäki (1996) found a new PM approach oriented around safety behavior

(Appendices continue)

(including weekly graphic feedback) at building construction sites in Finland significantly improved safety, and [Reber and Wallin \(1994\)](#) found a PM program around safety behavior in offshore oilfield diving significantly reduced OSHA-recordable occupational injuries and accidents. Research has also examined the impact of PM on turnover (six articles) and/or absenteeism (two articles) at the organization level, suggesting that turnover can be reduced by PA practices in general ([Galang, 2004](#); [Zheng et al., 2006](#)) and other “high involvement” practices such as electronic performance monitoring ([Batt, 2002](#)). These same articles also suggest that turnover at least partially mediates the relationship between these practices and financial performance. In addition, a survey about PA systems in municipal governments showed that the majority of respondents perceive the systems to be effective at retaining good employees, but they were judged as less effective for controlling absenteeism ([Roberts, 1995](#)). On the other hand, [Peretz and Fried \(2012\)](#) found, in an organization-level study across 21 countries, that congruence between societal cultural practices and the characteristics of PA practices (i.e., formality, focus on development, multiple sources of raters, and percentage of EEs evaluated) affects absenteeism and turnover, but there was greater support for absenteeism than for turnover.

b. Financial outcomes. Our review identified four articles that examined the link between PM and aspects of firm financial performance (all measured at the organization level). [Zheng et al. \(2006\)](#) operationalized firm performance as increased sales, market competitiveness, and expected growth (all measured via interview responses), and found that a sound PA system generated better “HRM outcomes” (e.g., turnover, commitment, competency) which, in turn, contributed positively to financial performance. [Yang and Klaas \(2011\)](#) measured financial performance as the ratio of operating profit to assets (which represents how effectively firm assets are utilized in achieving profitability) and found that pay dispersion was less negatively related to firm financial performance when the organization invests more in performance evaluation and feedback. Sales growth was measured in [Batt \(2002\)](#) and was found to be positively impacted by the PM practice

of electronic performance monitoring (among other high involvement HR practices), as mediated by turnover. Finally, [Goh and Anderson \(2007\)](#) examined the return-on-investment (ROI) of a PM learning curriculum, which outlined how managers were supposed to improve the performance of their people and how employees were expected to take responsibility for their own development. The ROI was based on five impact factors (personal productivity, team efficiency, improved quality, increased net sales, and reduced cost) and was found to be 122%.

c. Other outcomes. We also found a number of empirical articles that examined aspects of organizational performance that could not be clearly categorized under the above categories. Some of these examined the impact of PM on more general (or undifferentiated) aspects of organizational performance (e.g., [Irs & Türk, 2012](#), found that the PA system positively impacted school performance on a number of performance indicators). Several others examined subjective ratings of perceived organizational performance. For example, in [Daley \(1986\)](#), Iowa public employees reported on the extent to which the “organization is effective in accomplishing its objectives” as a result of PM. In [Rodwell and Teo \(2008\)](#), managing directors were asked to evaluate their organizations’ performance as compared with similar organizations and relative to market competitors over the past 3 years (these performance indicators were positively related to the adoption of PA practices). In [Raeder et al. \(2012\)](#), participants were asked to assess the performance of their organizations compared with others in the sector on six items: service quality, productivity, profitability, product to market time, rate of innovation, and stock market performance. In [Galang \(2004\)](#), respondents were asked how accurately each of the following described their respective companies on a 5-point scale: produces high quality goods, has a promising future, manages its people well, is flexible enough to change, has high quality people, has a strong unified culture, is very effective overall, has a very satisfied workforce, has a very productive workforce, and is seen as a leader in industry (PA practices were strongly related to these performance ratings).

(Appendices continue)

Appendix B

Empirical Estimates of Criterion–Criterion Relationships

This appendix empirically summarizes the research reporting bivariate relationships among our evaluative criteria. We computed the average sample-weighted correlation for any criterion–criterion relationships with at least two samples (see also Figure 2). Where possible, we report these average correlations separately by subcategories of criteria; where that is not possible, we acknowledge that (along with generally small numbers of samples) as a limitation in interpreting these results.

Employee Reactions → Employee Learning

In the training evaluation literature there is an assumption of (and empirical evidence for) a positive link between reactions and learning (see Alliger et al., 1997). We found several PM articles (nine samples across nine articles) that included these empirical relationships, revealing a positive relationship between employee reactions and employee learning (average $r = .23$). In these relationships, employee reactions most commonly (60% of the time) focused on fairness reactions; also examined were emotional reactions, satisfaction with PM, and perceived accuracy and usefulness. Learning was most commonly measured as motivation to improve (in four studies; for just these studies, average $r = .35$). Unfortunately, in several studies reactions and learning were measured via the same source at the same time. Overall this research suggests that employee reactions to PM relate positively to employee learning from aspects of PM (especially motivation to improve). However, in the case of justice reactions this relationship appears a bit more complicated. For example, Selvarajan and Cloninger (2012) observed that PM-related motivational learning can improve with perceived procedural and interactional fairness, but not with perceived distributive fairness. Taylor et al. (1998) actually reported a moderately negative relationship between distributive justice of PM and learning outcomes such as self-efficacy for skill improvement and goals for improved future performance.

