Journal of Management
Vol. 44 No. 6, July 2018 2209–2245
DOI: 10.1177/0149206318755303

Putting the *System* Into Performance Management Systems: A Review and Agenda for Performance Management Research

Deidra J. Schleicher
Texas A&M University
Heidi M. Baumann
Bradley University
David W. Sullivan
Texas A&M University
Paul E. Levy
University of Akron
Darel C. Hargrove
Central Michigan University
Brenda A. Barros-Rivera
Texas A&M University

It has been 13 years since the last comprehensive review of the performance management (PM) literature, and a lot has changed in both research and practice in that time. The current review updates (identifying new research directions post 2004) and extends this previous work by creating a systems-based taxonomy and conceptual model of PM. We then use this model to interpret and integrate the extensive work in this area and to identify fruitful and systems-based directions for future PM work. As input to our conclusions, we reviewed the last 36-plus years of PM research (1980–2017) and conducted a comprehensive coding of all empirical PM articles. We offer several specific directions for future PM research, with the ultimate goal of improving PM in practice.

Acknowledgment: This article was accepted under the editorship of Patrick M. Wright.

Supplemental material for this article is available with the manuscript on the JOM *website.*

Corresponding author: Deidra J. Schleicher, Department of Management, Mays Business School, Texas A&M University, 4113 TAMU, College Station, TX 77843-4113, USA.

E-mail: dschleicher@mays.tamu.edu

Keywords: *performance management; performance appraisal; performance evaluation; systems theory*

Performance appraisal (PA) and performance management (PM) have historically been areas of substantial focus in both research and practice.[1] The vast majority of organizations have formal PM systems (91% according to Cascio, 2006), and there is an incredibly voluminous research literature, including a number of previous reviews (e.g., Ilgen, Barnes-Farrell, & McKellin, 1993; Landy & Farr, 1980; Levy & Williams, 2004). Yet despite the omnipresence of PM in both practice and the research literature, there remain many unanswered questions about its effectiveness. In addition, it has been 13 years since the last comprehensive review of PA/PM (Levy & Williams, 2004), and in that time a number of new directions have emerged. Some of these developments have been the subject of vociferous debate in the literature and amongst practitioners, such as doing away with ratings or other formal aspects of PM (e.g., Adler et al., 2016). As Levy, Tseng, Rosen, and Lueke recently noted, "If you think that we don't have a problem with current PM approaches and strategies, you can do a simple Google search and tap into the uproar" (2017: 156). We believe an updated scholarly review is in order that can help shed light on these and other issues based on empirical evidence. Accordingly, we conducted an extensive and critical review of more than 36 years of the PM literature, with multiple goals and intended contributions.

First, given the time elapsed since the last comprehensive review (Levy & Williams, 2004), as well as the amount of PM research conducted in those intervening years and the extent of discourse on this topic, there is need for an update. This is particularly true given the unprecedented changes to PM systems in practice, with varying amounts of substantiation from the research literature (Levy et al., 2017). The PM landscape is very different than it was even 10 years ago. We aim to review new PM practices and areas of inquiry that have emerged in the scholarly literature and discuss how these compare to older research. We provide summary information regarding what questions PM research is currently studying, how it is studying these questions, and what we know (and do not know) as a result. This information is provided in both qualitative and numerical form; the latter is possible because we coded all empirical PM articles from 1980 to 2017 as input to this review, making this the most comprehensive review of empirical PM research to date.

Second, our review is distinguished from prior reviews by focusing on PM, as opposed to PA. As defined by Aguinis, PM "is a *continuous process of identifying, measuring, and developing the performance of individuals and teams and aligning performance with the strategic goals of the organization*" (2013: 2). The term *PM* is understood to have arisen when practitioners (and eventually scholars) began talking about transforming PA from an event to a process (see Bretz, Milkovich, & Read, 1992; Kinicki, Jacobson, Peterson, & Prussia, 2013; M. J. Williams, 1997). Whereas PA is generally understood to be "a discrete, formal, organizationally sanctioned event, usually not occurring more frequently than once or twice a year" (DeNisi & Pritchard, 2006: 254), PM is seen as a broader set of ongoing activities aimed at managing employee performance (M. J. Williams, 1997). In other words, PA can be thought of as a subset of PM (see also Levy et al., 2017). Levy and Williams (2004) referred to the scope of their review as PA; the current review is explicitly positioned more

broadly as PM. Consequently, we cover some PM components not included in prior reviews, such as setting performance expectations and employee coaching. We also emphasize the role of context (as an input to PM), affording the first real opportunity to evaluate progress in this area since Levy and Williams identified it as an emerging theme in research.

Third, our review introduces, and is organized around, a systems-based model of PM, which provides a much-needed taxonomy for this area. Our review of 36-plus years of research reveals no articulated consensus on what the main components of PM are or the primary variables composing them. The lack of such a taxonomy to direct work in this area has been explicitly mentioned as a difficulty in multiple research-practice panel discussions on PM (Cavanaugh, Levy, Schleicher, Anseel, Colquitt, & Hunt, 2013; Schleicher, Levy, Baumann, & Hartwell, 2012). To illustrate, some scholars place any activity related to managing employees under the PM label. For example, Roberts (2001) notes that

> PM involves the setting of corporate, departmental, team, and individual objectives; the use of PA systems; appropriate reward strategies and schemes; training and development strategies and plans; feedback, communication, and coaching; individual career planning; mechanisms for monitoring the effectiveness of PM system and interventions and even culture management. (as quoted in den Hartog, Boselie, & Paauwe, 2004: 558)

In addition, DeNisi and Smith (2014) propose that PM should encompass *all* human resource (HR) practices designed to give employees the means, motivation, and opportunity to improve firm-level performance. Such broad definitions are at odds with other accepted definitions in the literature (e.g., Aguinis, 2013) and, importantly, do not place meaningful boundaries around what PM is and is not, as they essentially exclude no HR practices as being outside of PM. Moreover, as displayed in our model, there is a lot more to PM than just practices; the context and individuals involved matter greatly in PM, and there is a need for a parsimonious way to categorize all of these other relevant variables. As we develop later, we believe the lack of such a taxonomy in the extant literature has hindered progress towards cumulative scientific knowledge. Accordingly, we provide a systems-based taxonomy of PM grounded in both the extensive empirical literature and theory and that can usefully guide both future research and practice in this area. This taxonomy is also useful for the current review, providing an effective framework for organizing the truly voluminous PM literature, including what we are currently studying, how this differs from what we studied in the past, and what we do and do not know about PM as a result.

Fourth, we make an important conceptual contribution by applying a systems theoretical framework to both assimilate current knowledge and chart a course for future research in PM. A system, simply put, is a set of interrelated elements, such that a change in one element affects other elements in the system, and an open system is one that also interacts with its environment (Katz & Kahn, 1978). The interrelated components of a system are designed to work together and function as a whole to achieve a common purpose (Boulding, 1956). Despite the fact that PM processes in organizations are often referred to as PM *systems*, historically PM has not been researched in this way. We feel that a systems approach is essential both for distilling knowledge about the effectiveness of PM from the extant literature (as such questions ultimately rely on an examination of multiple components and how they interrelate) and for identifying additional questions that we *should* be studying and ways in which we ought to be studying them. Both of these are made possible by an articulation and

application of systems theory principles (which we discuss below), and both lead to the ultimate goal of better understanding PM systems in organizations.

To illustrate the unique value of our approach, we provide three examples of recent "debates" within PM and discuss how our review and the systems-based model on which it is based affect the understanding of these questions and suggest the direction of future research that might help resolve them. These include assertions that (a) the informal aspects of PM are more important than the formal aspects (e.g., Pulakos & O'Leary, 2011), (b) performance ratings are inaccurate and not particularly useful and therefore should be done away with (e.g., Adler et al., 2016), and (c) PM processes should be streamlined to remove their "low value" aspects (Effron & Ort, 2010). After we describe our literature search, introduce our systems-based model, and summarize the research about each of the components of PM in turn, we integrate this research vis-à-vis systems theory and discuss the implications for understanding these and other timely issues in PM.

Literature Search and Coding of Empirical Articles

We comprehensively reviewed the last 36-plus years (1980–June 2017) of PM work, using a two-step literature search to find relevant articles.[2] First, we searched Business Source Complete, using the terms *performance management*, *performance appraisal*, and *performance evaluation*. Second, given our primary focus on recent research and to ensure we did not miss any relevant articles, we also manually searched a number of journals from 2004 to 2017 that either are considered top management outlets or are specialty outlets that publish a significant amount of PM-related work (according to our database of articles): *Academy of Management Journal, Academy of Management Review, Journal of Applied Psychology, Journal of Management, Organizational Behavior and Human Decision Processes, Personnel Psychology, The International Journal of Human Resource Management*, and *Public Personnel Management*. This resulted in 1,915 articles.

An important aspect of our review involved systematically coding and classifying all empirical PM articles into the components of our model and along multiple other dimensions (see Table 1). We carefully reviewed the above articles and parsed them into a subset that empirically studied PM (including case studies and qualitative research in addition to quantitative research). This process resulted in a set of 575 empirical articles (614 separate studies). Each of these studies was then coded by the authors (we also captured information on all variables studied in an article as well as the article's results). Coders were trained via multiple calibration sessions, and a subset of 20 articles coded by all five coders showed good agreement on coding. Following the coding process, an additional 62 studies were excluded for having very weak methodology,[3] resulting in a final sample of 552 empirical studies. Summary information on these articles is reported in Table 1.

Our Systems-Based Model of PM

We created our systems-based model of PM (see Figure 1) with two purposes in mind: to offer a taxonomy for organizing the myriad variations (i.e., all the moving parts) of PM in both research and practice and to provide a conceptual framework for integrating all of this research to better understand the effectiveness of PM and identify future research needs. We relied specifically on the congruence systems model by Nadler and Tushman (1980, 1984)

Table 1
Descriptive Information on Coded Articles

	All Empirical Studies			Tasks		Individuals		Formal Processes		Informal Processes		Inputs		Outputs	
	k	*n*	%	*k*	%	*k*	%	*k*	%	*k*	%	*k*	%	*k*	%
Total studies	552	393,298	100	104	18.84	273	49.46	235	42.57	176	31.88	110	19.93	229	41.49
Study setting															
Lab	161	43,792	29.17	32	30.77	104	38.10	62	26.38	47	26.70	16	14.55	80	34.93
Field	323	326,459	58.51	61	58.65	143	52.38	144	61.28	106	60.23	86	78.18	115	50.22
In situ	68	23,047	12.32	11	10.58	26	9.52	29	12.34	23	13.07	8	7.27	34	14.85
Study focal participant															
Employees	288	258,463	42.99	50	40.00	141	41.23	115	39.79	103	50.49	53	42.40	107	36.90
Students	135	19,644	20.15	31	24.80	85	24.85	54	18.69	39	19.12	11	8.80	77	26.55
Managers	200	61,687	29.85	34	27.20	111	32.46	88	30.45	54	26.47	30	24.00	94	32.41
Organization level	47	32,127	7.01	10	8.00	5	1.46	32	11.07	8	3.92	31	24.80	12	4.14
Data source															
Archival	45	69,202	6.65	4	2.90	24	7.14	21	7.02	10	4.61	12	9.16	23	7.80
Experiment	176	20,731	26.00	40	28.99	101	30.06	79	26.42	52	23.96	17	12.98	87	29.49
One-time survey	288	218,745	42.54	54	39.13	131	38.99	122	40.80	104	47.93	80	61.07	105	35.59
Qualitative	50	7,016	7.39	7	5.07	18	5.36	21	7.02	12	5.53	12	9.16	16	5.42
Quasi-experiment	46	6,460	6.79	19	13.77	27	8.04	28	9.36	11	5.07	2	1.53	27	9.15
Time-lagged or longitudinal	72	53,257	10.64	14	10.14	35	10.42	28	9.36	28	12.90	8	6.11	37	12.54
PM purpose															
Not stated	193	117,434	31.28	28	24.14	88	29.73	61	23.02	72	35.29	44	32.12	61	24.50
Stated explicitly	157	153,890	25.45	27	23.28	76	25.68	70	26.42	43	21.08	23	16.79	69	27.71
Inferred	180	56,350	29.17	51	43.97	101	34.12	90	33.96	55	26.96	26	18.98	103	41.37
Varied	87	65,624	14.10	10	8.62	31	10.47	44	16.60	34	16.67	44	32.12	16	6.43

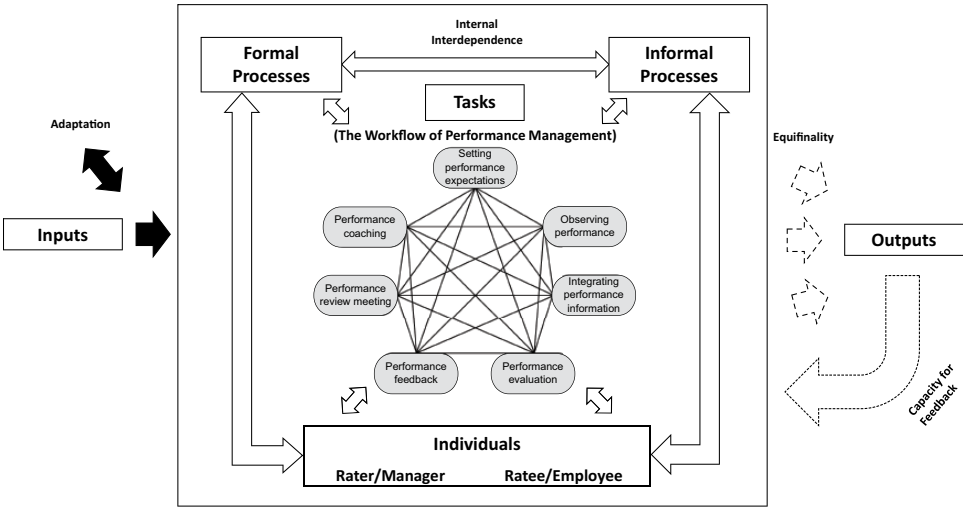
(continued)

Table 1 (continued)

