

Focal Article

The Evolution and Devolution of 360° Feedback

David W. Bracken

DWBracken & Associates

Dale S. Rose

3D Group

Allan H. Church

PepsiCo

In the 25+ years that the practice of 360° Feedback has been formally labeled and implemented, it has undergone many changes. Some of these have been positive (evolution) in advancing theory, research, and practice, and others less so (devolution). In this article we offer a new definition of 360° Feedback, summarize its history, discuss significant research and practice trends, and offer suggestions for all user communities (i.e., researchers, practitioners, and end users in organizations) moving forward. Our purpose is to bring new structure, discussion, and some degree of closure to key open issues in this important and enduring area of practice.

Keywords: 360 degree feedback, leadership development, assessment

Under whatever label, what we now commonly refer to as 360° Feedback has been around for a long time and has become embedded in many of our human resources (HR) processes applied at individual, group, and organizational levels. The practice of 360° Feedback¹ has been around so long now that we are all comfortable with it. In fact, we would contend that the profession has become so comfortable with this practice area that we have lost our focus on what it is and what it is not. What was once considered a “fad” by many in the 1990s and cutting edge is now well known and a

David W. Bracken, DWBracken & Associates, Atlanta, Georgia; Dale S. Rose, 3D Group, Emeryville, California; Allan H. Church, PepsiCo, Purchase, New York.

The authors wish to thank Carol Timmreck and Jeff McHenry for their invaluable suggestions.

Correspondence concerning this article should be addressed to David W. Bracken, DWBracken & Associates, 2271A Dunwoody Crossing, Atlanta, GA 30338. E-mail: dwbracken@gmail.com

¹ We use capitalization to indicate our desire to indicate a proper noun. This is by design and is intended to draw attention to our wish for 360° Feedback to have a well-defined meaning, which is offered later in this article.

standard tool (or “core HR process”) in many organizations (e.g., Church, Rotolo, Shull, & Tuller, 2014b). Although it is not necessarily a bad thing to have industrial–organizational (I-O) psychology practices become well established in the HR community, as the tool proliferates, it can begin to morph in some challenging ways. Given this context, and almost paradoxically as a result of 360° Feedback’s well-established nature, we believe it is time to take a step back and provide a comprehensive overview of the conceptual underpinnings and philosophical debates associated with the use of 360° Feedback today. Our intent with this article is to help put some definitional and conceptual guardrails around the current practice of what should be considered 360° Feedback going forward, with an eye toward igniting new theory, research, and even practice innovations in this area.

We will begin by taking a stance on defining the practice area in general. We believe that being specific about what 360° Feedback is (and is not) is the starting point for a healthy discussion and debate, creating clarity as to what it is we are debating. By doing so, we are excluding other feedback processes from the discussion, which probably alone will create a debate.

We will also highlight key issues and best practices, focusing on the practices that we see as being antithetical to the use of 360° Feedback to create sustained change, or examples of the misuse or misrepresentation of 360° Feedback that are creating unintended consequences and inhibiting its adoption by various users. The article concludes with some recommendations designed to guide future research and practice beyond just the best practice recommendations themselves.

Short History of the Evolution of 360° Feedback

Arguably the first systematic treatment of 360° Feedback occurred with the publication of Edwards and Ewen’s (1996) book, *360° Feedback: The Powerful New Model for Employee Assessment & Performance Improvement*. In this book, Edwards and Ewen describe the “coining” of the term “360° Feedback” back in the mid-1980s as an alternative to the term “multirater.” They then registered 360° Feedback as a trademark for their firm, TEAMS, Inc. (Edwards & Ewen, 1996, p. 4), and expended quite a bit of time and energy in an attempt to enforce that trademark until the firm was acquired.

The Edwards and Ewen (1996) bibliography includes only one article that uses the term “360° Feedback” in its title before the book was published (Edwards & Ewen, 1995). It would not have been in their best interest to quote other sources that used the term while defending the copyright. Around 1993, a number of references to 360° Feedback began to surface (e.g., Hazucha, Hezlett, & Schneider, 1993; Kaplan, 1993; London & Beatty, 1993; Nowack, 1993), and by the time the Edwards/Ewen book was published, the

term was quickly being accepted as a concept and practice (Church, 1995; “Companies Use,” 1993; O’Reilly, 1994; Swoboda, 1993).

One of the most significant characteristics of the Edwards and Ewen book draws from the total focus of their consulting firm (TEAMS, Inc.) on 360° Feedback systems and from the authors being two of the earliest practitioners who were well positioned to conduct large scale administrations across dozens of organizations and create a substantial research database in the process. As can be seen in the title of the book, their strong bias and message was that 360° Feedback not only is appropriate for use in assessment and appraisal but also can have major advantages over single source (supervisor) evaluations. They had large data sets to support their position.

The Edwards/Ewen book was quickly followed by a number of books that provided broader perspectives on the subject (and ignored the trademark claim as well)—*The Art and Science of 360° Feedback* (Lepsinger & Lucia, 1997), *Maximizing the Value of 360-Degree Feedback* (Tornow, London, & CCL Associates, 1998), and *The Handbook of Multisource Feedback* (Bracken, Timmreck, & Church, 2001)—the last of which did not use the trademarked term in deference to Edwards and Ewen, who had contributed two chapters. Compared with Edwards and Ewen (1996), the Lepsinger and Lucia (1997) and Tornow et al. (1998) books had a much deeper emphasis on developmental applications and hesitance regarding use in personnel decisions (e.g., appraisal, compensation, staffing, downsizing). In that same period, a Society for Industrial and Organizational Psychology debate on the use of 360° Feedback was transcribed and published by Center for Creative Leadership (CCL), *Should 360-Degree Feedback Be Used Only for Developmental Purposes?* (Bracken, Dalton, Jako, McCauley, & Pollman, 1997), reflecting the deep schism in the field regarding how 360° Feedback should and shouldn’t be used in regard to decision making (London, 2001).

Of those four books, only *The Handbook* provides an in-depth treatment of the history and evolution of the field up to that point. Hedge, Borman, and Birkeland (2001) devote an entire chapter to “The History and Development of Multisource Feedback as a Methodology,” recommended reading (along with Edwards & Ewen, 1996) for students of the field. Their treatment takes us back as far as 1920 but identifies Lawler’s article (Lawler, 1967) on the multitrait–multirater method as the dawn of what they call the “Modern Multirater Perspective.”²

Subsequent to the first generation of books on the topic, a dozen or so books of varying length, depth, and application hit the market, most

² We acknowledge that the roots of 360° Feedback can be traced back to many relevant sources, including performance management, employee surveys, and assessment centers, even perhaps to Taylor. That discussion is purposively excluded in this article, which is bounded largely by the introduction of 360° Feedback as a specific, identified process in itself.

notably a second edition of Lepsinger and Lucia's book (Lepsinger & Lucia, 2009) and a similar volume from CCL (Fleenor, Taylor, & Chappelow, 2008). Contributions from Jones and Bearley (1996) and Ward (1997) are examples of "how to" manuals. Sample book chapters include Bracken (1996) and Church, Walker, and Brockner (2002). We are aware of two journal editions that were dedicated to the topic, including *Human Resource Management* (Tornow, 1993) and *Group & Organization Management* (Church & Bracken, 1997).