Employee Reactions → Employee Transfer

One reason employee reactions to PM matter is that positive reactions can “transfer” to important criteria on the job, including improved job attitudes, views of one’s supervisor, and aspects of performance (Korsgaard et al., 1998; Youngcourt et al., 2007). Reactions are also thought important in the social exchange between PM partners (i.e., managers and employees, Pichler, 2012), suggesting they may relate in important ways to general attitudes and behaviors on the job. These were in fact the relationships most frequently examined in our review, with 24 samples across 23 articles examining relationships between employee reactions and employee transfer. The average r was .29. Employee reactions

most commonly focused on justice and other cognitive reactions (65% of studies); and the most common employee transfer variables were organizational commitment (where $r = .35$) and job satisfaction (where $r = .37$). Several of these studies were again prone to same method bias, thus potentially inflating these relationships. In addition, it is unknown whether the magnitude of these direct relationships would hold if one accounted for employee learning (a potential mediator; see following section). Interestingly, several studies in this area converged to suggest that perceived fairness of PM, as opposed to perceived value of the PM, drives employee turnover intentions specifically. For example, Burke (1996) found that employee intent to quit was (negatively) related to due process and perceived fairness of PA, but not to the meaningfulness of the personal development plan created. Related, Si and Li (2012) found that the extent to which PA is developmentally useful was negatively related to employee neglect but not to exit (i.e., turnover intentions). As further support for the importance of due process and fairness on turnover, Poon (2004) found that if employees believe ratings were manipulated because of raters’ personal bias and intent to punish employees, this leads to greater turnover intentions; but there is no effect on turnover intentions when employees believe ratings were manipulated for motivational purposes. This is interesting because in both cases the ratings were intentionally manipulated (thereby ostensibly decreasing the utility of the ratings and the PM system), but manipulation for motivational purposes presumably is seen as more fair than manipulation for personal bias or intent to punish, and fairness appears to trump utility in driving turnover intentions.

Employee Learning → Employee Transfer

In models of training evaluation, it is learning that is the most proximal determinant of transfer (Alliger et al., 1997; Kraiger et al., 1993). In the context of PM, this suggests that it is what employees learn from the PM experience that affects their overall attitudes and behaviors back on the job, and our review does reveal some empirical evidence of this. We found nine studies reporting this relationship, with an average r of .38 (interestingly this relationship did not vary based on whether the data were same- or different-source, $r = .38$ vs. .39, respectively). For the attitudinal and motivational subcategory of learning ($k = 7$), the relationship was even larger (average $r = .45$), for multiple subcategories of transfer (e.g., for job attitudes, $r = .46$; for performance $r = .44$). These estimates suggest that what employees learn from the PM experience (especially in terms of attitudinal and motivational learning) can indeed transfer into improved attitudes and performance back on the job.

(Appendices continue)

Manager Reactions → Manager Learning

Just as employees' reactions to PM should positively relate to learning from PM, so should managers' reactions. Our review showed that manager reactions are positively (albeit weakly) related to manager learning ($r = .14$). This estimate is based on seven distinct samples reported in three different articles. Manager reactions variables included satisfaction, fairness, difficulty, and discomfort; manager learning variables were limited to rating quality and distortion.

Manager Reactions → Manager Transfer

Our review showed manager reactions are also positively related to manager transfer ($r = .30$). But this estimate is based on only two studies published together in a single article on managers' reactions to the implementation of a procedurally just PM system (Taylor et al., 1998). Specifically, managers' satisfaction with the appraisal system related positively to a favorable working relationship with their employees. However, in both samples, the two variables were reported by managers at the same point in time, likely inflating the magnitude of the relationship.

Manager Learning → Manager Transfer

Our review showed manager learning was positively and strongly related to manager transfer ($r = .53$). This overall positive estimate is based on three studies. One study showed a very strong positive relationship (Gregory & Levy, 2012), but two of these studies (from the same article referenced in the previous section; Taylor et al., 1998) actually had a negative relationship (average $r = -.10$). The latter examined managers' self-reported distortion of appraisals and found that more distortion (coded as less learning here) related positively to working relationships with their employees. Manager learning is likely to be positively related to the quality of decisions subcategory of manager transfer, but as these results show, it might negatively impact the quality of relationships

with employees (especially given that greater learning may imply lower ratings).

Manager Learning → Employee Transfer

We also found three articles that reported relationships between aspects of manager learning about PM and employee transfer, showing a positive relationship (average $r = .51$). In particular, managers' learning with regard to providing feedback through the year and discussing past and future performance in the PA interview ultimately related to employee job satisfaction (Inderrieden et al., 2004); and managers' learning in terms of interactive behaviors that help employees convert tacit knowledge to explicit knowledge and managers' performance enhancement strategies both related positively to subordinate performance (Lakshman, 2014).

Predictors of Unit-Level Outcomes

There were a handful of studies that reported relationships between employee- or manager-level criteria from our model and unit-level criteria. Unfortunately, for only one criterion category (employee transfer, $k = 4$) was there a sufficient number of samples to aggregate. The average r here was .37, but this was marked by a bimodal distribution, with two effect sizes in the $r = .60$ range (the link between employee transfer variables and organizational climate and innovation) and two in the $r = .06-.10$ range (for the link between employee transfer and bottom-line measures of organizational performance). This pattern suggests, not surprisingly, that employee transfer criteria are more strongly related to the more proximal unit-level criteria of emergence enablers and operational outcomes than they are the more distal bottom-line measures of organizational performance.

Received May 15, 2016

Revision received September 19, 2018

Accepted October 2, 2018 ■