	All Empirical Studies			Tasks		Individuals		Formal Processes		Informal Processes		Inputs		Outputs	
	k	*n*	%	*k*	%	*k*	%	*k*	%	*k*	%	*k*	%	*k*	%
Stated PM purpose															
Administrative	134	118,343	32.76	27	27.55	66	30.70	63	30.73	41	34.17	21	30.43	60	28.85
Developmental	80	45,313	19.56	18	18.37	29	13.49	45	21.95	23	19.17	12	17.39	39	18.75
Research	167	45,408	40.83	49	50.00	105	48.84	80	39.02	52	43.33	30	43.48	98	47.12
Selection	11	36,110	2.69	2	2.04	7	3.26	6	2.93	1	0.83	1	1.45	5	2.40
Other	17	7,197	4.16	2	2.04	8	3.72	11	5.37	3	2.50	5	7.25	6	2.88
Terminology used															
PE	75	71,134	12.54	7	6.14	43	14.10	26	10.24	27	13.92	12	9.84	36	14.12
PA	413	258,583	69.06	87	76.32	205	67.21	186	73.23	131	67.53	81	66.39	181	70.98
PM	47	74,547	7.86	10	8.77	25	8.20	23	9.06	20	10.31	15	12.30	16	6.27
PA/PM interchangeably	46	10,914	7.69	6	5.26	22	7.21	14	5.51	12	6.19	10	8.20	16	6.27
Distinguished between PA and PM	17	15,210	2.84	4	3.51	10	3.28	5	1.97	4	2.06	4	3.28	6	2.35
Component studied as															
Independent variables	831	331,843	61.97	96	82.76	234	64.82	201	77.61	146	65.47	100	74.07	54	21.86
Moderators	168	84,456	12.53	8	6.90	60	16.62	35	13.51	40	17.94	18	13.33	7	2.83
Mediators	50	15,240	3.73	0	0.00	22	6.09	3	1.16	15	6.73	6	4.44	4	1.62
Dependent variables	292	89,685	21.77	12	10.34	45	12.47	20	7.72	22	9.87	11	8.15	182	73.68
Tested component interactions	145		26.27	24	23.08	69	25.27	62	26.38	46	26.14	32	35.96	62	27.07

Note: PM = performance management; PE = performance evaluation; PA = performance appraisal.

Figure 1
A Systems-Based Model of Performance Management



for developing our framework. Nadler and Tushman (1980) developed this model explicitly to be less abstract and more pragmatic than general systems theories. In addition to specifying *inputs* and *outputs* to the system (as all systems models do), their model breaks down the process component into four interdependent factors: *tasks, individuals, formal processes*, and *informal processes*. Thus, our taxonomic model specifies that PM systems are made up of the *inputs* and *outputs* of PM as well as four interdependent factors composing the process components of PM: *tasks* of PM, *individuals* involved in PM, *formal* processes of PM, and *informal* processes of PM (see Figure 1). These six components represent the most general level of our PM systems taxonomy. Subcategories of this taxonomy were further distilled, using systems theory (Katz & Kahn, 1978; Nadler & Tushman, 1980) and the review and coding we did of the 552 empirical PM articles. The extended taxonomy resulting from this process is displayed in Table 2 and discussed further in a later section. Each section below discusses these six components in turn, summarizes directions since the last review and the conclusions that can be drawn from recent research, and identifies areas for future study. These sections consider each component *in isolation*, for purposes of clarity and comprehensiveness regarding our taxonomy (and because that is how the literature typically studies them).

Following the review of each PM component in isolation, we consider this research from the perspective of *systems principles*. Systems theory includes a number of principles regarding how elements of the system are likely to interact. Some of these principles are graphically depicted in Figure 1 (and there is a table in the online supplemental material that provides a more detailed explanation of additional principles not captured here). Together, these principles emphasize a much more complex and dynamic view of PM than has typically been addressed in extant research. As such, and as we discuss in this later section, they provide an important and unique conceptual foundation that helps us make better sense of the extant research and chart important directions for future research.

Table 2
Taxonomy of System Components and Subcategories

Component	Subcategory	Variables and Sample Research
Individuals	Ratee abilities, skills, and performance	• Performance mean, trend, and variation (Reb & Greguras, 2010) • Proactive behavior (Grant, Parker, & Collins, 2009) • Organizational citizenship behavior (Whiting, Podsakoff, & Pierce, 2008) • Cognitive ability (Blume, Rubin, & Baldwin, 2013)
	Ratee demographic, cultural, and position-related factors	• Gender (Lyness & Heilman, 2006) • Race (Hernandez, Avery, Tonidandel, Hebl, Smith, & McKay, 2016) • Age (Kooij, Guest, Clinton, Knight, Jansen, & Dikkers, 2013) • Physical attractiveness (Bande Vilela, Varela González, Fernández Ferrín, & Luisa del Río Araújo, 2007) • Cultural values (Ellis, 2012) • Need (Webb Day, Holladay, Johnson, & Barron, 2014) • Line versus staff jobs (Lyness & Heilman, 2006) • CEO position (Silva & Tosi, 2004) • Union status (Nurse, 2005)
	Ratee individual differences	• Core self-evaluations (Kamer & Annen, 2010) • Feedback orientation (Linderbaum & Levy, 2010) • Performance goal orientation (Culbertson, Henning, & Payne, 2013) • Autonomy orientation (Kuvaas, 2007) • Dispositional intrinsic motivation (Kuvaas, 2006)
	Rater abilities, skills, and cognitive processing	• PM-related skills (Becker, Antuar, & Everett, 2011) • Credibility (Kinicki, Prussia, Wu, & McKee-Ryan, 2004) • Use of rejection threshold (Pesta, Kass, & Dunegan, 2005) • Self-serving bias (Goerke, Möller, Schulz-Hardt, Napiersky, & Frey, 2004) • Own appraisal (Latham, Budworth, Yanar, & Whyte, 2008)
	Rater demographic, cultural, and position-related factors	• Gender (van der Leeuw, Overeem, Arah, Heineman, & Lombarts, 2013) • Ethnic origin (Rahman, 2006) • Cultural value orientation (Ng, Koh, Ang, Kennedy, & Chan, 2011) • Bicultural identity integration (Mok, Cheng, & Morris, 2010) • Interdependence (Saffie-Robertson & Brutus, 2014) • Experience (Bernardin & Villanova, 2005) • Power (Ferguson, Ormiston, & Moon, 2010) • Charismatic leadership (Tuytens & Devos, 2012b) • Transformational leadership (Gregory & Levy, 2011)
	Rater individual differences and beliefs	• Openness to experience (Ogunfowora, Bourdage, & Lee, 2010) • Agreeableness (Randall & Sharples, 2012) • Conscientiousness (Dewberry, Davies-Muir, & Newell, 2013) • Self-monitoring (Jawahar, 2005) • Implicit person theory (Heslin & VandeWalle, 2011) • Just world beliefs (Skarlicki & Turner, 2014)
	Rater motivational differences	• Motivation to assist in identification of strengths and weaknesses (Ng et al., 2011) • Rating goal (e.g., harmony, fairness; Wang, Wong, & Kwong, 2010) • Motivation to respond without prejudice (Hernandez et al., 2016) • Self-efficacy (Wood & Marshall, 2008)
	Rater and ratee interactions and dyadic variables	• Similarity (Pearce & Xu, 2012) • Leader-member exchange (Elicker, Levy, & Hall, 2006) • Quality of coaching relationship (Gregory & Levy, 2012)
Formal Processes	Rating format and approach	• Relative versus absolute ratings (Wong & Kwong, 2005) • Forced distribution rating system (Lawler, 2003) • Social comparative appraisal (Goffin, Jelley, Powell, & Johnston, 2009) • Display of positive versus negative information (Wong & Kwong, 2005)

(continued)

Table 2 (continued)

Component	Subcategory	Variables and Sample Research
	Rater training	• Frame-of-reference training (Gorman & Rentsch, 2009) • Source-monitoring training (Martell & Evans, 2005) • Training on written and verbal feedback (Erdemli, Sümer, & Bilgiç, 2007) • Whole brain training (Selden & Sowa, 2011)
	Dimensions of performance	• Implicit weighting of dimensions (Whiting et al., 2008) • Task versus citizenship performance (Lievens, Conway, & De Corte, 2008) • Acceptability of dimensions (Taormina & Gao, 2009)
	Technology and online systems	• Online system (Payne, Horner, Boswell, Schroeder, & Stine-Cheyne, 2009) • E-mail versus paper evaluation (Kurtzberg, Naquin, & Belkin, 2005)
	Rater source and multisource ratings	• Measurement equivalence in multisource ratings (Woehr, Sheehan, & Bennet, 2005) • Ceiling effect across multisource raters (Kulas & Hannum, 2006) • Integration between multisource feedback and performance appraisal (Haines & St-Onge, 2012) • Multisource appraisal (Mamatoglu, 2008) • Category of rater (Kulas & Hannun, 2006)
	Rater accountability	• Formal justification expectations (Curtis, Harvey, & Ravden, 2005) • Downward versus upward accountability (Mero, Guidice, & Brownlee, 2007) • Face-to-face versus written accountability (Mero et al., 2007)
	Ratee participation	• Involvement during design and implementation (Becker et al., 2011) • Self-ratings (Inderrieden, Allen, & Keaveny, 2004)
	Due process characteristics	• Goal setting (Kataoka, Cole, & Flint, 2006)
	Feedback characteristics	• Midyear developmental review (Becker et al., 2011) • Favorable narrative comments (David, 2013) • Interactionally just narrative comments (David, 2013) • Frequency of formal feedback (Bhave, 2014) • Feedback constructiveness (Alder & Ambrose, 2005) • Reflection step (Anseel, Lievens, & Schollaert, 2009)
Informal Processes	Ratee participation	• Voice during performance review meeting (Inderrieden et al., 2004) • Two-way communication (Kavanagh, Benson, & Brown, 2007) • Involvement in goal setting (Kavanagh et al., 2007)
	Rater accountability	• Anticipation of future collaboration (Randall & Sharples, 2012)
	Due process characteristics	• Continuous feedback (Kataoka et al., 2006)
	Feedback characteristics	• Feedback richness (Selvarajan & Cloninger, 2012) • Feedback helpfulness (Kuvaas, 2011)
	Feedback environment	• Feedback environment dimensions (Norris-Watts & Levy, 2004)
	Politics	• Perceptions of politics (Dhiman & Maheshwari, 2013)
Inputs	Industry/sector	• Specific industry (e.g., health care; Clarke, Harcourt, & Flynn, 2012) • For profit versus not for profit (Rodwell & Teo, 2008)
	Ownership type and organizational structure	• Public versus private (Abu-Doleh & Weir, 2007) • Management controlled versus owner controlled (Silva & Tosi, 2004) • Locally versus foreign owned (Akuratiyagamage, 2005) • Organizational size (Giumetti, Schroeder, & Switzer, 2015) • Unionized setting (Brown & Warren, 2011)

(continued)

Table 2 (continued)

Component	Subcategory	Variables and Sample Research
	Global context	• Multinational versus domestic (Singh, Mohamed, & Darwish, 2013) • Internationalization (Fee, McGrath-Champ, & Yang, 2011) • Global integration (Claus & Briscoe, 2009) • National culture (Peretz & Fried, 2012) • Specific culture (e.g., China; Taormina & Gao, 2009)
	Organizational culture and climate	• Team-based culture (Lievens et al., 2008) • Group-oriented culture (Haines & St-Onge, 2012) • Climate of trust in senior management (Farndale, Hope-Hailey, & Kelliher, 2011)
	Resources	• Charismatic leadership (Tuytens & Devos, 2012a) • Leadership support (Ayers, 2013) • Senior management concern for PM (Wood & Marshall, 2008) • Organizational investment in PM (Yang & Klaas, 2011)
	Strategy and PM purpose	• Strategic integration (Haines & St-Onge, 2012) • Vertical alignment (Decramer, Smolders, & Vanderstraeten, 2013) • Horizontal alignment (Decramer et al., 2013) • Strategy alignment with HRM (Lee, Lee, & Wu, 2010) • PM purpose (W. Zheng, Zhang, & Li, 2012)
	In the context of other HRM practices	• Role of PM above and beyond other human resource practices (Jiang, Wang, & Zhao, 2012) • In the context of organizational socialization efforts (Taormina & Gao, 2009) • Assessment center use (Macan, Mehner, Havill, Meriac, Roberts, & Heft, 2011) • Human resource configuration (Medcof & Song, 2013)
Outputs	Performance rating • Feedback content • Feedback sign • Development plan • Administrative decisions • Career planning • Documentation	
Tasks	Setting performance expectations • Observing performance • Integrating performance information • Performance evaluation • Performance feedback • Performance review meeting • Performance coaching	

Note: PM = performance management; HRM = human resource management.

The Tasks of PM

Borrowing from Nadler and Tushman (1980), the tasks component of our model refers to the "workflow of PM," or the activities involved in PM. These represent the basic or inherent work to be done by the system, particularly in light of its strategy. As Nadler and Tushman note, because "it's assumed that a primary (although not the only) reason for the system's existence is to perform the tasks consistent with strategy, the task is the starting point" (41). Nadler and Tushman further note that "the assessment of the adequacy of other components depends to a large degree on an understanding of the nature of the tasks to be performed" (41). For that reason, we begin our review of the PM literature with this component. We feel additional clarity is provided to our review by focusing first on the tasks that compose the PM process—not how these tasks are done, the context in which they are done, or the outcomes of these tasks (which all speak to other components of our systems-based model). We also believe this component is particularly important from a taxonomic perspective, as it allows us to place some important boundaries around what PM is.