Other resources include compendia and reviews of assessment tools, such as Van Velsor, Leslie, and Fleenor (1997) and Leslie (2013). An additional type of perspective on the state of the art of 360°s can be found in the 3D Group's series of benchmark studies, of which the most recent (3D Group, 2013) is the most comprehensive and robust, and in comprehensive review articles such as Nowack and Mashihhi (2012).

Definition of 360° Feedback

Part of the evolution of an area of theory and practice in I-O psychology involves achieving greater clarity and refinement in the constructs and operational definitions of what is being measured and/or executed. In the early days of 360° Feedback in the 1980s–1990s, when the application was taking off in industry, there were a plethora of names, approaches, and applications (multirater, multisource, full circle, 450 feedback, etc.), which reflected a wide range in scope. Over the years we have seen many different forms of feedback spring from the origins of 360° Feedback as well (e.g., 360° interview approaches, 360° personality measures, 360° customer surveys, 360° pulse surveys, etc.). In order to draw some lines in the sand and help reposition the practice as a formal construct, we offer a new definition of 360° Feedback. It is our hope that this will provide greater clarity regarding what should (and should not) be included in the discussion, practice, and promotion of the field of I-O psychology going forward. The definition is as follows:

360° Feedback is a process for collecting, quantifying, and reporting coworker observations about an individual (i.e., a ratee) that facilitates/enables three specific data-driven/based outcomes: (a) the collection of rater perceptions of the degree to which specific behaviors are exhibited; (b) the analysis of meaningful comparisons of rater perceptions across multiple ratees, between specific groups of raters for an individual ratee, and for ratee changes over time; and (c) the creation of sustainable individual, group, and/or organizational changes in behaviors valued by the organization.

As with any definition, ours is an attempt at balancing precise language with an effort toward as much brevity as possible. Although brevity has its value, some explanation of, and rationale for, each component of our definition may prove useful in evaluating its utility:

1. In our definition, a 360° Feedback "process" includes all the steps that affect the quality (reliability, validity, execution, and acceptance) of the

feedback, from design through use. If its purpose is to create change in behaviors valued by the organization, it must be designed to align with organizational behavioral requirements. These requirements will be based on such factors as the strategic needs of the business, current or desired cultural norms, leadership mission and vision, capabilities required, and so forth. Because of this need for relevance to the business, fundamentally, the data must be shared, reviewed, and ultimately used for decision making at some base level. If it is not “used” (e.g., to influence development and/or other decisions), then it is not a 360° Feedback “process” and is just an event. We propose that feedback methods that do not have these features can be called “alternate forms of feedback” (AFF). It should be noted that we are suggesting a new term here (AFF) versus using existing alternate labels for 360° Feedback such as multirater or multisource, which will be explained later in this article.

2. Our definition specifically includes observations from coworkers only, with the intention of excluding individuals not directly involved in work (e.g., family members, friends). We would include internal customers, external customers, suppliers, consultants, contractors, and other such constituents as long as the individuals’ interactions with the ratee are almost exclusively work related.
3. Although we debated the use of the term “measurement” at length and eventually elected not to include it in our definition (because it implies a testing mentality that can result in other issues, as we discuss later in the article), we did incorporate the concept of measurement by using the term “quantifying.” We believe strongly that 360° Feedback processes should be defined solely as a questionnaire-based methodology and require a *numerical assessment* of the ratee directly by the rater. As a result of this definition, we are intentionally excluding data that are collected by an interviewer who typically interviews coworkers and then translates those interviews into an evaluation of some sort of the target leader, whether quantified or not. Although we recognize that interviews may be valuable for some professional coaching or individual development interventions as a complement to 360° Feedback processes, interviews themselves should not be considered 360° Feedback. Data generated from truly qualitative interviews would not allow comparisons between rater groups on the same set of behaviors. Interviews conducted in a true qualitative framework (see Kirk & Miller, 1986) would provide no standard set of behaviors with which to compare the views of various subgroups (peers, direct reports, etc.) for a given feedback recipient. A final point to consider with these sorts of interviews is that, practically speaking, it is very unlikely that an adequate number (see Greguras & Robie, 1998) of interviews will be

conducted to allow strong conclusions about the perspective of each rating group. The requirement of quantification for a “true” 360° Feedback approach also allows us to accurately perform the kinds of evaluations and comparisons referenced in the subsequent parts of the definition, with the assumption that, in order to conduct those evaluations and comparisons, the ratings must be sufficiently reliable to justify these uses. Using traditional rating processes allows us to test for reliability, which is usually not the case for qualitatively collected data. Other AFFs also find their way into feedback studies (e.g., Kluger & DeNisi, 1996) that are later cited and used in the context of 360° Feedback critiques (e.g., Pfau & Kay, 2002) in a true “apples and oranges” application.

4. Anonymity (or confidentiality in the purest sense of the term if the data have been collected online and coded to individual demographics and unique identifiers) has long been assumed to be the foundation of 360° Feedback processes because it promotes greater honesty in responses. This is particularly the case for certain rater groups such as direct reports and peers where individuals are most at risk of retaliation. Nonetheless, we could not make a blanket statement that anonymity is required for a process to be labeled as 360° Feedback. The most obvious example of lack of anonymity is the common practice of reporting the responses of the manager (boss), matrix manager, skip level manager, or even the self as a single data point, even though the reporting rules were clearly communicated to these raters beforehand. There is also, of course, the counterargument that if data were to be collected from someone, there should be a commitment to use it if possible (without violating confidentiality and anonymity being the operative concern). Regardless, reporting results this way would not alter the fundamental character of the process to the degree that it could not be called 360° Feedback, as it still meets the two goals as stated above.
5. We found it easy to agree that 360° Feedback involves observations of behavior (as opposed to traits, attitudes, internal states, values, or task performance), including behavioral operationalizations of competencies. We believe that the rater should not be asked to report what he/she believes is going on inside a leader’s mind, often evidenced by stems that include verbs such as “understands,” “believes,” “knows,” or “considers” (Murphy & Cleveland, 1995). Task performance is often out of the line of sight of raters, not an indication of ability, and often not developable (or at least not in the same way as leadership behaviors and capability). As for competencies, skills, and knowledge, although we do not believe these should be rated directly per se, clearly they can be rated in the

context of specific behaviors exhibited that demonstrate, align, or manifest these in the workplace. Rating competency clusters (dimensions), and not the individual behaviors in the dimension, do not satisfy this requirement and create additional measurement errors (Facteau & Craig, 2001; Healy & Rose, 2003). Unfortunately, this is often seen as a means of simplification of the process (i.e., by asking people to rate only nine dimensions and not the 30–40 items that compose those nine dimensions), and we feel it is not within the guidelines of good practice.

6. The ability to conduct *meaningful comparisons of rater perceptions* both between (inter) and within (intra) groups is central and, indeed, unique to any true 360° Feedback process. This is what differentiates 360° Feedback from organizational surveys, though, in truth, the underlying mechanism for change is essentially the same and dates back to the roots of organizational development (OD) and action research (e.g., Burke, 1982; Church, Waclawski, & Burke, 2001; Nadler, 1977). This element of our definition acknowledges that 360° Feedback data represent rater perceptions that may contradict each other while each being true and valid observations. Interestingly enough this assertion is critical for ensuring the validity of 360° Feedback from a practice point of view (Tornow, 1993) despite academic research suggesting the opposite; that is, combinations of ratings data other than naturally occurring work groups (direct reports, peers, customers) might be better from a true score measurement perspective (Mount, Judge, Scullen, Systma, & Hezlett, 1998). A true 360° Feedback assessment, under our definition, must be designed in such a way that differences in rater perceptions are clearly identified and meaningful comparisons can be made between perceptions of different rater groups, agreement (or lack thereof) within a rater group, and even individual raters (e.g., self, manager) where appropriate. Thus, measures or displays of range and distribution can be important elements, along with internal norms, in an effective and well-designed 360° Feedback program (Mount et al., 1998).
7. We believe that a 360° Feedback process needs to create sustainable change, both proximal and distal (Bracken, Timmreck, Fleenor, & Summers, 2001), starting at the individual level and sometimes with the explicit expectation that group and/or organizational change will also occur. This is the case if (a) an individual's change affects others (a few or many) due to his/her influence and/or (b) if a large number of individuals participate in the 360° Feedback initiative (Bracken & Rose, 2011; Church & Waclawski, 2001b). Church et al. (2002) outline a model with case examples for how 360° Feedback can be used for driving change at the micro-, meso-, and macrolevels that aligns nicely with our approach here.

In order to help researchers and practitioners understand the distinction we are trying to make here between what we are calling 360° Feedback and AFF, we have provided some simple questions to consider:

1. Is the data collection effort quantitative and focused on observable behaviors or some other types of skills or attributes?
2. Is the feedback process conducted in a way that formally segments raters into clearly defined and meaningful groups, or are people mismatched in the groups?
3. Is the content tied to behaviors that the organization has identified as important (e.g., based on values, leadership competencies tied to business strategies, manager quality, new capabilities, etc.)?
4. Is the feedback collected (vs. invited) from an adequate number of qualified raters to establish reliability, which can vary by rater group (Greguras & Robie, 1998)?
5. Are the feedback process and the reporting designed to display quantifiable data that provide the user (i.e., the ratee, coach, manager, or HR leader) with sufficiently clear and reliable insights into inter- and intra-group perceptions?
6. Are the results shared with the employee and other key stakeholders as defined by the process up front (e.g., manager, HR, senior leadership, colleagues, customers, etc.) or kept solely for the organization? If only for the organization, how are they being used, if at all? In other words, if a feedback process has no use, then it produces information, not feedback.

Why a Review of the Field of 360° Feedback Is Needed

Even Edwards and Ewen's (1996, p. 4) definition of 360° Feedback equated it to "multisource assessment." So it should be no surprise that many versions of multirater processes have used the label of 360° Feedback, with consulting groups creating various permutations under that label (or variations such as "450 Feedback," and even "720 Feedback") that strayed significantly from the original concept and the one we propose in our definition. The practice of 360° Feedback quickly became a commodity with many attempts and many failures to capture the demand. As a result, the popularity of the concept has created more and more distance from its theoretical and research underpinnings. For example, today it is commonly referenced in other types of books on HR, talent management, and related topics (e.g., Cappelli, 2008; Efron & Ort, 2010; MacRae & Furnham, 2014) but is often done so in the context of a specific purpose rather than focused on as a methodology in and of itself. It is our objective to highlight key trends that exemplify that evolution (or devolution).

We also see 360° Feedback getting swept up in the recent wave of commentaries on performance appraisals, performance management, and ratings in general (e.g., Adler et al., 2016; Buckingham & Goodall, 2015; Pulakos, Hanson, Arad, & Moye, 2015). We agree that 360° Feedback should be part of those discussions because they have many of the same characteristics and/or contribute to the use and outcomes of those processes and can also have an integral role in their success or failure (Bracken & Church, 2013; Campion, Campion, & Campion, 2015). These discussions have also brought to the surface prime examples of professional representations of how 360° Feedback can be used to potentially improve talent management (e.g., Campion et al., 2015) as well as less responsible but highly visible portrayals of the field (e.g., Buckingham & Goodall, 2015) that we will discuss further.

What Is Going Well With 360° Feedback?

Technology

Although technology has had some negative applications (Rose, English, & Thomas, 2011), it is not debatable that technology has caused a remarkable growth in 360° Feedback utilization. In the late 1980s and early 1990s, the introduction of personal computers allowed 360° Feedback data to be analyzed and presented with increased accuracy and at a much larger volume (Bracken, Summers, & Fleenor, 1998). In that era, feedback recipients typically distributed 20 or so paper surveys to a group of raters along with a postage-paid return envelope that would be sent to a “service bureau” where the surveys would be scanned and then compiled into an individual feedback report. This was a huge improvement in the technology employed for 360° Feedback, and it allowed for significant growth in usage due to the decrease in cost and the ability to scale the process to large numbers of leaders across entire organizations.

The remaining processing challenges that the PC could not quite fix were almost completely cured by the Internet in the early 2000s (Summers, 2001). For instance, moving from PC to Internet-based processing increased typical survey response rates from 60% when paper surveys were the norm to the present day when it is common to set targets and often achieve response rates of 90% or higher.

The primary advantages of Internet-based tools relate to a remarkable improvement in connectivity and communication: Survey invitations are immediate, reminders can be sent to nonresponders on a regular basis, lost passwords can be easily recovered, and new raters can be easily added. Data quality has also improved considerably with Internet-based tools. Response rates can be monitored and deadlines extended if inadequate data are available to assure confidentiality, data entry errors are limited to the responders themselves, and open-ended comments are longer.

Technology has also created a method to enhance rater accountability, a problem cited by London, Smither, and Adsit (1997) that has been short on solutions. Beyond immediate notification of rater errors (e.g., incomplete responses), online technologies can also be designed to provide the respondent with immediate feedback on the distribution of his/her ratings and/or notification of “invalid” rating patterns (e.g., extremely high or low rating average, responses all of one value), with the possible suggestion or even requirement to change the ratings. The technology can also “nudge” feedback providers to enter a comment. Costar and Bracken (2016) report that rater feedback does cause some respondents to modify their ratings and may contribute to a reduction in leniency error across a population.