The PM tasks identified as part of our model are grounded in both inductive (from our extensive review of PM articles) and deductive (based on other authors' lists of key PM activities or responsibilities; e.g., Aguinis, 2013; Kinicki et al., 2013) approaches. The result is seven key PM tasks: *setting performance expectations, observing employee performance, integrating performance information, the rendering of a formal summative performance evaluation, generating and delivering performance feedback, the formal performance review meeting*, and *performance coaching* (see Figure 1). It is meaningful that the word *performance* occurs in each task, especially for distinguishing PM from the broader area of talent management. PM refers to the managing of *performance* required in employees' *current* roles. Identifying potential and developing employees for future roles are essential elements of talent management (Silzer & Church, 2009), but we believe it further obfuscates an already complex field to suggest these are core elements of PM. It is also meaningful that this list of tasks excludes the more distal actions organizations might take to bring about performance change (such as training or development programs) as well as the administration of consequences based on performance (e.g., compensation). These more distal actions are part of our PM model but are conceptualized as *outputs*, rather than tasks, of PM in our model (see Figure 1). This classification is in line with how some others (e.g., Effron & Ort, 2010; Kinicki et al., 2013) have defined PM, but this issue has never been made explicit. Finally, as Figure 1 depicts, these seven tasks are all interrelated and often occur contemporaneously (although they may not all be present in any particular system). In each of the following subsections, we briefly review research focusing on each task in isolation.

Setting performance expectations. This task has not been explicitly included in prior PA/PM reviews, and focus on this task constitutes 14% of empirical PM studies. Research from other areas clearly suggests this task is important, including meta-analytic evidence for the direct (negative) link between role ambiguity and job performance (Tubre & Collins, 2000) and extensive evidence for the (positive) impact of goal setting on individual (Mento, Steel, & Karren, 1987) and group (Kleingeld, van Mierlo, & Arends, 2011) performance. Yet it is notable that goal setting has not generally been studied in the context of PM specifically (for exceptions, see Bipp & Kleingeld, 2011; Kuvaas, 2006; and Pritchard, Harrell, DiazGranados, & Guzman's, 2008, productivity measurement and enhancement system). We view it as promising, however, that a somewhat higher percentage of more recent PM articles are examining this task compared to earlier articles (17% vs. 10%). As one example, Pampino, MacDonald, Mullin, and Wilder (2004), in an experimental design, found that task clarification and goal setting as part of PM improved employee reactions and performance.

Observing employee performance. We know from older research that raters should have adequate opportunity to observe employees' performance (Ilgen et al., 1993), as this is essential for rating quality (e.g., Rothstein, 1990). Yet there has been very little recent work on this component of PM (just 2% post 2004, compared to 17% of earlier research). Very little is known about how busy managers do or should incorporate the observation of employee performance into their responsibilities. Although we know that diary keeping, for example, leads to more positive rater reactions and also allows for better recall of performance information and greater discrimination among employees (e.g., DeNisi & Peters, 1996), we lack field studies on diary keeping with actual managers that would shed light on ways to make this practical and manageable.

Integrating performance information. There has also been very little recent research on the task of integrating performance information (just 2% of work post 2004), reflecting a likely shift away from cognitive processing approaches in the study of PA/PM. Although there is work in the area of judgment and decision-making that could inform this task (Reisberg, 2006), it is problematic from a systems perspective that it has not been explicitly studied in the PM context. We believe that more research on this task is needed, especially in light of some evidence that performance ratings may reveal more about the rater than the ratee (Murphy, 2008; Scullen, Mount, & Goff, 2000).

Rendering a formal summative performance evaluation. Although PM includes more than just evaluating performance, evaluation is still critical, given that PM "involves managing employee efforts based on *measured* outcomes" (den Hartog et al., 2004: 557). Our review shows that evaluation not only is the most frequently studied task overall (80% of all articles) but also is studied somewhat more often in recent articles than in older articles (84% vs. 74%, respectively). Relevant to the debate over the value of formal evaluations (Adler et al., 2016), there are a handful of empirical studies that explicitly examine the impact of having a formal evaluation, with mixed results: some found it can positively affect organizational climate and job satisfaction (Kaya, Koc, & Topcu, 2010) as well as product quality as an indicator of firm performance (Nayyab, Hamid, Naseer, & Iqbal, 2011); others found no effect of formal evaluation on outcomes such as employee satisfaction, commitment, and turnover (Pullin & Haidar, 2003; C. Zheng, Morrison, & O'Neill, 2006).

Generating and delivering performance feedback. Our review suggests that feedback is the second most common task studied (31% of recent studies). Research shows that only 4% to 10% of the variance in feedback sign is accounted for by the performance rating (Culbertson, Henning, & Payne, 2013), supporting the necessity of studying feedback as a separate task from the evaluation itself. We know from research that feedback *can* be effective for improving performance, assuming it does not threaten one's self-esteem (Kluger & DeNisi, 1996) and is perceived accurately and accepted by the recipient (Kinicki, Prussia, Wu, & McKee-Ryan, 2004; Thurston & McNall, 2010), and that follow-up feedback is an essential part of goal setting and has been meta-analytically shown to improve performance beyond goal setting alone, especially for complex tasks (Neubert, 1998). Yet recent work also suggests that feedback alone may not be sufficient to improve performance via developmental efforts, especially given that much feedback tends to be positive as opposed to focused on suggestions for improvement (van der Leeuw, Overeem, Arah, Heineman, & Lombarts, 2013); active reflection on the feedback is likely required for performance improvement (Anseel, Lievens, & Schollaert, 2009). Recent work has also found that organizations implementing feedback interventions see improved employee reactions and performance (Pampino et al., 2004), including safety behavior (Culig, Dickinson, Lindstrom-Hazel, & Austin, 2008) and job satisfaction (via a positive organizational climate; Kaya et al., 2010). Regarding multisource feedback, meta-analysis has shown such feedback can marginally improve employee performance, with larger effects observed when peers and direct reports are included, when it occurs more than once a year, and when it is done for developmental as opposed to administrative purposes (Smither, London, & Reilly, 2005). In addition, implementation of 360-degree feedback can increase employee perceptions of achievement and support cultures as well as general PA system effectiveness (Mamatoglu, 2008).

Performance review meeting. The performance review meeting component of PM has been examined in 11% of recent empirical studies. This research largely focuses on interactions between manager and employee and frequently applies a fairness (J. A. Colquitt et al., 2013) lens to the meeting and its outcomes, establishing some important justice principles to be followed when such meetings are held. In a notable large-scale, 4-year longitudinal study, Linna, Elovainio, Van den Bos, Kivimäki, Pentti, and Vahtera (2012) were able to isolate participants who had not previously received a formal performance review meeting but received one at Time 2. They found that such meetings positively predicted justice perceptions (procedural and interactional) but only when the meeting was experienced as helpful; when experienced as unhelpful, it actually hurt justice perceptions. Moreover, during times of negative changes in one's work life, useful performance review sessions were especially important in preventing the deterioration of justice perceptions. This research suggests that at least if done well, this task can play an important role in the overall PM system.

Performance coaching. Like others (e.g., Gregory & Levy, 2010, 2011), we see performance coaching as a critical part of PM and yet one that previous reviews have not included, perhaps because research in this area has notably lagged practice (see e.g., Gregory & Levy, 2012; Gregory, Levy, & Jeffers, 2008). There is a distinction between employee coaching and other forms of workplace coaching, such as executive coaching. This task in our model is about employee coaching, which Gregory and Levy (2010) define as that which takes place between a manager and his or her direct reports; it occurs amidst the other day-to-day manager-employee interactions. Heslin, VandeWalle, and Latham (2006) have identified (and developed a scale to measure) three dimensions of such coaching: guidance, facilitation, and inspiration. Much of the research in this area examines only perceptual outcomes (e.g., satisfaction with the coaching); there continues to be a lack of evidence that managerial coaching actually results in improved performance (see, e.g., Gregory & Levy, 2011). Although a recent meta-analysis (Theeboom, Beersma, & van Vianen, 2014) suggests that coaching can be useful for improving employee skills and performance, even with just a few sessions, this is based on coaching by professional coaches. The effectiveness of managers in this role would hinge on their related skills, and as of yet we know very little about managers' abilities and willingness to coach employees (although there is some evidence that managers are particularly weak in this; see Eichinger, Lombardo, Ulrich, & Cannon, 2004).

In summary, our review reveals a great deal of variability in terms of which of the seven PM tasks are the focus of empirical research. The formal summative performance evaluation task component is still the most frequently studied task (and in fact is studied in a higher percentage of articles in recent research compared to earlier research) followed by feedback (which was also more prevalent in recent research compared to earlier research). The only two task components studied less frequently in recent research compared to earlier research are observing employee performance (2% vs. 18%, respectively) and integrating performance information (2% vs. 8%, respectively), which likely reflects a shift away from a cognitive processing perspective. Importantly, this suggests that we as a field possess more knowledge about some PM tasks than others. Given that these tasks operate jointly as part of a system, this is an issue to be addressed.

Explicitly focusing on the tasks of PM in research should be helpful in understanding which components are most important for overall system effectiveness. Some PM practices are believed to be more beneficial than others (see Haines & St-Onge, 2012; Lawler &

McDermott, 2003), yet as Lawler (2003) has noted, historically there has been little clarity about what practices make a PM system effective. Our coding uncovered 13 articles (all post 2004) that empirically examined the impact of one or more of these seven PM tasks. That is, these studies examined whether the *existence* of a particular PM task component (most commonly formal evaluation and feedback) affected outcomes related to individuals and organizations. This research generally suggests (with some mixed results) that each of these tasks can be important in isolation. Yet we lack research that examines these tasks jointly. In fact, our coding did not uncover a single empirical article examining the importance of the existence of more than one of these tasks considered jointly. Thus, a central conclusion from our review is that we lack sufficient evidence regarding how essential each of these PM tasks is (vis-à-vis the other tasks) and how they fit (e.g., interact with one another) within the broader PM system. We return to this issue in the final section of our paper.

The Individuals of PM

In systems models, *individuals* refer to the characteristics and nature of the organizational members (their knowledge, skills, needs, demographics, etc.) who perform the tasks, which are key factors thought to influence their behavior on the tasks (Nadler & Tushman, 1980). In the context of PM, the primary individuals involved have traditionally been raters (especially managers) and ratees (employees). Our coding uncovered 273 studies (49% of empirical PM studies) that tested relationships with one or more variables from the individuals component, making this the most studied component in our model. When breaking down the research by studies examining ratee, rater, or both roles, a focus on ratees has increased (from 42% before 2004 to 50% since 2004), while studies on rater variables have decreased (from 43% to 34%); research considering both roles has remained constant at a relatively lower level (15% to 16%). As such, it appears that researchers have heeded calls to move away from a predominant focus on rater characteristics and cognition (e.g., Ilgen et al., 1993; Levy & Williams, 2004). Below, we summarize the overall conclusions from this extensive research on the individuals of PM. Because of the sheer volume of research in this and some other areas of PM (273 studies on the individuals of PM alone), it is necessary to review this research here at a fairly broad level to set the stage for our systems-based discussion. Yet because many readers may be interested in the details of this recent research, a much more extensive summary can be found in the appendices in the online supplemental material.

This research suggests we have learned much about the importance of several categories of ratee and rater variables within PM (see also Table 2). Relevant ratee variables include performance, which affects both evaluations and feedback (e.g., van der Leeuw et al., 2013); ratee demographic and cultural variables, which affect ratings and preferences for PM (e.g., Ellis, 2012); and ratee individual differences (including feedback orientation; Linderbaum & Levy, 2010), which affect feedback and ratee reactions (e.g., Culbertson et al., 2013; Kuvaas, 2007). Relevant rater variables include individual differences and cultural variables, which affect attitudes toward PM and rating behavior and biases (e.g., Ng, Koh, Ang, Kennedy, & Chan, 2011; Saffie-Robertson & Brutus, 2014); rater abilities and skills, which affect a number of different PM tasks (e.g., Goerke, Möller, Schulz-Hardt, Napiersky, & Frey, 2004; Pesta, Kass, & Dunegan, 2005); and raters' goals and other motivational variables (including self-efficacy; e.g., Ng et al., 2011; Wong & Kwong, 2007). Finally, regarding dyadic variables, we know rater-ratee similarity affects evaluations (e.g., Erez, Schilpzand, Leavitt,

Woolum, & Judge, 2015; Pearce & Xu, 2012) and that several aspects of PM are affected by relational variables such as leader-member exchange and the quality of the coaching relationship (e.g., Elicker, Levy, & Hall, 2006; Gregory & Levy, 2012).

Our review of the research on individuals of PM shows there is plenty of evidence that the characteristics of the organizational members involved in the PM tasks (i.e., managers and employees) exert important effects, primarily because they are key influencers of behaviors on PM tasks, as systems theory would suggest (Nadler & Tushman, 1980). However, research seems to examine only some types of individual variables for some types of PM tasks, at the exclusion of other tasks. For example, we know that ratee abilities, skills, and performance have a substantial impact on the tasks of performance evaluation and feedback, but far fewer studies have considered the importance of these variables for other PM tasks. We also know that ratee individual differences can influence the feedback and performance review meeting tasks, but these variables have not been studied much in the context of other PM tasks. In addition, the examination of ratee and rater demographic variables has been confined almost exclusively to the performance evaluation task. Raters' abilities and skills is the one category of individual variables that has been examined more broadly across tasks, with research showing these variables are important factors for all of the various PM tasks. Coaching in particular appears to be a task area ripe for studying additional manager and employee individual variables. Heslin et al. discussed the importance of examining supervisor individual differences with regard to coaching, noting that the "stark reality is that managers often differ substantially in their inclination to coach their subordinates" (2006: 871). Thus, much variation in informal coaching behaviors as well as adherence to formal coaching processes set by the organization may be present as a result of manager variables from this component of our model. Additional research should also address the influence of individual variables on the conflict that likely ensues when a manager takes on the roles of both rater and coach (see Cleveland, Murphy, & Williams, 1989). Finally, despite the fact that Aguinis's (2013) definition of PM references performance of teams as well as individuals, there is a striking absence of empirical work on team-based PM. Thus, we wish to identify this as an important area for future empirical work.