Feedback reports can be distributed rapidly and (relative to paper reports) more cheaply on a global scale. Responders can be sorted into rating categories definitively rather than relying on responders to choose the correct relationship to the feedback recipient (for example, it is very common with paper surveys that responders choose “peer” instead of “direct report” when identifying themselves). Last, technology has radically improved the options for analyzing and presenting results.

Follow-Up Support

While improved survey technologies significantly increased usage of 360° Feedback, this growth created the unfortunate byproduct euphemistically known as the “desk drop.” First used in the 1990s in the consulting industry, the “desk drop” approach to 360° Feedback involves collecting the data, compiling the report, and then essentially “dropping it” on a leader’s desk with no additional follow-up support (Church & Waclawski, 2001b; Scott & London, 2003). In short, the leader receives no assistance with interpretation, development planning, resources, or follow-up accountability and tracking mechanisms. Although prevalent in the early days of 360° Feedback implementation when the methodology was still new and the focus was more on selling the concept and implementation methods than outcomes (Church & Waclawski, 1998a), the challenge still exists in practice today. We recently heard of a 360° Feedback process owner at a large corporation admit that 27% of the more than 4,000 managers participating in the program never even downloaded their reports. Although the latest 360° Feedback benchmark research (3D Group, 2013) has indicated that the “desk drop” approach still is used by 7% of organizations, there has been a positive trend report over the last 10 years such that one-on-one coaching support has increased to upward of 70% of companies (3D Group, 2003, 2004, 2009, 2013).

Further, we know from other types of benchmark research focused on senior executive and high-potential assessment efforts (e.g., Church & Rotolo, 2013; Church, Rotolo, Ginther, & Levine, 2015; Silzer & Church, 2010)

that top development companies take full advantage of their 360° Feedback programs as well for both development and decision-making (e.g., identification and succession planning) purposes. Finally, as noted earlier, although not the focus of current research per se, the practice literature would suggest that 360° Feedback has become more integrated into ongoing HR, leadership development, and talent management processes and programs more broadly (e.g., Church, 2014; Church et al., 2014b; Efron & Ort, 2010; McCauley & McCall, 2014; Scott & Reynolds, 2010; Silzer & Dowell, 2010). Thus, while usage has increased considerably, many companies have recognized the importance of moving beyond the “desk drop,” and they continue to invest in making the review of feedback a thoughtful effort at gaining self-awareness.

Use for Development

As an early innovator and longtime advocate for the power and potential for 360° Feedback to develop leaders, the CCL has always held that feedback is owned only by the feedback recipient and should only be used to guide development (Bracken et al., 1997; Tornow et al., 1998). The good news for CCL is that their message has been, and continues to be, received: Consistently over the last 10 years, 36% or more of companies report using 360° Feedback *exclusively* for development (3D Group, 2003, 2013). Although we would argue conceptually that development *only* is a misnomer (Bracken & Church, 2013), we absolutely agree that 360° Feedback should always be first and foremost a developmental tool. This view lines up with the data that show that 70% to 79% of companies use 360° Feedback for development, though not *exclusively* for development (3D Group, 2004, 2013). It is, therefore, encouraging that many companies that are looking for a development tool and want to start off with something that feels “safe” may be more inclined to use 360° Feedback as a development only tool.

With the increase in usage for more than just development (3D Group, 2013), it appears that “only” part of the CCL position may be giving way to a “development plus” approach to using feedback results. There can be dangers associated with this assumption and approach as well, particularly if an organization decides to change course or intent in the use of its 360° Feedback program (and data) over time. The risks increase dramatically if rules are changed midstream (and even retroactively) versus applied newly going forward. A number of cases in the OD literature have been written about such challenges (e.g., Church et al., 2001), and there are ethical, cultural, and even potential legal implications to such decisions. This is why purpose is so important in a 360° Feedback process design. If there is an intent to change from development only to decision making (in a more formal, process driven manner) over time, the content, system, tools, and communications must be designed up front to accommodate this transition.

We firmly believe that there are many situations where 360° Feedback should *not* be used for decision making due to a lack of readiness, climate, or, frankly, lack of support. For instance, Hardison et al. (2015) reported all branches of the U.S. military use some form of 360° Feedback, but none of them use it for evaluation purposes. They recommended that the military not adopt 360° Feedback for use in decision making moving forward but continue with their use for leadership development.

Sometimes feedback processes are introduced as “development only” and then evolve into systems that support decision making. PepsiCo introduced the upward feedback tool (MQPI, or Manager Quality Performance Index) in 2008, which was specifically divorced from 360° Feedback but aligned to the same process to enable direct input on manager behaviors as rated by direct reports in their performance management system (Bracken & Church, 2013; Church, Rotolo, Shull, & Tuller, 2014a). As part of the launch of this feedback process, the tools were clearly separated by purpose. One was linked to leadership development, and the other, directly to the people results component of performance. For the first year of the MQPI, in order to enhance its acceptance, the tool was rolled as a “pilot” where results were delivered to managers but not to their bosses above them or even HR. In other words, Year 1 represented a free feedback year and level setting exercise on the new behaviors. This promise was maintained, and the data were never released. In 2009, the following year, the results were used to inform performance ratings and have been used going forward. Today, the program remains highly successful, with significant pull across all parts of the organization (and an annual cycle of ~10,000 people managers). Many individuals point to that free pilot year and the level of transparency as key enabling factors.

Use for Decision Making

When properly implemented, 360° Feedback has significant value that can extend well beyond self-driven development. Increasingly, 360° Feedback has been used for a wider and wider range of decisions beyond development, including performance management, staffing, promotions, high-potential identification, succession planning, and talent management. For instance, whereas only 27% of companies reported using 360° Feedback for performance management in 2003 (3D Group, 2003), by 2013 the number of companies using it for performance management jumped to 48% (3D Group, 2013). As we will note in more depth below, we believe that the “debate” over proper use of 360° Feedback (Bracken et al., 1997; London, 2001) is over. In addition to the benchmark data noted above, 360° Feedback, as noted earlier, has also become one of the most used forms of assessments for both public and private organizations (Church & Rotolo, 2013; United

States Office of Personnel Management, 2012). Therefore, our efforts should be directed toward finding ways to improve our methodologies and not discussing whether it “works” or not (Smither London, & Reilly, 2005).

Forms of 360° Feedback are being used increasingly for certification of occupations such as that of physician in countries including Great Britain and Canada (Lockyer, 2013; Wood, Hassell, Whitehouse, Bullock, & Wall, 2006). Using studies from those countries, Donnon, Ansari, Alawi, and Violato (2014) cite 43 studies with feedback from physician peers, coworkers, patients (and families), and self-assessments demonstrating acceptable psychometric properties for use in performance assessment. Ferguson, Wakeling, and Bowie (2014) examined 16 studies of the effectiveness of 360° Feedback evaluations among physicians in the United Kingdom and concluded that the most critical success factors were the use of credible raters, inclusion of narrative comments, and facilitation of the feedback to influence how the physician responds to the feedback, leading to its acceptance and full use. It is also interesting to see a renewed discussion of threats to the validity of these work-based assessments as they affect physician careers, including opportunity to observe, rater training, assessment content (items), and even rating scales (Massie & Ali, 2016).

Moving Beyond Individual-Level Data

Another way 360° Feedback has evolved is by pushing beyond individual-level data. Although I-O psychologists often look at 360° Feedback as individual-level assessments only, the tool is a highly valuable asset for shaping, driving, and evaluating organizational change (Bracken & Rose, 2011; Church et al., 2014a, 2014b; Church et al., 2001). For example, 360° Feedback content has been used to articulate aspirational, strategically aligned goals throughout the leadership team and then used to evaluate and develop leaders toward those strategically aligned goals (Rose, 2011). Other examples include 360° Feedback's role in diversity and culture change efforts (e.g., Church et al., 2014a), mergers and acquisitions (e.g., Burke & Jackson, 1991), and a host of other applications (e.g., Burke, Richley, & DeAngelis, 1985; Church, Javitch, & Burke, 1995; Church et al., 2002; Nowack & Mashih, 2012).

These are positive signs of the evolution of 360° Feedback as a process and its transition from fad to a well-respected and integrated tool for OD. Along with our efforts to promote best practices in the use of 360° Feedback in support of creating sustainable change (Bracken & Church, 2013; Bracken & Rose, 2011; Bracken, Timmreck, & Church, 2001), some other notable publications have contributed to the ongoing healthy evolution of 360° Feedback practices. Champion et al.'s (2015) rejoinder to Pulakos et al. (2015) makes some wide-ranging points addressing many of the major

talking points that continue to make 360° Feedback a valuable choice for individual and organizational change.

What Is Wrong With 360° Feedback Today?

Although it is clear that the application of 360° Feedback-like methods have proliferated in the past 25–30 years since the term was introduced, reflection on the very same practice and research noted above shows considerable variability in the ways in which the methodology is being implemented. In fact, it would appear as though the approach has become so popular and commonplace as to have devolved from what some might consider being an aligned practice area at all. Similar to some of the issues raised with organizational engagement survey work (e.g., Church, 2016), the popularity and subsequent rise of 360° Feedback may also be its demise. From our perspective there are a number of trends we have observed in practice and in the practice literature (along with the glaring absence of academic papers focused on research agendas) that suggest both the science and the practice of 360° Feedback has devolved. Listed below are several areas we have identified that need to be addressed in the field, with some specific recommendations of our own at the conclusion.

Skipping the Purpose Discussion

In our experience, a large percentage of the challenges that occur around 360° Feedback processes today and in the past, whether within the academic, practitioner (internal or external consultant), or the end-user (e.g., HR or line manager) community, can be traced to a lack of clarity as to purpose. It is a simple premise and one that is consistent with any business plan or change management agenda. It also ties to any organizational change agenda (Burke & Litwin, 1992). A lack of clarity of purpose will result in misunderstandings among everyone involved. Sometimes the issue is that the goals of the 360° Feedback effort were never fully defined. In other cases, even when it might have been defined, those leading and implementing the design of the system failed to take appropriate measures to communicate expectations and outcomes to ratees and key stakeholders in the process. Although this has been discussed at length in the literature at the maturation of the practice area (e.g., Bracken, Timmreck, & Church, 2001; Church et al., 2002), the message seems to have been lost by many.

Although the specific purpose of a 360° Feedback program may vary (e.g., tied to a leadership program, culture change effort, or an executive assessment suite) at the broadest level, it is possible to classify most efforts into a group of broad categories. In [Table 1](#), we offer a classification of purposes.

Note that “development only” is only one of many purposes. Even so, our experience over the years with all types of organizations has shown that

Table 1. Typical Uses of 360° Feedback Processes

Purpose	Common characteristics	Examples
Assessment only	<ul style="list-style-type: none"> • KSA focused • Shorter instruments • Ratee may not receive results/report • May not have developmental expectations/requirements • May be paired with formal performance evaluation process 	<ul style="list-style-type: none"> • Physician credentialing in Canada and UK • Used to support performance management ratings • ID of bottom 10% for remedial action or termination • Selection for programs (e.g., Hi-Po)
Development only	<ul style="list-style-type: none"> • Common for $N = 1$ administrations • Often OTS instruments • Often totally confidential (report to leader only, maybe coach as well) • Pre-/postintervention (coaching, program) 	<ul style="list-style-type: none"> • Recommended by major OTS providers • Often integrated into LD programs, training • Pre-/postcoaching effectiveness measure
Assessment and development	<ul style="list-style-type: none"> • Often used for groups of leaders ($N > 1$) • Custom content more common than OTS • May be used for decision making (e.g., PMP, Succession, Hi-Po) 	<ul style="list-style-type: none"> • Increasing number of organizations (see 3D Group, 2003, 2004, 2009, 2013) • PepsiCo (Bracken & Church, 2013)
Organization change	<ul style="list-style-type: none"> • Culture change driven by leader behaviors • Often tied to values • Typically requires census administration across organization or unit • See Bracken and Rose (2011) for examples 	<ul style="list-style-type: none"> • Startups creating a culture (e.g., spinoffs) • Startups transitioning to larger organization (growth, acquisition) • Organizations with new leadership at the top • New competency/ leadership model • Culture change due to context changes
Program evaluation	<ul style="list-style-type: none"> • Focus is on group performance • Individuals may or may not get results or be held accountable for change 	<ul style="list-style-type: none"> • Leadership development program evaluation • Coaching effectiveness measure • Overall 360° process evaluation

Note. KSA = knowledge, skills, and abilities; OTS = off the shelf; PMP = performance management processes; Hi-Po = high potential; LD = leadership development.

even “development only” processes typically produce feedback that is used in some sort of decision-making process (e.g., as part of a broader talent management effort, in talent review or succession planning discussions, in performance management processes, or to determine developmental experiences, including training; Bracken & Church, 2013). Research on the talent management and assessment practices of top development companies, for example, has noted that, aside from development only, 360° Feedback was the most commonly used tool (followed closely by personality tests and interviews) for high-potential identification, confirmation of potential, and succession planning purposes (Church & Rotolo, 2013). When combined across responses, though, roughly 60% of companies were using the data for both development and decision making with either high potentials or their senior executives. Only a third were claiming “development only” practices. Even if the feedback is simply being used to inform who receives access to further training and other development resources that others may not receive, that is, in itself, a differentiated decision with respect to talent (Bracken & Church, 2013) that impacts careers. Unless the individual receiving the feedback never (and we mean literally never) shares his or her data or any insights from that data with anyone inside the organization, it is not “development only.” They simply cannot get the results out of their frame of reference on the individual once they have been seen. In virtually every situation where organizations fund these types of 360° Feedback efforts, it is not realistic to expect “development only” feedback to remain totally private and confidential, thereby possibly affecting ratees differently and influencing the perceptions of others (Ghorpade, 2000). Thus, from our perspective, organizations should focus on the true purpose, align to that end state, and, finally, design and communicate transparently accordingly. This speaks to the importance of the concept of transparency today in how we approach of I-O tools and programs.