The Formal and Informal Processes of PM

Our model, adopted from Nadler and Tushman's (1980) approach, recognizes *formal processes* and *informal processes* as two distinct components of a system. Formal processes refer to structures, processes, and procedures that are explicitly developed to get individuals to perform (PM) tasks consistent with (PM) strategy; these are often documented in writing (Nadler & Tushman, 1980). In contrast, informal processes are unwritten, implicit, and emerge over time; they reflect PM-related processes themselves (e.g., presence of informal feedback) as well as contextual factors that affect processes (e.g., political climate surrounding PM). Because our review suggests that this formal/informal distinction has not always been articulated clearly in the literature, conceptually or methodologically, we discuss formal and informal processes together.

Formal process was the second most studied component in PM, with 235 studies (43% of studies). Informal processes were examined less often (176 studies, or 32%); however, we saw informal process research slightly increasing over time (53% post 2004), whereas formal process research is decreasing (only 40% post 2004). We also saw a promising shift where

more recent articles emphasize the importance of informal process above and beyond or in concert with formal system elements (e.g., Hofstetter & Harpaz, 2015; Tuytens & Devos, 2012a); for example, both supportive HR practices and supportive supervisors are needed to obtain positive outcomes from PM (Shahnawaz & Juyal, 2006). Below, we briefly summarize the overall conclusions from this extensive research on the formal and informal processes of PM; a much more detailed discussion of this research can be found in Appendix B in the online supplemental material.

Our review suggests that research on formal process relating to rating approach/format has substantially decreased, with the exception of forced distribution rating systems (FDRS) studies, which reveal some positive and several negative effects of FDRS (e.g., Lawler, 2003; Schleicher, Bull, & Green, 2009). Rater training research continues, with both traditional (e.g., frame of reference; Gorman & Rentsch, 2009) and newer (e.g., "whole brain"; Selden & Sowa, 2011) forms showing promise for improving PM. Studies on rater source and multisource ratings remain prevalent, with mixed support for the value of multiple raters in PM (e.g., Lawler & McDermott, 2003; Woehr, Sheehan, & Bennett, 2005). Use of technology in PM has begun to be investigated, with results showing some benefits of online systems and no differences or even disadvantages in some other areas (Kurtzberg, Naquin, & Belkin, 2005; Payne, Horner, Boswell, Schroeder, & Stine-Cheyne, 2009). Both rater accountability (e.g., Curtis, Harvey, & Ravden, 2005; Mero, Guidice, & Brownlee, 2007) and ratee participation (e.g., Inderrieden, Allen, & Keaveny, 2004; Kavanagh, Bensen, & Brown, 2007) have been shown to be important variables that can be enhanced through both formal and informal processes. Due process characteristics of PM (Levy, Cavanaugh, Frantz, & Borden, 2015), which contribute to fairness/justice perceptions, can also be found in formal and informal forms (e.g., Kataoka, Cole, & Flint, 2006, describe formal goal setting and informal continuous feedback combining as part of providing adequate notice). Feedback also occurs in both formal and informal forms, and richer feedback (in terms of timeliness, specificity, and frequency) is linked with positive outcomes (e.g., Northcraft, Schmidt, & Ashford, 2011; Selvarajan & Cloninger, 2012). The feedback environment specifically has emerged as important and represents a fully informal process element consisting of an employee's feedback interactions with his or her supervisor and coworkers (Steelman, Levy, & Snell, 2004). Finally, politics surrounding PM, another informal element, is associated with multiple negative outcomes (e.g., Poon, 2004).

In general, our review suggests that the various formal and informal process elements of PM differ in how consistently they affect PM effectiveness. Employee participation is one factor that appears unequivocally beneficial, showing positive results across both formal and informal forms of participation and in relation to multiple different PM tasks. Rater training also consistently produces positive outcomes, and one study found a 194% return on investment for the organization's PM training program (Goh & Anderson, 2007). Yet for rating format and approach, the mixed results across studies suggest that this is an area where "best practices" are likely to be elusive; rather, as we discuss in the context of systems theory principles below, the question of which evaluation format and approach is best is likely to depend on the specifics of the particular context.

Our review also suggests that research should acknowledge that PM processes can simultaneously exist in both forms (formal and informal). Farndale, Hope-Hailey, and Kelliher (2011) used a framework by Nishii and Wright (2008) of three lenses through which HR practices (including PM) can be viewed: (a) the practice *intended by HR* as embedded in

policy documents, (b) the practice as *enacted by different managers* throughout the organization, and (c) the *employee's experience* of the practice irrespective of intentions of his or her manager and the HR department. The first lens represents formal processes in our model, whereas the second and third lenses represent informal processes. We would encourage future research to adopt this multiple lens model in seeking to understand the effects of various process elements of PM. This framework also serves to highlight the manager as the central enactor of PM (see also den Hartog et al., 2004), acting as an interpreter of formal practices and influencing informal processes. As such, we would encourage future PM research to focus on the manager as a key participant (our review actually showed that formal and informal process research was somewhat *less* likely to focus on managers as participants compared to PM research overall). Related, our review suggests *compliance* with various formal PM policies and procedures is another interesting and understudied aspect and one that can likely be explained by many informal factors in the organization. In support of this, Anderson and Johnson (2005) found that the context influenced PA compliance more so than organizational mandates. We observed an underlying assumption in many studies that supervisors actually enact PM as dictated by the organization, but this is an untested assumption. Managerial compliance should be focused on as a substantive variable in future PM research, especially given that it likely serves as a key linking mechanism between formal and informal processes.

The Inputs of PM

Inputs are the "givens" of the system: the materials the organization has to work with and the context in which it conducts its work; as such, they place demands and constraints on systems (Nadler & Tushman, 1980). Inputs include the environment (factors outside the PM system or the organization itself), resources, and strategy (organizational strategy, HR strategy, and in this context, PM strategy; Collins-Camargo, Chuang, McBeath, & Bunger, 2014). For some time the literature has acknowledged that such variables would be important for understanding PM (e.g., Bretz et al., 1992; Ilgen et al., 1993; Klein, Snell, & Wexley, 1987; Murphy & Cleveland, 1991) while simultaneously observing that very little research exists on them. Levy and Williams labeled such factors as "distal variables," which they note "are broadly construed as contextual factors that affect many human resource systems, including [PA]" (2004: 885). They concluded that these variables continue to receive very little research attention, other than a small number of studies on the impact of culture, climate, and technology in PA (e.g., Hebert & Vorauer, 2003; Miller, 2003). Fortunately, our review suggests substantially increased attention to the role of some of these variables in PM research over the past 13 years. Specifically, our coding revealed 110 studies (20% of those in our review) that have empirically examined the effect of one or more variables that we would classify as system inputs. Whereas 44 (40%) of these studies were published in 2003 or before, 66 (60%) of them were published in 2004 or later, suggesting that recent research has indeed increased its focus on PM inputs. Below, we summarize the overall conclusions from this research; a more extensive and detailed discussion can be found in Appendix C in the online supplemental material.

Our review indicates that research has examined PM across multiple industries and sectors and in the context of other HR practices (all considered inputs). It has also studied a number of other types of variables classified as inputs, including ownership type and organizational

structure (e.g., Abu-Doleh & Weir, 2007); national culture (e.g., Peretz & Fried, 2012); organizational culture and climate (e.g., Farndale et al., 2011); resources, including leadership support (e.g., Wood & Marshall, 2008); and organizational strategy and PM strategy (e.g., Ayers, 2013). From this research, we can conclude the following. First, PM system inputs matter; they affect both the other components of PM (e.g., tasks, processes) and its ultimate effectiveness. Our review shows there is plenty of empirical support for this, confirming that inputs act as both demands and constraints on multiple aspects of the PM system. Second, it is clear that PM can be effective across different contexts but that the specifics of PM do and should vary across contexts. National culture in particular is a strong influence on PM (e.g., Peretz & Fried, 2012), as is the organization culture (e.g., Haines & St-Onge, 2012), especially as a moderator that constrains or enables various aspects of PM. Alignment between the PM system and both national culture and organizational culture has been shown empirically to be important (Peretz & Fried, 2012). Less clear empirically is the role of organizational strategy as an input. Theoretically, this should have sweeping implications for the PM system; indeed, PM is often seen as a tool through which an organization implements its strategy (Aguinis, 2013). However, our review suggests that these conclusions are in fact more theoretical than they are empirically based. Across the few studies that empirically examined alignment between organizational strategy and PM, the results have been mixed (see Appendix C in the online supplemental material). Thus, it appears that the empirical evidence for the importance of organizational strategy in PM is actually quite thin. PM strategy (including the purpose of PM), on the other hand, appears unequivocally important, both theoretically and empirically (e.g., Ayers, 2013), and recent research has identified a critical distinction between intended (from the perspective of management or the organization) and perceived and experienced (from the perspective of employees) purposes of PM (e.g., Krats & Brown, 2013; Youngcourt, Leiva, & Jones, 2007). We have more to say about the role of PM purpose when we discuss systems principles in the final section below.

Other HR practices are also an important input, and PM should be studied alongside these to assess the role of PM on outcomes over and above these other practices. We found more than a dozen articles since 2004 that have taken such an approach (see Appendix C in the supplemental material). About one third found that the effects of PM were not significant in the context of other HR practices such as recruitment, selection, compensation/reward, training and development, and job design. However, we also observed that in many of these (and other) studies, the PM variables were operationalized in very limited ways (e.g., dichotomous existence, percentage of employees affected), which may at least partially explain these results. Those articles finding the opposite, that PM was significantly related to outcomes in the context of other HR practices, were more likely to be employee- (rather than organizational-) level research (with some exceptions). This research also suggests that other HR practices can moderate the effects of PM, but we need more work examining such interactions.

Overall, we believe it is a positive sign that research on PM inputs has increased substantially since the last review (Levy & Williams, 2004). Yet there is still less work on inputs than any other PM component; as such, we encourage continued research in this area. Especially needed are examinations of more "macro" inputs, such as economic conditions, which can place demands and constraints on PM. Unlike some HR practices, economic crises seem to motivate an increasing interest in PM (Cabrera, Fernaud, Díaz, Rodríguez, Vilela, & Sánchez, 2014). In higher education, for example, decreased resources resulting from less government funding have increased adoption of PM in attempts to increase efficiency and effectiveness

(Decramer, Smolders, & Vanderstraeten, 2013). More research is needed on the effects of these and other macro PM inputs.

The Outputs of PM

The final component of our model of PM, outputs, refers to what the system produces (from the workflow of PM tasks). Typical outputs of PM in research and practice include *performance rating(s)*, the *feedback generated and delivered* (e.g., feedback valence or favorability, specific content, which could be evaluated along multiple dimensions such as agreement or perceived accuracy, usefulness, richness), the creation of a *development plan* or other *performance improvement plan, career planning, recommendations regarding administrative decisions* (including rewards, promotion, training, termination, etc.), and *documentation for legal purposes*.

In our review, 41% of studies included the empirical evaluation/measurement of one or more of the above outputs. As might be expected, research has not paid equal attention to all of these outputs; performance ratings in particular have been described as receiving "disproportionate focus" (Adler et al., 2016). Indeed, we found that a full 76% of output studies examined performance ratings; the next most common output was recommendations regarding administrative decisions, especially pay (15% of studies), followed by feedback (6%), career planning (3%), and development plans (2%). Performance ratings continue to be evaluated largely along psychometric criteria (this is true of both older and more recent research), including the level of rating (i.e., elevation or leniency), rating accuracy, and congruence/agreement between raters. Thus, not only does our review suggest a disproportionate focus on this particular output (i.e., performance ratings) but it is also the case that when this output is examined, the focus is disproportionately on psychometric criteria specifically. Missing is research examining how useful ratings are for informing developmental plans or other administrative decisions, for example; this suggests the importance of studying relationships among outputs in understanding PM effectiveness. This is especially critical in light of claims related to the limited usefulness of performance ratings (e.g., Murphy, 2008; Scullen et al., 2000), claims based exclusively on psychometric criteria.

In addition, the majority of research looking at outputs of PM other than performance ratings tends to link merely the existence of these outputs with reactions to the PM system. For example, research has found that receiving positive feedback as an output of PM affects overall reactions to the PA system (e.g., Jackson, 1996) as well as evaluations of the supervisor (e.g., Kacmar, Wayne, & Wright, 2009); when goal establishment is an output of PM, the feedback system is seen as more fair (Evans & McShane, 1988); and when salary increases are an output of PM, FDRSs are seen as more effective (Lawler, 2003). However, this research does not specifically measure how good each of these outputs is (e.g., the quality of the feedback or the goals established, or the fairness or validity of the salary increases) in its own right. We discuss this in greater detail below.

A Systems Theory–Based Integration of PM Research

Our goal with this review has been updating and extending prior reviews by explicitly focusing on a systems-based model of PM. This conceptual framework is useful for distilling knowledge about PM from the extant literature and for identifying additional issues that

should be researched. The following sections provide a more in-depth and integrative consideration of how systems theory and its principles change what we do and can know about PM (including the three recent assertions we mentioned in the introduction: informal aspects of PM are more important than formal aspects, Pulakos & O'Leary, 2011; performance ratings are inaccurate and not particularly useful and therefore should be done away with, Adler et al., 2016; and PM processes should be streamlined to remove their "low value" aspects, Effron & Ort, 2010). This systems theory discussion, along with the component information summarized above, is distilled into a list of key recommendations for future research provided in Table 3.

A Systems-Based Taxonomy of PM

In introducing their systems model, Nadler and Tushman (1980) noted that the pragmatic value of such a model is in identifying which factors (out of a possibly infinite set) are most critical or important for understanding the functioning of the system. The need for this in the PM literature has been previously noted (Lawler, 2003), and it has become more pronounced over time, as we have evolved from PA to the broader area of PM and as the relevant literature continues to grow at an exponential rate. At the broadest level, our model shows that understanding PM is a function of six critical components: inputs, tasks, individuals, formal processes, informal processes, and outputs. Yet fully delivering on the promise of such a taxonomy requires filling in this broader framework with more specific subcategories and variables from the PM literature. Accordingly, we created the more detailed taxonomy of the relevant categories of PM variables shown in Table 2.