Making Design Decisions That Don't Align With Purpose

We strongly believe that design and implementation decisions in 360° Feedback should be derived from a purpose statement agreed on by all critical stakeholders (including the most senior leader sponsoring the effort). Of course, that will be difficult to perform if the purpose definition step is skipped or assumed to be understood.

In this article, we refer to many design and implementation factors that have been either proposed or proven to determine the ability of a 360° Feedback system to deliver on its stated objectives (purpose). Smither et al. (2005) created a list of eight factors that can affect the efficacy of creating behavior change, but some of those are less within the individual's and organization's direct control than others (e.g., personality of ratee, feedback

orientation, readiness for change). In our recommendations section, we will list other design factors that have been shown to affect the likelihood of achieving a system's purpose and should be included in research descriptions and lists of decisions to be made when designing and implementing a 360° Feedback process. Consider the choice of rating scale as one example of a design decision that is probably treated not as a true decision but more as a default carried over from employee surveys to use the *agree/disagree* format (Likert, 1967). The 3D Group benchmark research confirms that the 5-point Likert format is by far the most commonly used response scale (87% of companies) in 360° Feedback processes (3D Group, 2013). Importantly, 360° Feedback data have been shown to be very sensitive to different rating scales (Bracken & Paul, 1993; English, Rose, & McLellan, 2009). This research illustrates that rating labels can quite significantly influence the distribution, variability, and mean scores of 360° Feedback results. Whereas, in nearly all cases where the rating format will be a 5-point Likert scale, it would be problematic to combine results from studies using different response formats (Edwards & Ewen, 1996; Rogelberg & Waclawski, 2001; Van Velsor, 1998).

An all too common situation is to let purpose and design features (such as response scale) float “free form” in the minds of raters and ratees through a lack of definition and training. Such is the case in the vast majority of 360° Feedback systems where rater training continues to be the exception, not the rule (3D Group, 2003, 2004, 2009, 2013). The Likert (*agree/disagree*) scale provides no frame of reference for the rater for guiding the rating decision (i.e., ipsative vs. normative vs. frequency vs. satisfaction vs. etc.). Assuming that there is little training or guidance provided, the rater is left to his/her own internal decision rules. Also, unfortunately, the ratee is similarly left with no framework for attempting to interpret the feedback. The good news there is that it should force the ratee to actually have to discuss the results with the raters; no matter how hard we try to design a sound feedback process, using anonymous surveys still creates an imperfect communication method that requires that the ratee actually discuss his/her results with the feedback providers to confirm understanding (Goldsmith & Morgan, 2004).

In the context of discussing the importance of the alignment of design and implementation decisions and purpose, we should mention Nowack and Mashih's (2012) engaging review of the current state that provides positions on 15 common questions that reflect a mix of issues that are raised by designers and implementers (e.g., rater selection, report design) and coaches (e.g., reactions, cross-cultural factors, and neuroscience) in the area. Although we appreciate the content discussed, in our opinion, their endorsement of leaving final decisions to the user reflects the ongoing challenges in the field that we are highlighting of having no one “right” answer to any question. In many ways, this “user friendly” approach is very common among practitioners

who are faced every day with a new request to morph a typical 360° Feedback process into a new mold to address a new need. This trend, although helpful at some level, can also enable a drift among practitioners who find it increasingly difficult to identify and stay true to any set of best practice guidelines. We would prefer to draw some specific lines in the sand. With this article, we are attempting to do so.

Making Generalizations Regarding the Effectiveness of 360° Feedback

One of the more egregious affronts to good science in our field is when poorly designed articles that make broad generalizations about the effectiveness or (more often) the ineffectiveness of 360° Feedback processes are widely cited with little regard to the veracity of the conclusions given the research methodology (e.g., Pfau & Kay, 2002). Some studies have become widely cited outside the I-O community despite questionable relevance (e.g., Kluger & DeNisi, 1996). Even well-researched studies, such as the Smither et al. (2005) meta-analysis, are easily quotable regarding results that we feel need to be challenged, given the difficulty of meaningfully combining data from studies with so much variation in design. Even so, methodological decisions can affect conclusions, such as reporting criterion-based validity coefficients that combine perspective groups and not differentiating sources of performance measures (Bynum, Hoffman, Meade, & Gentry, 2013).

Here is an example of what we know happens all too often. Pfau and Kay (2002) cite Ghorpade (2000) by saying,

Ghorpade also reported that out of more than 600 feedback studies, one-third found improvements in performance, one-third reported decreases in performance and the rest reported no impact at all.

The problem is that Ghorpade is actually citing Kluger and DeNisi (1996), whose meta-analysis used search values of “feedback” and “KR” (knowledge of results) and did not specifically use studies that were actually examining multisource (360°) feedback (according to titles). Although Pfau and Kay’s misattributed quote of Ghorpade did say only “feedback,” both their paper and Ghorpade’s are specifically about 360° Feedback, so the citation of Kluger and DeNisi (1996) is misplaced but very quotable by unsuspecting/unquestioning readers (or worse, those who are simply looking for a headline to support their position). We would say the same of Nowack and Mashihhi’s (2012) citation of the same article in their review of 360° Feedback research in terms of being out of place in this context.

Misrepresentation of Research

In a recent, highly visible publication, Buckingham and Goodall (2015) resurrect Scullen, Mount, and Goff (2000) in the cover article regarding

Deloitte's revamp of their performance management system. In the reference to Scullen et al. (2000), they note, "Actual performance accounted for only 21% of the variance" (Buckingham & Goodall, 2015, p. 42). This 21% figure, as well as the 55% attributed to rater characteristics, continues to be repeated in various contexts in blogs and articles as evidence of the ineffectiveness of the 360° Feedback to account for "true" performance variance (e.g., Kaiser & Craig, 2005). Ironically, if a selection instrument demonstrated this level of prediction (an uncorrected correlation of about .46), it would be considered a highly predictive assessment and would compare admirably with the best selection tools available.