This taxonomic framework has proven useful for organizing the truly voluminous research on the many moving parts of PM, and our review based on this framework has provided important summaries regarding cumulative knowledge in these component areas of PM. Our overarching conclusions here are that both theory and empirical research clearly demonstrate that each of the six system components exerts a significant influence on PM. In addition, across components, it appears that research has overemphasized the examination of individual variables (most likely owing to the decidedly "micro" focus of early PA and PM research), and it has historically underemphasized the examination of inputs, which are crucially important from a systems theory perspective in general (as they provide resources and place both demands and constraints on the system; Nadler & Tushman, 1980) and when considering the changing organizational context of PM specifically. On the positive side, recent research has witnessed a significant increase in the study of inputs; indeed, it is this system component that saw the greatest change in focus over time. In addition, research has certainly broadened in recent years to examine multiple tasks identified as part of our PM system model, including feedback, coaching, and the setting of objectives, not just performance evaluation; yet performance evaluation remains the most frequently studied task. Thus, it is not just that more recent research has shifted task focus but that it tends to include a *broader* set of tasks. This is in keeping with a systems-based view of PM, which identifies multiple integrated tasks composing the essential workflow of the system. Such a view also suggests that questions about the necessity of formal performance ratings (Adler et al., 2016) are best considered in conjunction with the other tasks of PM, not in isolation, an idea that we develop further below.

Table 3
Recommended Directions for Future Research on Performance Management

Focus	Research Needs and Sample Research Questions
Tasks of PM	● More research, in a PM context, is needed for certain tasks, including setting performance expectations, observing employee performance (e.g., field studies on diary keeping with actual managers to understand how to make this practical and manageable), integrating performance information, and coaching (e.g., Does coaching actually improve employee performance? What affects managers' abilities and willingness to coach employees?). ● We need a better understanding, in general, of how busy managers accomplish all these various PM tasks. ● Research should directly test how essential each task is to the effectiveness of PM overall (e.g., Are formal review meetings even necessary? Do they do more harm than good?). ● How do the tasks of PM interact with one another and other system components to influence PM effectiveness? Does each PM task provide incremental value beyond other tasks? How do multiple PM tasks work together to bring about performance improvement? How does the relative effectiveness of multiple PM tasks vary as a function of contextual elements? What is the impact on other tasks of, for example, removing formal performance ratings?
Individuals of PM	● Individuals research needs to move out of laboratory contexts and beyond student participants; it also needs to collect dyadic data (i.e., employee-manager dyads). ● Individual variables should be examined in relation to all of the separate PM tasks (e.g., the influence of ratee individual differences has not been studied much beyond the feedback and review meeting tasks; rater and ratee demographic variables have been confined to the performance evaluation task; manager individual differences need to be studied in coaching task). ● We need more research on the implications of holding multiple roles (e.g., rater and ratee; rater and coach). ● We need significantly more research on PM with teams.
Formal and Informal Processes	● Field research is needed to generalize lab results for rater accountability and ratee participation. ● Process research should examine managers as key respondents (given their role as interpreters of formal processes and significant sources of informal processes). Focus on managerial compliance as a substantive variable in PM process research. ● Researchers should note explicitly whether they are taking a formal or informal approach to a particular process variable (conceptually and operationally). ● Mixed results for rating format suggests it depends on the specifics of the particular context and should be studied interactively with context variables. ● We need to simultaneously examine both formal and informal processes within the same study: the relative importance of each, how these two forms of process interact, and how they mutually influence one another. Both direct and moderating effects should be examined in such models. Studies should examine the trade-offs that organizations might make (e.g., Can rich informal feedback substitute for formal feedback during a performance review meeting?). ● Apply Nishii and Wright's (2008) multiple lens model in studying the effects of various process elements of the PM system.
Inputs of PM	● There needs to be continued and increased research on inputs of PM (especially more "macro" inputs). ● Additional research on the role of organizational strategy is needed, given equivocal results. ● We need to consider how some input variables might be meaningfully manipulated in laboratory contexts (see Lievens, Conway, & De Corte, 2008) to supplement the need for multicontext field research. ● Researchers must describe the context of their studies along key contextual (input) variables, including national culture, industry, type of firm, organizational strategy and culture, and especially purpose of the PM system. ● PM must be studied in the context of other HR practices (and vice versa) to help isolate the role of PM on individual and organizational outcomes over and above these other practices. Also, how do other HR practices moderate the effects of PM?
Outputs of PM	● Research must begin to focus on system outputs other than the performance rating, including the quality of administrative decisions, feedback, career planning, and development plans that result from the system (e.g., the quality of the feedback or goals established, the fairness or validity of resulting salary increases). ● Relationships among output variables should be studied to better understand PM effectiveness. Research should examine aspects of performance ratings beyond psychometric criteria, including how useful they are for informing development plans or other administrative decisions.

(continued)

Table 3 (continued)

Focus	Research Needs and Sample Research Questions
Congruence Hypothesis	• Research should explicitly focus on the congruence between multiple aspects of the PM system as a predictor of system effectiveness. • The fit between inputs (especially PM strategy/purpose) and other system components is particularly important to examine. We need significantly more research on how various purposes of PM (beyond simply administrative vs. developmental) interact with the other components of PM systems to determine congruence and ultimately effectiveness.
Internal Interdependence Principle	• Research questions should involve the simultaneous examination of multiple PM system components. • Research should examine the interdependencies (and conflicts) that likely exist when individuals hold multiple roles within PM. • Regarding interdependence of tasks, does removing formal performance ratings, for example, place more or less of a burden on the feedback and coaching tasks? • Until this type of research is conducted, we recommend a moratorium on calls to do away with any tasks of PM.
Capacity for Feedback Principle	• Future research needs to focus on identifying what PM outputs (and what evaluative dimensions of these outputs) are most relevant for understanding different questions about PM system effectiveness. • Research should examine aspects of performance ratings beyond psychometric criteria, including how useful they are for informing development plans or other administrative decisions. This suggests a need to study relationships among output variables to better understand PM effectiveness.
Equifinality Principle	• Rather than focusing on single best ways, future research could more productively focus on identifying the various configurations of systems components that lead to a particular desired system output, including the trade-offs inherent in deciding amongst these various paths. These suggestions represent significant departures from traditional PM research.
Adaptation Principle	• We need research on how PM systems do and should adapt (i.e., change) to changing inputs such as economic conditions and the changing nature/context of work.
General Methodological Implications	• "Midrange" categories of variables important to PM (see taxonomy in Table 2) could serve as foundations for future meta-analyses to build more conclusive and cumulative scientific knowledge about PM. • Understanding the extensive contingencies within PM necessitates designs that measure variables from multiple system components at once and test interactions. • We need more research in the field (within PM systems) and with employee and manager participants. • Researchers should increase the methodological rigor of PM research (especially more macro PM research); if cross-sectional surveys are necessary, consider multiple data sources (see Jiang, Wang, & Zhao, 2012). • Beyond sampling at the PM system level, variance in many PM system components can often be found across work groups, departments, or divisions, and it may be more feasible to collect data at this level. • There is a need for better measures of PM systems themselves, beyond dichotomous and percent of employees affected. The detailed taxonomy in Table 2 should prove helpful in this regard.

Note: PM = performance management; HR = human resource.

Related to the contention that informal aspects of PM are more important than formal (Pulakos & O'Leary, 2011), our review suggests this distinction is often muddied in the research literature. Our taxonomy (see Table 2) can help clarify this, paving the way for future research that can more definitively answer this. We found several examples in the literature of research that essentially takes a "mixed" approach without explicitly acknowledging it and, for example, uses a measure that combines both formal and informal elements of process to examine a single variable. Given the different underlying mechanisms and empirical patterns for these two types of variables (see Appendix B in the supplemental material), this practice obfuscates understanding and hinders cumulative knowledge. Moreover, especially for factors

that can be represented in both formal and informal forms (e.g., employee participation in PM), understanding the varying relationships based on form is likely to be important for informing practice. Thus, we believe it is essential that PM researchers begin to note explicitly whether they are taking a formal or an informal approach (conceptually and operationally) to a particular variable and the implications for interpreting their results.

We believe our taxonomy can play an ongoing role in building and integrating cumulative and conclusive scientific knowledge in PM. Effron and Ort (2010) have suggested that any element of PM that cannot be justified based on conclusive scientific knowledge should be removed (with the goal of streamlining unnecessarily burdensome systems). Kepes, Bennett, and McDaniel (2014) similarly discuss how essential such cumulative knowledge is for the practice of evidence-based management. It has struck us in our review of the extensive PM literature that applying such a criterion (conclusive and cumulative scientific knowledge) would not leave a great deal of specific PM elements intact (with the likely exception of the established importance of ratee/employee participation and rater/manager accountability). Indeed, Buckingham and Goodall (2015) discuss how Deloitte recently reached this same conclusion after reviewing the literature and greatly streamlined their PM processes as a result. This is clearly not because as a field we *lack research* on PM. As our articles gathered for this review (which numbered over 1,900) can attest to, there has been plenty of PM research conducted, written about, and published in the past 36 years alone. Yet our review reveals that the resulting scientific knowledge about PM tends to be at the level of either very specific findings from numerous "one-off" studies or very broad conclusions. The current taxonomy should prove very useful in identifying "midrange" categories of variables important for PM (see Table 2), which can then serve as foundations for future meta-analyses.

Meta-analyses in our field serve as both a method and a metric for ascertaining cumulative scientific knowledge; in fact, Kepes et al. (2014) place meta-analysis at the top of their hierarchy of evidence. Regarding their purpose as a metric for cumulative scientific knowledge, meta-analyses are generally less feasible and less prevalent to the extent that areas are not characterized by agreed-upon taxonomies of variables. To test this, we gathered some comparative data regarding meta-analyses in PM versus other topics in the organizational sciences. In the PM literature specifically, despite a very large number of empirical articles, there are relatively few meta-analyses. The ratio of meta-analyses to articles in PM is only .01 (about 1 meta-analysis for every 100 empirical articles), which is *5 times smaller* than the comparable ratios for cognitive ability and personality, literatures that have clearly benefitted from more established taxonomies (and have also benefitted from meta-analysis). We are hopeful that the systems-based taxonomy we created and used to organize this review will be helpful in categorizing both extant and future research to allow for more meta-analyses and, consequently, greater cumulative knowledge about PM. It is also time that we build more conclusive knowledge by moving beyond questions about PM components in isolation. Doing so requires consideration of other aspects of systems theory, as we detail next.

The Congruence Hypothesis

A key assumption underlying the Nadler and Tushman (1980) model is that the various components of a system exist in states of relative balance, consistency, or "fit" with each other. They argue that the effectiveness of the system is based on the quality of this congruence, and their *congruence hypothesis* specifically states that the greater the total degree of

congruence or fit between components of the system, the more effective will be the system. So an important question is What can we conclude from our review about congruence amongst system components in PM?

At the most general level, there is much research evidence that a particular PM approach may be good (i.e., effective) for a particular organization operating in a particular context and for a particular set of purposes, yet what works or is a key requirement of PM in one context may not be so in another context (as the congruence hypothesis would predict). Indeed, the presence of extensive contingencies in PM calls into question the very notion of "best practices." This is an important point, as researchers and practitioners continue to imply that a listing of simple best practices for PM is possible (e.g., focusing on informal aspects is more important than formal, Pulakos & O'Leary, 2011; identifying "low value" aspects of PM that can be removed, Effron & Ort, 2010). It is possible to build a list of evidence-based effective PM practices, but it will not be a simplistic list, and to do so we clearly need more research that examines the fit between inputs and other systems components, as well as interactions amongst system components in general (we provide more detail on this need in the Internal Interdependence section below). Related, our review reveals that "congruence" itself is not extensively examined in the PM literature, with some notable exceptions (e.g., Peretz & Fried, 2012). We encourage future research to explicitly focus on the congruence between multiple aspects of the PM system as a predictor of system effectiveness.

There are also important congruence-related implications stemming from our review of specific PM components. First, some tasks of PM have been studied in other areas yet are underrepresented in the PM literature specifically; for example, setting performance expectations has been extensively studied in the motivation literature but not in PM, and integrating performance information has been studied in judgment and decision-making but not extensively in PM. Although one might argue that knowledge from these ancillary areas has been imported to understand PM, if congruence amongst tasks and between tasks and other system components is important for PM effectiveness, we cannot simply apply what is known from other literatures that study these components in isolation. We must instead study these tasks in the context of PM specifically and in concert with other tasks to assess congruence. Thus, we would strongly encourage additional PM-specific research on these tasks that have been relatively understudied in the recent PM literature, namely, setting performance expectations and observing and integrating performance information.

Second, tasks must be performed consistent with strategy. This is a fundamental principle of Nadler and Tushman's (1980) congruence model, and it suggests the particular necessity of studying fit between various PM tasks (and the formal and informal processes governing how these tasks are done) and input variables such as strategy. Although our review showed equivocal empirical support for the importance of overall organizational strategy, perhaps more important in this regard is the *PM strategy* of the organization (e.g., Ayers, 2013). PM strategy includes the philosophies and assumptions and especially the various purposes for which PM is conducted, and we argue that the fit between PM purpose(s) and other system components is particularly critical. Interestingly, prior research has often examined the purpose of PM as merely an instructional set (e.g., rating performance for administrative vs. developmental reasons), often in lab settings. We believe that purpose should be seen not merely as one of many decisions made with regard to the specifics of PM process, as implied by much of the research in this area. Rather, we see PM purpose as playing a much broader

and more far-reaching role than simply a procedural decision, more appropriately conceptualized as an input to the entire PM system and specifically as a crucial part of the PM strategy. Borrowing from Nadler and Tushman, we note that purpose, as part of the PM strategy, "is critical because it determines the [PM] work to be performed by the organization [the tasks] and it defines desired [PM] outputs" (43), for example, administrative decisions, development plans, career planning, and documentation. As such, we need a great deal more research on how various PM purposes (beyond simply administrative vs. developmental) interact with the other components of PM systems to determine congruence and ultimately effectiveness.

Other Relevant Systems Theory Principles

Systems theory includes a number of other specific principles that together emphasize a much more complex and dynamic view of PM than has typically been addressed in the extant research. We do not include an exhaustive list of systems theory principles here (see Katz & Kahn, 1978) but rather discuss those with the most significant implications for PM research.