Note that the measure of employee performance used by the Scullen et al. (2000) was not an independent performance measure but was instead defined by the common rating variance from the 360° Feedback ratings. If we took the time to truly explain how this "works," we might characterize this performance measure as "synthetic," that is, not a true independent measure of a person's performance. If 360° Feedback is going to be studied as a predictor of leader performance, we would like to see independent, reliable measures of performance used as criteria. It is safe to say that the users of our research (i.e., organizations) would also define "performance" as being actual activity that is observable and measurable.

Scullen et al. (2000) contains multiple design elements that limit its generalizability, but let's stop and revisit one design factor that is important in both the performance management and the 360° Feedback realms (Ghorpade, 2000) and one this study also mentions: rater training. In the performance management world, the training of managers on conducting performance reviews is a best practice (whether followed or not) and could be a factor in legal defensibility (American National Standards Institute/Society of Human Resource Management, 2012). In the 360° Feedback world, it is extremely rare (3D Group, 2013).

Interestingly, in explaining their findings, Scullen et al. (2000) suggest that rater training can significantly impact the findings. In examining differences between rater groups, Scullen et al. found that supervisor ratings accounted for the most variance in performance ratings (38%) compared with all other rater groups. The authors attributed that finding to the likely fact that supervisors have more "training and experience in rating performance" than the other raters (p. 966). Their data and conclusions on this point indicate (a) that the background (e.g., experience) of the rater will probably affect the correlation with performance ratings and (b) that correlations between 360° Feedback scores and performance ratings are also enhanced by training as a rater, thereby acknowledging methodological effects that are not otherwise considered in reporting the overall results but are controllable.

The main point is that 360° Feedback as a performance measure has requirements in its design and implementation that are unique and complex, and each of these can profoundly influence outcomes in both research and practice (Bracken & Church, 2013; Bracken & Rose, 2011; Bracken & Timmreck, 2001; Bracken, Timmreck, Fleenor, & Summers, 2001). One reason for that is that 360° Feedback is an assessment process where the data are generated from observation by others, as in assessment centers. Unlike most traditional assessments, they are not self-report measures; when self-ratings are collected, they are usually not included in the overall scoring. Therefore, reliability and validity considerations go far beyond just the instrument itself (though the instrument is also an important determinant, of course).

To criticize any given 360° Feedback process or set of practices as inadequate for use in decision making and then use 360° Feedback methodologies in the design and implementation of those processes that are themselves deficient (e.g., no rater training) is hypocritical. We actually agree that 360° Feedback should not be used for decision making when designed and implemented as described by Scullen et al. (2000), and they were not explicitly used for that purpose. But the reported overlap with performance could easily have been improved using different implementation decisions. The Scullen et al. (2000) study was published in the preeminent journal in our field. And we should also be vigilant as to how our studies are used for certain agendas, including claims that 360° Feedback data are not sufficiently reliable to contribute to personnel decisions. We categorically continue to believe that 360° Feedback data are superior to single source (supervisory) ratings when collected correctly (Edwards & Ewen, 1996; Mount et al., 1998).

Ignoring Accountability

Those who lament the lack of perceived effectiveness in 360° Feedback processes must first define what success is. Bracken, Timmreck, Fleenor, and Summers (2001) define it in terms of sustained behavior change. That is a measurable outcome, and we know that lack of accountability is one of its many barriers (Goldsmith & Morgan, 2004). It is enlightening and discouraging to go back and read the seminal article on this topic, “Accountability: The Achilles’ Heel of Multisource Feedback” (London, Smither, & Adsit, 1997). This article could have been written yesterday because little has changed since it was first published.

London et al. (1997) cite three main needs to establish accountability: (a) rater accountability to use results, (b) rater accountability for accuracy and usefulness, and (c) organizational accountability for providing the resources that support behavior change. They also offer recommendations for improving accountability for all those stakeholder groups, most of which have been

underachieved. Let's look at the evolution/devolution that we observe occurring in each area.

As for ratee accountability to use results, we do see an evolution in accountability being built in to 360° Feedback processes through integration into HR systems, particularly performance management processes (PMP). This makes sense in part given the documentation aspects required (interestingly enough it is also one of the arguments against the removal of ratings and reviews entirely). The 3D Group benchmark studies (3D Group, 2003, 2004, 2009, 2013) have shown a persistent positive trend in organizations reporting use of 360° Feedback for decision-making purposes. In addition, HR managers are much more likely than ever before to be given copies of 360° Feedback reports (3D Group, 2013). Church et al. (2015), for example, have reported that 89% of managers in their survey of top development companies share some level of assessment data (360° being the most commonly used tool) with managers, and even 71% of the senior leadership of these organizations also have access to some type of detail or summary.

There is implicit accountability when managers ("bosses") receive copies of the feedback report, which is also on the rise. We believe that involving bosses in the entire 360° Feedback process is a best practice that should ideally include selecting raters, discussing results, planning discussions with other raters, and prioritizing development. Many practitioners see the boss as the most important figure in an employee's development, and preventing them from being involved in the process certainly impedes their ability to support and create development activities. One additional benefit of requiring boss participation in all phases lies in standardization that, in turn, removes claims of real and perceived unfairness toward ratees. From an OD (e.g., Church & Waclawski, 2001b) and even talent management (Effron & Ort, 2010) perspective, other authors would agree with these recommendations as well. Of course, sharing results with a boss (or senior executive even if just a summary) reinforces our argument about 360° Feedback not being for "development only."

On the other side, clearly, we see a form of devolution (or at least stagnation) when 360° Feedback processes are designed with the requirement or strong recommendation that ratees not share their results with others (e.g., Dalton, 1997). A powerful source of accountability comes from sharing and discussing results with raters, especially subordinates and managers. Similar to telling friends/family of New Year's resolutions, making commitments to actions that are made public create an opportunity for supporters to help with achieving the goal(s). In addition, in organizations where 360° Feedback serves as a major or the sole component of evaluating behavior change following a leadership assessment and/or developmental intervention, the sharing of results along with resulting action plans becomes critical.

Although sharing data can have significant benefits, we acknowledge that sharing the full feedback report can be anxiety provoking in organizational cultures with low levels of trust or in organizations experiencing high levels of change or stress (e.g., layoffs, high-level leadership changes, mergers and acquisitions, industry-wide disruption; Ghorpade, 2000). As with organizational survey results, the context of the organizational system is key, along with prior history with feedback tools in general, including whether they have been misused in the past (Church & Waclawski, 2001a). Even in these environments, however, and in cases where sharing results is not mandatory, feedback recipients should be made aware of how they could benefit from sharing results in a healthy and productive manner. Further, program managers and organizational leadership at all levels would be well served by recognizing that the best case is to move their organization to a point where 360° Feedback results are shared and performance-related behavior can be discussed openly and honestly.