Internal interdependence. This principle emphasizes the interconnectedness and interdependence of system components, such that any one component can and does affect other components. This has several implications both for interpreting the results of our review and for future research. First, as our review shows, although there is research demonstrating that the various PM tasks are important in isolation, there is little evidence regarding how essential each task is vis-à-vis the others or, particularly importantly, how they interact with one another and within the broader PM system. In fact, there was not a single study that empirically examined more than one task as variables (i.e., the absence/presence of the task). For both theoretical and practical reasons, future research should be directed towards establishing, for example, whether one PM task provides incremental value beyond other tasks, how multiple PM tasks work together in bringing about performance improvement, and how the relative effectiveness of multiple PM tasks varies as a function of context. It would be interesting to examine, given the recent discussions about the value of formal versus informal approaches to PM (Pulakos & O'Leary, 2011), whether removing formal performance ratings (Adler et al., 2016) places more or less of a burden on the feedback task (e.g., What is the strain on the feedback task from removing formal ratings? Does a system then require more or less feedback and coaching?). Importantly, in lieu of this type of systems-based research that jointly examines multiple PM tasks, we believe strongly that any calls to do away with any of these tasks of PM (e.g., formal performance ratings; Adler et al., 2016) are premature.

Second, regarding individuals of PM, historically raters and ratees have been seen as two distinct and independent groups. However, viewing PM as a more complex and integrated system suggests the need for a more nuanced approach to the roles involved, with a number of implications for future PM research on individuals. For example, individuals holding a rater role are almost always also ratees themselves; thus, these roles are not mutually exclusive. In support of this, Latham, Budworth, Yanar, and Whyte (2008) found that managers' own performance rating served as an anchor for the ratings they provided to their subordinates, indicating that a person's experience as a ratee affects how he or she enacts the role of rater. Relatedly, how one reacts to the PM system as a ratee can be influenced by the additional system knowledge the person obtains from also holding the role of rater (J. R. Williams

& Levy, 2000). Other interdependencies likely exist when an individual holds both of these roles, and we see this as an important area for future research to pursue.

Third, this interdependence principle has implications for Pulakos and O'Leary's (2011) contention that informal aspects of PM are more important than formal aspects. To best address this assertion, PM researchers would need to simultaneously examine both formal and informal processes within the same study, and recent research has begun to do this. Rather than suggesting that informal processes are more important for PM effectiveness, however, our review of this work shows that formal and informal processes are likely to interact in many ways. For example, informal factors (e.g., culture) can serve to mitigate or exacerbate the effects of formal factors (e.g., rating approach) on PM outcomes (in this regard it is interesting that informal process variables are studied as moderators more frequently than any other component of our model). The effectiveness of formal processes is likely to vary based on informal factors and vice versa. Thus, our review suggests that formal and informal elements of PM are best understood in conjunction with one another and, in turn, that the Pulakos and O'Leary contention is premature. This is particularly true when one further considers that there is insufficient research that pits these two categories against one another and that the formal/informal distinction itself has not been articulated very clearly in the extant research (conceptually or methodologically). Accordingly, research that simultaneously considers informal and formal PM processes must continue in order to build additional knowledge about the relative importance of each, how these two forms of process interact, and how they mutually influence one another. Future research should consider both direct and moderating effects in models including both formal and informal process variables.

Fourth, regarding PM outputs, we have learned from research that the performance rating output specifically is a complex function of multiple aspects of PM systems. For example, ratings are a function of not only ratee performance mean, trend, and variation; performance history also interacts with appraisal purpose (Reb & Greguras, 2010; an input) and multiple aspects of the rater (individual factors). In addition, we now know that raters use different tactics (inflation/deflation) to achieve certain rating goals (e.g., identification, harmony, fairness motivation; Wang, Wong, & Kwong, 2010), suggesting that the quality of rating outputs can actually shed light on raters' goals. Furthermore, raters' own previous appraisals (an output of their own PM process) can serve as anchors for the appraisals they conduct of others (Latham et al., 2008). Each of these findings suggests a great deal of interdependence amongst system components in determining even a relatively simple output such as performance ratings, highlighting the necessity of adopting multivariate approaches in research on this and other PM outputs.

Capacity for feedback. With this principle, outputs of the system are seen as providing key information (or feedback) about how well the system operates (Nadler & Tushman, 1980). Thus, information about the (PM) outputs can be used to control or correct the (PM) system (Katz & Kahn, 1978). As such, identifying and evaluating specific PM outputs becomes essential in understanding how well components of the PM system are operating.

Our review suggests that the majority of research looking at outputs of PM other than performance ratings tends to link the *existence* of these various outputs with reactions or evaluations of the PM system. Research has found, for example, that (a) receiving positive

feedback as an output of PM affects overall reactions to the PA system (e.g., Jackson, 1996) as well as evaluations of the supervisor (e.g., Kacmar et al., 2009); (b) when goal establishment is an output of PM, the feedback system is seen as more fair (Evans & McShane, 1988); and (c) when salary increases are an output of PM, FDRSs are seen as more effective (Lawler, 2003). However, this research does not specifically measure how good each of these outputs is (e.g., the quality of the feedback or the goals established, or the fairness or validity of the salary increases) in its own right. This is needed in the research literature and in practice in order for the examination of these outputs to deliver on their *capacity for feedback* and self-correcting potential.

Equifinality. The systems principle of equifinality suggests that different configurations of various system components can lead to the same output or outcome. As Katz and Kahn observed, "The general principle, which characterizes all open systems, is that there does not have to be a single method for achieving an objective" (1978: 171). "A system," according to Katz and Kahn, "can reach the same final state from differing initial conditions and by a variety of paths" (170). In the context of PM, this principle suggests that if improved employee performance, or appropriate administrative decisions, or increased firm performance, for example, are the desired outcomes of the PM system, there is no "one best" or universal way to get there. Instead, there are multiple paths, including alternative combinations of different PM tasks as well as a variety of formal and informal processes. These are very interesting possibilities that represent a significant departure from typical PM research. This principle would suggest, for example, that rather than informal processes being more important than formal processes (Pulakos & O'Leary, 2011), different configurations of formal and informal processes can lead to similar outcomes. Thus, future research might be more productively focused on identifying these various configurations. Studies could examine trade-offs that organizations would make in deciding amongst these various paths (e.g., Can rich informal feedback substitute for formal feedback during a review meeting?).

Adaptation. The principle of adaptation argues that system inputs and outputs must be maintained at a favorable balance with the environment. Thus, PM systems must adapt to changing environmental conditions (inputs) to maintain their effectiveness. However, there has been no substantive empirical research on how PM systems do and should adapt (i.e., how they change) in response to changing environmental conditions, such as economic conditions (Cabrera et al., 2014), or the changing context of work (e.g., Godin, 2009; Howard, 1995). Interestingly, as Alan Colquitt (2017) illustrates in his recent book on PM, the historical form of adaptation of PM in organizations has been an additive one. PM systems have evolved over time to simultaneously address several different purposes, all reflecting varying historical influences (e.g., more objectively measuring performance to provide a better basis for promotion decisions; legal defensibility; management by objectives; focus on planning future performance). As A. Colquitt eloquently notes, "All of these influences were additive, like barnacles attaching themselves to a log, one on top of the other" (10). The result is an underlying and inherent challenge with PM systems in many organizations: we simply ask them to do too many things. There are likely to be much more functional approaches to adaptation of PM systems, and we need future PM research to direct its efforts in this regard.

Methodological Considerations and Implications

We conclude our review by identifying some key methodological recommendations for future PM research that stem from our systems-based model. First, and arguably most important, understanding the extensive contingencies within PM necessitates research designs that measure variables from multiple system components at once and test interactions. As we noted in our summaries, this is not the norm in the extant PM research. Research on individual factors was the most likely to do this, with 25% of those studies empirically testing interactions between individual factors and other PM system components. Understanding the effectiveness of PM systems requires this type of multivariate interactional research, and the field needs more of it.

Second, progress in understanding PM *systems* (as opposed to just components of PM) likely necessitates more research being conducted in the field and with employee and managerial participants. As Table 1 shows, study context (lab vs. field) and participant type (employee or manager vs. student) varied substantially across components of our model. For example, input and informal process studies were more likely (than other areas) to be conducted in the field and less likely to be lab studies; informal process studies were also more likely to engage employees as participants and less likely to use students. A comparatively larger proportion of research on individuals (38%) occurred in a lab setting; and 25% of the individuals studies used students as participants. Although it is admittedly easier to collect data on individual variables in the lab, it seems difficult to fully recreate the complexities and nuances of individual roles involved in PM in a laboratory context and with students as participants. We also found that only 5% of individuals research examined employee-manager dyads. Given the importance of this relationship for PM (e.g., Dusterhoff, Cunningham, & MacGregor, 2014), we believe more dyadic research is needed. Perhaps most troubling, we also found that for topics where more definitive conclusions could be reached (i.e., rater accountability and ratee participation), nearly all of the studies were conducted in the lab. Thus, much more field research is needed to test the generalizability of these results.

Third, although this varied across system components, there is significantly more PM research that exists at the organization level (wherein organizations, not individuals, are the unit of analysis) in the recent literature compared to the older literature. This suggests at least the capacity (if not the explicit intent) to better understand PM systems as a whole. This was particularly prevalent with inputs and formal process research, which makes sense because, theoretically, these variables generally vary only across organizations or at least across departments within a single organization. On the other hand, we sometimes saw formal process variables measured at the individual level (as perceptions of employees; e.g., in Liu & Liu, 2011, employees rated their agreement that team participation was an appraisal dimension in their organization). Although it is reasonable to examine perceptions of formal process variables, strong theoretical rationale should be provided for why one would expect variance across individuals on formal processes. Although in general we encourage continued focus on organization-level PM data, an obvious challenge to conducting such work is the typical sample requirements—often many organizations or at least units within organizations. For example, studies by Haines and St-Onge (2012) and Farndale et al. (2011), which both examined the impact of culture/climate on PM, included 312 public and private sector organizations and 4,422 employees across 22 business units, respectively. In addition, this organization-level research often seems to come with some methodological limitations, such as one-time

cross-sectional surveys (used in the majority of input studies in our database), that need to be overcome in future research. Yet some researchers are more sophisticated in their choices regarding data sources. Jiang, Wang, and Zhao (2012), for example, conducted surveys in 106 firms but utilized three different sources for the data: HR managers provided responses on HR practices, production/operations managers provided responses on innovation, and employees provided responses on creativity. We would encourage future PM researchers to maintain high standards of methodology in collecting such organization-level data.

Fourth, we wish to clarify that we are not suggesting that sampling at the PM system level is the only viable option for research. Researchers could study variables from separate components of our systems model and examine interactions among them without having to sample multiple intact PM *systems*. (As an example, Kuvaas, 2007, examined the tasks of setting performance expectations and performance feedback, explicitly for a developmental purpose, under a contingency factor related to individual ratees–autonomy orientation.) The field can build a more conclusive science of multiple components of PM systems if we can start to map patterns of interactions among the components; this is true even if the primary studies are not conducted at the system level. However, building this knowledge is not possible if researchers confine themselves to studying only one variable or do not examine interactions among variables from different components of the system. In addition, variance in many PM system components can often be found across work groups, departments, or divisions, and it is generally more feasible to collect data at this level.

Fifth, if inputs are important for the effectiveness of PM, it seems essential that researchers describe the context of their studies vis-à-vis these sorts of variables (e.g., country, industry, organization type, culture, purpose of PM). We found a great deal of variance in how well authors discuss the PM context of their study (see Kleingeld, Van Tuijl, & Algera, 2004, for a particularly detailed example). Most studies (82%) provided *some* information on context, but it was quite limited: country most frequently, followed by industry and public versus private organization. Given that our systems approach affords a particularly important role to purpose, we also investigated the frequency with which the purpose of the PM system was noted in the study description. Only 25% of all empirical articles explicitly stated the purpose(s) of PM in the organization when describing the context of the study. Unfortunately, our parsing of the research into older and more recent research suggests that more recent articles are less, not more, likely to explicitly state the purpose of PM (21% vs. 39%). This is a crucial contextual variable for interpreting the effects of any aspect of PM, and it is perplexing that as we have come further in the PM literature and in our understanding of the importance of context in general (Johns, in press), researchers are less likely to describe this essential contextual element. We strongly encourage authors of future research articles on PM to be very explicit with regard to the context of the PM system under study, including national culture, industry, type of firm, organizational strategy and culture, and especially purpose of the PM system.

A final general and particularly pressing need identified from our review is better measures of PM systems themselves. The way researchers (particularly strategic HR researchers) talk about and measure PM often ignores its complexity. For example, Raeder, Knorr, and Hilb (2012) measured PM at the organization level in terms of the percentage of employees assessed; Kooij, Guest, Clinton, Knight, Jansen, and Dikkers (2013) measured it at the individual level by asking whether the employee had had a formal PA in the last 12 months (scored as yes/no/not sure). Such measures ignore the "type" of PM system, and future

research should endeavor to create a measure of PM systems that goes beyond the mere existence of *any PM* in the organization and instead reflects the important dimensions of PM as identified in research. The current review could be very helpful in this regard (see Table 2), and efforts should be directed at using the results summarized here to create a theoretically and empirically based measure of PM systems to be used in organizational survey research. We believe, on the basis of our review, that such a measure would become indispensable to future PM research.

Conclusions

We reviewed more than 36 years of the PM literature, using a systems-based model to distill knowledge from the extant (and especially recent) research and to identify productive directions for future research. Our overarching conclusion is that despite progress in the study of PM, there is much work yet to be done in developing a comprehensive, conclusive, and systems theory–grounded body of scientific knowledge about PM systems that can better inform PM in practice. However, given the passion for these topics that we have witnessed across both research and practice and across more than 36 years of the literature, we remain optimistic that as a field, we will get there.

Notes

1. We use the terms *performance appraisal* and *performance management* somewhat interchangeably when referring to the body of literature only. This is because the scope of our review includes research and practice in both areas. But it is not that these are synonymous terms. As we discuss later in our review, the latter is broader and includes more components than the former; hence, we use this latter label to describe the overall focus of our review.

2. Although going back to 1980 is important for both establishing the components of our model (designed to be comprehensive with regard to the literature in general) and providing context for what is "new" about more recent PM work, the bulk of our paper focuses on the literature post 2004, given the date of the last comprehensive review (Levy & Williams, 2004). Thus, the information on each component of our model reviews this more recent research (post 2004), and our identification of recent trends relies on the same time frame. But when we summarize what is "known" about each component, we combine the most recent research with the older work to provide more conclusive and comprehensive statements about what is established knowledge. Of course, when we draw comparisons between the older and the more recent research, we are contrasting pre 2004 and post 2004.

3. This referred to studies that were so weak in methodology that their results were entirely uninterpretable (e.g., scales with no validity, no discussion of methodology used), not merely to articles with relatively weaker designs (e.g., all single-source cross-sectional data, or failing to account for the nested nature of data).

References

Abu-Doleh, J., & Weir, D. 2007. Dimensions of performance appraisal systems in Jordanian private and public organizations. *The International Journal of Human Resource Management*, 18: 75-84.

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*, 9: 219-252.

Aguinis, H. 2013. *Performance management* (3rd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Akuratiyagamage, V. M. 2005. Identification of management development needs: A comparison across companies of different ownership—foreign, joint venture and local in Sri Lanka. *The International Journal of Human Resource Management*, 16: 1512-1528.

Alder, G. S., & Ambrose, M. L. 2005. An examination of the effect of computerized performance monitoring feedback on monitoring fairness, performance, and satisfaction. *Organizational Behavior and Human Decision Processes*, 97: 161-177.

Anderson, J. R., & Johnson, N. B. 2005. On the relationship between work contexts, mandates and compliance behaviours of supervisors. *Journal of Change Management*, 5: 381-390.

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.

Ayers, R. S. 2013. Building goal alignment in federal agencies' performance appraisal programs. *Public Personnel Management*, 42: 495-520.

Bande Vilela, B., Varela González, J. A., Fernández Ferrín, P., & Luisa del Río Araújo, M. 2007. Impression management tactics and affective context: Influence on sales performance appraisal. *European Journal of Marketing*, 41: 624-639.

Becker, K., Antuar, N., & Everett, C. 2011. Implementing an employee performance management system in a non-profit organization. *Nonprofit Management and Leadership*, 21: 255-271.

Bernardin, H. J., & Villanova, P. 2005. Research streams in rater self-efficacy. *Group & Organization Management*, 30: 61-88.

Bhave, D. P. 2014. The invisible eye? Electronic performance monitoring and employee job performance. *Personnel Psychology*, 67: 605-635.

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.

Blume, B. D., Rubin, R. S., & Baldwin, T. T. 2013. Who is attracted to an organisation using a forced distribution performance management system? *Human Resource Management Journal*, 23: 360-378.

Boulding, K. E. 1956. General systems theory: The skeleton of science. *Management Science*, 2: 197-208.

Bretz, R. D., 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.

Brown, T. C., & Warren, A. M. 2011. Performance management in unionized settings. *Human Resource Management Review*, 21: 96-106.

Buckingham, M., & Goodall, A. 2015. Reinventing performance management. *Harvard Business Review*, 93(4): 40-50.

Cabrera, D. D., Fernaud, E. H., Díaz, R. I., Rodríguez, N. D., Vilela, L. D., & Sánchez, C. R. 2014. Factores relevantes para aumentar la precisión, la viabilidad y el éxito de los sistemas de evaluación del desempeño laboral [Relevant factors to increase the accuracy, feasibility and success of job performance evaluation systems]. *Papeles del Psicólogo*, 35: 115-121.

Cascio, W. F. 2006. 10 global performance management systems. In G. K. Stahl & I. Björkman (Eds.), *Handbook of research in international human resource management*: 176-196. Northampton, MA: Edward Elgar.

Cavanaugh, C., Levy, P. E., Schleicher, D. J., Anseel, F., Colquitt, A., & Hunt, S. 2013. *Performance management researchers and practitioners…You need to hear this, 2.0!* Panel discussion at the annual conference of the Society for Industrial and Organizational Psychology, Houston, TX.

Clarke, C., Harcourt, M., & Flynn, M. 2012. Clinical governance, performance appraisal and interactional and procedural fairness at a New Zealand public hospital. *Journal of Business Ethics*, 117: 667-678.

Claus, L., & Briscoe, D. 2009. Employee performance management across borders: A review of relevant academic literature. *International Journal of Management Reviews*, 11: 175-196.

Cleveland, J. N., Murphy, K. R., & Williams, R. E. 1989. Multiple uses of performance appraisal: Prevalence and correlates. *Journal of Applied Psychology*, 74: 130-135.

Collins-Camargo, C., Chuang, E., McBeath, B., & Bunger, A. C. 2014. Private child welfare agency managers' perceptions of the effectiveness of different performance management strategies. *Children and Youth Services Review*, 38: 133-141.

Colquitt, A. L. 2017. *Next generation performance management: The triumph of science over myth and superstition*. Charlotte, NC: Information Age.

Colquitt, J. A., Scott, B. A., Rodell, J. B., Long, D. M., Zapata, C. P., Conlon, D. E., & Wesson, M. J. 2013. Justice at the millennium, a decade later: A meta-analytic test of social exchange and affect-based perspectives. *Journal of Applied Psychology*, 98: 199-236.

Culbertson, S. S., Henning, J. B., & Payne, S. C. 2013. Performance appraisal satisfaction: The role of feedback and goal orientation. *Journal of Personnel Psychology*, 12: 189-195.

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.

Curtis, A. B., Harvey, R. D., & Ravden, D. 2005. Sources of political distortions in performance appraisals: Appraisal purpose and rater accountability. *Group & Organization Management*, 30: 42-60.

David, E. M. 2013. Examining the role of narrative performance appraisal comments on performance. *Human Performance*, 26: 430-450.

Decramer, A., Smolders, C., & Vanderstraeten, A. 2013. Employee performance management culture and system features in higher education: Relationship with employee performance management satisfaction. *The International Journal of Human Resource Management*, 24: 352-371.

den Hartog, D. N., Boselie, P., & Paauwe, J. 2004. Performance management: A model and research agenda. *Applied Psychology*, 53: 556-569.

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.

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.

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.

Dewberry, C., Davies-Muir, A., & Newell, S. 2013. Impact and causes of rater severity/leniency in appraisals without postevaluation communication between raters and ratees. *International Journal of Selection and Assessment*, 21: 286-293.

Dhiman, A., & Maheshwari, S. K. 2013. Performance appraisal politics from appraisee perspective: A study of antecedents in the Indian context. *The International Journal of Human Resource Management*, 24: 1202-1235.

Dusterhoff, C., Cunningham, J. B., & MacGregor, J. N. 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.

Effron, M., & Ort, M. 2010. *One page talent management: Eliminating complexity, adding value*. Boston: Harvard Business School.

Eichinger, R. W., Lombardo, M. M., Ulrich, D., & Cannon, K. 2004. *100 things you need to know: Best people practices for managers & HR*. Minneapolis: Lominger.

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.

Ellis, D. R. 2012. Exploring cultural dimensions as predictors of performance management preferences: The case of self-initiating expatriate New Zealanders in Belgium. *The International Journal of Human Resource Management*, 23: 2087-2107.

Erdemli, C., Sümer, H. C., & Bilgiç, R. 2007. A comparison of written feedback and written plus verbal feedback methods in performance management. *Türk Psikoloji Dergisi*, 22(60): 71-85.

Erez, A., Schilpzand, P., Leavitt, K., Woolum, A. H., & Judge, T. A. 2015. Inherently relational: Interactions between peers' and individuals' personalities impact reward giving and appraisal of individual performance. *Academy of Management Journal*, 58: 1761-1784.

Evans, E. M., & McShane, S. L. 1988. Employee perceptions of performance appraisal fairness in two organizations. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement*, 20: 177-191.

Farndale, E., Hope-Hailey, V., & Kelliher, C. 2011. High commitment performance management: The roles of justice and trust. *Personnel Review*, 40: 5-23.

Fee, A., McGrath-Champ, S., & Yang, X. 2011. Expatriate performance management and firm internationalization: Australian multinationals in China. *Asia Pacific Journal of Human Resources*, 49: 365-384.

Ferguson, A. J., Ormiston, M. E., & Moon, H. 2010. From approach to inhibition: The influence of power on responses to poor performers. *Journal of Applied Psychology*, 95: 305-320.

Giumetti, G. W., Schroeder, A. N., & Switzer, I. S. 2015. Forced distribution rating systems: When does "rank and yank" lead to adverse impact? *Journal of Applied Psychology*, 100: 180-193.

Godin, B. 2009. National innovation system: The system approach in historical perspective. *Science, Technology, & Human Values*, 34: 476-501.

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.

Goffin, R. D., Jelley, R. B., Powell, D. M., & Johnston, N. G. 2009. Taking advantage of social comparisons in performance appraisal: The relative percentile method. *Human Resource Management*, 48: 251-268.

Goh, F. A., & Anderson, M. C. 2007. Driving business value from performance management at Caterpillar. *Organization Development Journal*, 25: P219-P226.

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.

Grant, A. M., Parker, S., & Collins, C. 2009. Getting credit for proactive behavior: Supervisor reactions depend on what you value and how you feel. *Personnel Psychology*, 62: 31-55.

Gregory, J. B., & Levy, P. E. 2010. Employee coaching relationships: Enhancing construct clarity and measurement. *Coaching: An International Journal of Theory, Research and Practice*, 3: 109-123.

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: Practice and Research*, 63: 67-88.

Gregory, J. B., & Levy, P. E. 2012. Employee feedback orientation: Implications for effective coaching relationships. *Coaching: An International Journal of Theory, Research and Practice*, 5: 86-99.

Gregory, J. B., Levy, P. E., & Jeffers, M. 2008. Development of a model of the feedback process within executive coaching. *Consulting Psychology Journal: Practice and Research*, 60: 42-56.

Haines, V. Y., & St-Onge, S. 2012. Performance management effectiveness: Practices or context? *The International Journal of Human Resource Management*, 23: 1158-1175.

Hebert, B. G., & Vorauer, J. D. 2003. Seeing through the screen: Is evaluative feedback communicated more effectively in face-to-face or computer-mediated exchanges? *Computers in Human Behavior*, 19: 25-38.

Hernandez, M., Avery, D. R., Tonidandel, S., Hebl, M. R., Smith, A. N., & McKay, P. F. 2016. The role of proximal social contexts: Assessing stigma-by-association effects on leader appraisals. *Journal of Applied Psychology*, 101: 68-85.

Heslin, P. A., & VandeWalle, D. 2011. Performance appraisal procedural justice: The role of a manager's implicit person theory. *Journal of Management*, 37: 1694-1718.

Heslin, P. A., VandeWalle, D., & Latham, G. P. 2006. Keen to help? Managers' implicit person theories and their subsequent employee coaching. *Personnel Psychology*, 59: 871-902.

Hofstetter, H., & Harpaz, I. 2015. Declared versus actual organizational culture as indicated by an organization's performance appraisal. *The International Journal of Human Resource Management*, 26: 445-466.

Howard, A. E. 1995. *The changing nature of work*. San Francisco: Jossey-Bass.

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.

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.

Jackson, C. 1996. An individual differences approach to the halo-accuracy paradox. *Personality and Individual Differences*, 21: 947-957.

Jawahar, I. M. 2005. Do raters consider the influence of situational factors on observed performance when evaluating performance? Evidence from three experiments. *Group & Organization Management*, 30: 6-41.

Jiang, J., Wang, S., & Zhao, S. 2012. Does HRM facilitate employee creativity and organizational innovation? A study of Chinese firms. *The International Journal of Human Resource Management*, 23: 4025-4047.

Johns, G. in press. Advances in the treatment of context in organizational research. *Annual Review of Organizational Psychology and Organizational Behavior*. doi:10.1146/annurev-orgpsych-032117-104406

Kacmar, K. M., Wayne, S. J., & Wright, P. M. 2009. Subordinate reactions to the use of impression management tactics and feedback by the supervisor. *Journal of Managerial Issues*, 21: 498-517.

Kamer, B., & Annen, H. 2010. The role of core self-evaluations in predicting performance appraisal reactions. *Swiss Journal of Psychology*, 69: 95-104.

Kataoka, H. C., Cole, N. D., & Flint, D. A. 2006. Due process model of procedural justice in performance appraisal: Promotion versus termination scenarios. *Psychological Reports*, 99: 819-832.

Katz, D., & Kahn, R. L. 1978. *The social psychology of organizations* (2nd ed.). New York: Wiley.

Kavanagh, P., Benson, J., & Brown, M. 2007. Understanding performance appraisal fairness. *Asia Pacific Journal of Human Resources*, 45: 132-150.

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. *The International Journal of Human Resource Management*, 21: 2031-2051.

Kepes, S., Bennett, A. A., & McDaniel, M. A. 2014. Evidence-based management and the trustworthiness of our cumulative scientific knowledge: Implications for teaching, research, and practice. *Academy of Management Learning & Education*, 13: 446-466.

Kinicki, A. J., Jacobson, K. J., Peterson, S. J., & Prussia, G. E. 2013. Development and validation of the performance management behavior questionnaire. *Personnel Psychology*, 66: 1-45.

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.

Klein, H. J., Snell, S. A., & Wexley, K. N. 1987. Systems model of the performance appraisal interview process. *Industrial Relations: A Journal of Economy and Society*, 26: 267-280.

Kleingeld, A., van Mierlo, H., & Arends, L. 2011. The effect of goal setting on group performance: A meta-analysis. *Journal of Applied Psychology*, 96: 1289-1304.

Kleingeld, A., van Tuijl, H., & Algera, J. A. 2004. Participation in the design of performance management systems: A quasi-experimental field study. *Journal of Organizational Behavior*, 25: 831-851.

Kluger, A. N., & DeNisi, A. 1996. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119: 254-284.

Kooij, D. T. A. M., Guest, D. E., Clinton, M., Knight, T., Jansen, P. G. W., & Dikkers, J. S. E. 2013. How the impact of HR practices on employee well-being and performance changes with age. *Human Resource Management Journal*, 23: 18-35.

Krats, P., & Brown, T. C. 2013. Unionised employee's reactions to the introduction of a goal-based performance appraisal system. *Human Resource Management Journal*, 23: 396-412.

Kulas, J. T., & Hannum, K. M. 2006. Investigation of 360-degree instrumentation effects: Application of the Rasch rating scale model. *Journal of Applied Measurement*, 8: 404-416.

Kurtzberg, T. R., Naquin, C. E., & Belkin, L. Y. 2005. Electronic performance appraisals: The effects of e-mail communication on peer ratings in actual and simulated environments. *Organizational Behavior and Human Decision Processes*, 98: 216-226.

Kuvaas, B. 2006. Performance appraisal satisfaction and employee outcomes: Mediating and moderating roles of work motivation. *The International Journal of Human Resource Management*, 17: 504-522.

Kuvaas, B. 2007. Different relationships between perceptions of developmental performance appraisal and work performance. *Personnel Review*, 36: 378-397.

Kuvaas, B. 2011. The interactive role of performance appraisal reactions and regular feedback. *Journal of Managerial Psychology*, 26: 123-137.

Landy, F. J., & Farr, J. L. 1980. Performance rating. *Psychological Bulletin*, 87: 72-107.

Latham, G. P., Budworth, M., 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.

Lawler, E. E. 2003. Reward practices and performance management system effectiveness. *Organizational Dynamics*, 32: 396-404.

Lawler, E. E., & 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. *The International Journal of Human Resource Management*, 21: 1351-1372.

Levy, P. E., Cavanaugh, C., Frantz, N. B., & Borden, L. A. 2015. The role of due process in performance appraisal: A 20 year retrospective. In R. S. Cropanzano & M. L. Ambrose (Eds.), *Oxford handbook of justice in the workplace*: 605-620. New York: Oxford University Press.

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." In M. R. Buckley, A. R. Wheeler, & J. R. B. Halbesleben (Eds.), *Research in personnel and human resources management*, vol. 35: 155-213. Bingley, England: Emerald.

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.

Lievens, F., Conway, J. M., & De Corte, W. 2008. The relative importance of task, citizenship and counterproductive performance to job performance ratings: Do rater source and team-based culture matter? *Journal of Occupational and Organizational Psychology*, 81: 11-27.

Linderbaum, B. A., & Levy, P. E. 2010. The development and validation of the Feedback Orientation Scale (FOS). *Journal of Management*, 36: 1372-1405.

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. *The International Journal of Human Resource Management*, 23: 1360-1375.

Liu, N., & Liu, M. 2011. Human resource practices and individual knowledge-sharing behavior: An empirical study for Taiwanese R&D professionals. *The International Journal of Human Resource Management*, 22: 981-997.

Lyness, K. S., & Heilman, M. E. 2006. When fit is fundamental: Performance evaluations and promotions of upper-level female and male managers. *Journal of Applied Psychology*, 91: 777-785.

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.

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.

Martell, R. F., & Evans, D. P. 2005. Source-monitoring training: Toward reducing rater expectancy effects in behavioral measurement. *Journal of Applied Psychology*, 90: 956-963.

Medcof, J. W., & Song, L. J. 2013. Exploration, exploitation and human resource management practices in cooperative and entrepreneurial HR configurations. *The International Journal of Human Resource Management*, 24: 2911-2926.

Mento, A. J., Steel, R. P., & Karren, R. J. 1987. A meta-analytic study of the effects of goal setting on task performance: 1966-1984. *Organizational Behavior and Human Decision Processes*, 39: 52-83.

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.

Miller, J. S. 2003. High tech and high performance: Managing appraisal in the information age. *Journal of Labor Research*, 24: 409-424.

Mok, A., Cheng, C. Y., & Morris, M. W. 2010. Matching versus mismatching cultural norms in performance appraisal: Effects of the cultural setting and bicultural identity integration. *International Journal of Cross Cultural Management*, 10: 17-35.

Murphy, K. R. 2008. Explaining the weak relationship between job performance and ratings of job performance. *Industrial and Organizational Psychology*, 1: 148-160.

Murphy, K. R., & Cleveland, J. N. 1991. *Performance appraisal: An organizational perspective*. Needham Heights, MA: Allyn & Bacon.

Nadler, D. A., & Tushman, M. L. 1980. A model for diagnosing organizational behavior. *Organizational Dynamics*, 9(2): 35-51.

Nadler, D. A., & Tushman, M. L. 1984. A congruence model for diagnosing organizational behavior. In D. A. Kolb, J. M. Rubin, & J. M. McIntyre (Eds.), *Organizational psychology: Reading on human behavior in organizations*: 587-603. Englewood Cliffs, NJ: Prentice Hall.

Nayyab, H. H., Hamid, M., Naseer, F., & Iqbal, M. 2011. The impact of HRM practices on the organizational performance: The study of banking sector in Okara, Punjab (Pakistan). *Interdisciplinary Journal of Contemporary Research in Business*, 3: 661-672.

Neubert, M. J. 1998. The value of feedback and goal setting over goal setting alone and potential moderators of this effect: A meta-analysis. *Human Performance*, 11: 321-335.

Ng, K. Y., Koh, C., Ang, S., Kennedy, J. C., & Chan, K. Y. 2011. Rating leniency and halo in multisource feedback ratings: Testing cultural assumptions of power distance and individualism-collectivism. *Journal of Applied Psychology*, 96: 1033-1044.

Nishii, L. H., & Wright, P. 2008. Variability at multiple levels of analysis: Implications for strategic human resource management. In D. B. Smith (Ed.), *The people make the place*: 225-248. New York: Erlbaum.

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.

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.

Nurse, L. 2005. Performance appraisal, employee development and organizational justice: Exploring the linkages. *The International Journal of Human Resource Management*, 16: 1176-1194.

Ogunfowora, B., Bourdage, J., & Lee, K. 2010. Rater personality and performance dimension weighting in making overall performance judgments. *Journal of Business and Psychology*, 25: 465-476.

Pampino, R. N., MacDonald, J. E., Mullin, J. E., & Wilder, D. A. 2004. Weekly feedback vs. daily feedback: An application in retail. *Journal of Organizational Behavior Management*, 23: 21-43.

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.

Pearce, J. L., & Xu, Q. J. 2012. Rating performance or contesting status: Evidence against the homophily explanation for supervisor demographic skew in performance ratings. *Organization Science*, 23: 373-385.

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.

Pesta, B. J., Kass, D. S., & Dunegan, K. J. 2005. Image theory and the appraisal of employee performance: To screen or not to screen? *Journal of Business and Psychology*, 19: 341-360.

Poon, J. M. 2004. Effects of performance appraisal politics on job satisfaction and turnover intention. *Personnel Review*, 33: 322-334.

Pritchard, R. D., Harrell, M. M., DiazGranados, D., & Guzman, J. M. 2008. The productivity measurement and enhancement system: A meta-analysis. *Journal of Applied Psychology*, 93: 540-567.

Pulakos, E. D., & O'Leary, R. S. 2011. Why is performance management broken? *Industrial and Organizational Psychology*, 4: 146-164.

Pullin, L., & Haidar, A. 2003. Performance contract management in regional local government—Victoria. *Asia Pacific Journal of Human Resources*, 41: 279-297.

Raeder, S., Knorr, U., & Hilb, M. 2012. Human resource management practices and psychological contracts in Swiss firms: An employer perspective. *The International Journal of Human Resource Management*, 15: 3178-3195.

Rahman, S. A. 2006. Attitudes of Malaysian teachers toward a performance-appraisal system. *Journal of Applied Social Psychology*, 36: 3031-3042.

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.

Reb, J., & Greguras, G. J. 2010. Understanding performance ratings: Dynamic performance, attributions, and rating purpose. *Journal of Applied Psychology*, 95: 213-220.

Reisberg, D. 2006. *Cognition: Exploring the science of the mind* (3rd ed.). New York: W.W. Norton.

Roberts, I. 2001. Reward and performance management. In I. Beardwell & L. Holden (Eds.), *Human resource management: A contemporary approach*: 506-558. Edinburgh: Pearson.

Rodwell, J. J., & Teo, S. T. T. 2008. The influence of strategic HRM and sector on perceived performance in health services organizations. *The International Journal of Human Resource Management*, 19: 1825-1841.

Rothstein, H. R. 1990. Interrater reliability of job performance ratings: Growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology*, 75: 322-327.

Saffie-Robertson, M. C., & Brutus, S. 2014. The impact of interdependence on performance evaluations: The mediating role of discomfort with performance appraisal. *The International Journal of Human Resource Management*, 25: 459-473.

Schleicher, D. J., Bull, R. A., & Green, S. G. 2009. Rater reactions to forced distribution rating systems. *Journal of Management*, 35: 899-927.

Schleicher, D. J., Levy, P. E., Baumann, H. M., & Hartwell, C. J. 2012. *Calling all performance management researchers: You need to hear this!* Panel discussion presented at the annual conference of the Society for Industrial and Organizational Psychology, San Diego, CA.

Scullen, S. E., Mount, M. K., & Goff, M. 2000. Understanding the latent structure of job performance ratings. *Journal of Applied Psychology*, 85: 956-970.

Selden, S., & Sowa, J. E. 2011. Performance management and appraisal in human service organizations: Management and staff perspectives. *Public Personnel Management*, 40: 251-264.

Selvarajan, T. T., & Cloninger, P. A. 2012. Can performance appraisals motivate employees to improve performance? A Mexican study. *The International Journal of Human Resource Management*, 23: 3063-3084.

Shahnawaz, M. G., & Juyal, R. C. 2006. Human resource management practices and organizational commitment in different organizations. *Journal of the Indian Academy of Applied Psychology*, 32: 171-178.

Silva, P., & Tosi, H. L. 2004. Determinants of the anonymity of the CEO evaluation process. *Journal of Managerial Issues*, 16: 87-102.

Silzer, R., & Church, A. H. 2009. The pearls and perils of identifying potential. *Industrial and Organizational Psychology*, 2: 377-412.

Singh, S., Mohamed, A. F., & Darwish, T. 2013. A comparative study of performance appraisals, incentives and rewards practices in domestic and multinational enterprises in the country of Brunei Darussalam. *The International Journal of Human Resource Management*, 24: 3577-3598.

Skarlicki, D. P., & Turner, R. A. 2014. Unfairness begets unfairness: Victim derogation bias in employee ratings. *Organizational Behavior and Human Decision Processes*, 124: 34-46.

Smither, J. W., London, M., & Reilly, R. R. 2005. Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings. *Personnel Psychology*, 58: 33-66.

Steelman, L. A., Levy, P. E., & Snell, A. F. 2004. The Feedback Environment Scale: Construct definition, measurement, and validation. *Educational and Psychological Measurement*, 64: 165-184.

Taormina, R. J., & Gao, J. H. 2009. Identifying acceptable performance appraisal criteria: An international perspective. *Asia Pacific Journal of Human Resources*, 47: 102-125.

Theeboom, T., Beersma, B., & van Vianen, A. E. 2014. Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context. *The Journal of Positive Psychology*, 9: 1-18.

Thurston, P. W., & McNall, L. 2010. Justice perceptions of performance appraisal practices. *Journal of Managerial Psychology*, 25: 201-228.

Tubre, T. C., & Collins, J. M. 2000. Jackson and Schuler (1985) revisited: A meta-analysis of the relationships between role ambiguity, role conflict, and job performance. *Journal of Management*, 26: 155-169.

Tuytens, M., & Devos, G. 2012a. Importance of system and leadership in performance appraisal. *Personnel Review*, 41: 756-776.

Tuytens, M., & Devos, G. 2012b. The effect of procedural justice in the relationship between charismatic leadership and feedback reactions in performance appraisal. *The International Journal of Human Resource Management*, 23: 3047-3062.

van der Leeuw, R. M., Overeem, K., Arah, O. A., Heineman, M. J., & Lombarts, K. M. J. M. H. 2013. Frequency and determinants of residents' narrative feedback on the teaching performance of faculty: Narratives in numbers. *Academic Medicine*, 88: 1324-1331.

Wang, X. M., Wong, K. F. E., & Kwong, J. Y. Y. 2010. The roles of rater goals and ratee performance levels in the distortion of performance ratings. *Journal of Applied Psychology*, 95: 546-561.

Webb Day, J., Holladay, C. L., Johnson, S. K., & Barron, L. G. 2014. Organizational rewards: Considering employee need in allocation. *Personnel Review*, 43: 74-95.

Whiting, S. W., Podsakoff, P. M., & Pierce, J. R. 2008. Effects of task performance, helping, voice, and organizational loyalty on performance appraisal ratings. *Journal of Applied Psychology*, 93: 125-139.

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.

Williams, M. J. 1997. Performance appraisal is dead. Long live performance management. *Harvard Management Update*, 2: 1-6.

Woehr, D. J., Sheehan, M. K., & Bennett, W. 2005. Assessing measurement equivalence across rating sources: A multitrait-multirater approach. *Journal of Applied Psychology*, 90: 592-600.

Wong, K. F. E., & Kwong, J. Y. Y. 2005. Between-individual comparisons in performance evaluation: A perspective from prospect theory. *Journal of Applied Psychology*, 90: 284-294.

Wong, K. F. E., & Kwong, J. Y. Y. 2007. Effects of rater goals on rating patterns: Evidence from an experimental field study. *Journal of Applied Psychology*, 92: 577-585.

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.

Yang, H., & Klaas, B. S. 2011. Pay dispersion and the financial performance of the firm: Evidence from Korea. *The International Journal of Human Resource Management*, 22: 2147-2166.

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.

Zheng, C., Morrison, M., & O'Neill, G. 2006. An empirical study of high performance HRM practices in Chinese SMEs. *The International Journal of Human Resource Management*, 17: 1772-1803.

Zheng, W., Zhang, M., & Li, H. 2012. Performance appraisal process and organizational citizenship behavior. *Journal of Managerial Psychology*, 27: 732-752.