Perhaps the single most powerful evidence for the value in following up with raters is provided by Goldsmith and Morgan (2004). Having accumulated tens of thousands of data points from companies in all geographies and industries, the findings are compellingly clear that follow-up with raters is highly related to perceived behavior change (or lack thereof) from the feedback providers. Lack of follow-up often leads to negative behavior trends. Ignoring or discounting this apparently powerful process factor is another example of devolution. This same finding has been well-documented in survey action planning research as well (Church & Oliver, 2006; Church et al., 2012). The impact is in sharing and doing something with the results from the feedback process. Not sharing results back from the process or even sharing but having respondents feel that nothing has been done to take action will lead to negative attitudes over time.

Rater accountability, however, can be a difficult requirement when anonymity to the focal participant (or technically confidentiality to the process if the feedback surveys are collected via some externally identified process) is a core feature of almost all 360° Feedback processes. Ratings from direct reports, peers, and often customers/clients are almost always protected by requiring an aggregate score based on ratings from three or more raters. Although this in theory allows for greater honesty, it can also provide an excuse to over- or underrate for various political or self-motivated reasons. These are in fact some of the most commonly cited reasons by laypeople (and sometimes HR professionals) in organizations for why 360° Feedback is inherently flawed and can be manipulated (as if performance management or high-potential talent ratings do not suffer from the same political and psychometric forces). Either way, the outcomes of a 360° Feedback process, whether positive or negative, come back to affect the raters in multiple ways.

In this context, we believe that raters will feel more accountable if they see that their input is valuable to both the individual and the organization. At a basic level, it probably will also determine whether they respond at all and, if they do, their propensity to be honest, depending on whether honesty was rewarded or punished in prior administrations. As noted earlier, technology also provides us with means to integrate rater feedback into the survey taking process to create another form of direct or indirect accountability (Costar & Bracken, 2016).

One example of a subtle message to raters of a lack of accountability is the old proverb that “feedback is a gift.” The message seems to be that once the rater hits the “submit” button on the computer, his/her responsibility is over. Instead of a gift, we would prefer people begin to say, “feedback is an investment,” an investment whose quality will affect the rater because he/she will continue to work with the ratee, and the development of the ratee will benefit the rater, the team, and the organization. In that way, the investor can expect a return on his/her investment. It has both short- and long-term implications, and people in organizations need to see that. This is where positioning (and purpose) can help once again.

As for organizational accountability, we have already mentioned one way that the organization answers the call for accountability at that level—that is, providing “coaching” resources to ratees, even if only in the form of one-time feedback facilitators. Another major form of support comes from its investment (or lack thereof) in the 360° Feedback system itself, including the technology, associated training, communications, staffing, and help functions, along with communicating clear expectations for each stakeholder (leadership, raters, ratees, managers; Effron, 2013). In some cases, this can involve the additional integration of 360° Feedback into broader leadership development and talent management efforts as well (e.g., Oliver, Church, Lewis, & Desrosiers, 2009).

In addition, managers, as representatives of the organization, need to be held accountable for helping the feedback recipients use their results constructively and provide the guidance and resources required for useful follow-through on development plans. The fact is that managers are often neither prepared nor rewarded for performing the role of “coach” in the context of 360° Feedback processes, let alone day-to-day performance management and employee development (Buckingham & Goodall, 2015; Pulakos et al., 2015).

How Can We Facilitate Evolution and Circumvent Devolution of 360° Feedback?

In this section we offer a set of recommendations for researchers, practitioners, and end users to enhance our understanding and the effectiveness of 360° Feedback going forward.

For Researchers

Fewer, better segmented meta-analyses. Although we acknowledge the major contributions of studies such as Smither et al. (2005) that bring together large volumes of research on this topic, we believe that limitations in meta-analysis as a research tool may actually impede the acceptance of 360° Feedback processes by creating cryptic storylines and masking success stories. Meta-analyses are limited by the quality and quantity of data in the primary research. But given the complexity of organizational system dynamics involved in 360° Feedback applications, meta-analyses of 360° Feedback processes report findings that are watered down by effects, direct and indirect, known and unknown, which are nonetheless reported out as “fact.” Although the possible remedies are many, one would be to segment the meta-analyses by common, major independent variables (e.g., purpose)

Here is a list of some of the design factors that have likely effects in the outcome of any study, though they may not be the primary independent or moderator variables being examined:

- Purpose of the process (e.g., leader development, leader assessment, input into HR process(es), organizational culture change, performance management; Bracken et al., 1997; Church et al., 2001, 2014b; DeNisi & Kluger, 2000; Kraiger & Ford, 1985; Smither et al., 2005; Toegel & Conger, 2003)
- Geographic region (Atwater, Wang, Smither, & Fleenor, 2009; Hofstede & McRae, 2004)
- Definition and communication/transparency of desired outcome (e.g., behavior change; improved job performance; training decisions; Antonioni, 1996; Atwater, Waldman, Atwater, & Cartier, 2000; Nowack, 2009)
- Off-the-shelf or custom instrument design (Mount et al., 1998; Toegel & Conger, 2003)
- Item format and content type (e.g., standard items, traits, category level ratings; Healy & Rose, 2003)
- Rating scale detail (English, Rose, & McLellan, 2009; Heidermeier & Moser, 2009)
- Rater selection method(s) (including rater approval methodology and rules e.g., HR, boss, other; Farr & Newman, 2001)
- Rater composition (types, proportions, limits, requirements—all direct reports vs. select direct reports, functional differences between line and staff, etc.; Church & Waclawski, 2001c; Greguras & Robie, 1998; Smither et al., 2005; Viswesvaran, Schmidt, & Ones, 1996)
- Rater training (Y/N; if Y, type, evaluation of said training; Antonioni & Woehr, 2001; Woehr & Huffcutt, 1994)

- Response rates (impact on reliability and validity of the data; Church, Rogelberg, & Waclawski, 2000)
- Participant selection (e.g., random, training program, suborganization [department], special group [e.g., high potential; Hi-Po], total organization, voluntary vs. required; Mount et al., 1998)
- Follow-up with raters (discouraged/prohibited, encouraged, required; Goldsmith & Morgan, 2004; London et al., 1997; Walker & Smither, 1999)
- Who receives reports (self, manager, coach, OD/talent management professionals, HR business partners, second level leaders/business unit CEOs, etc.; Dalton, 1997; DeNisi & Kluger, 2000)
- How reports are distributed—push versus pull methodology (including the percentage of leaders who actually looked at their reports)
- Whether a coach is provided (part of coaching engagement, internal or external, feedback review only (one time), none; Smither, London, Flautt, Vargas, & Kucine, 2003; Smither et al., 2005)
- Other methods (beyond coach) to sustain behavior change (e.g., follow-up mini surveys, Goldsmith & Morgan, 2004; cell phone “nudges,” Bahr, Cherrington, & Erickson, 2015)

Less self-versus-other research. Self-ratings versus non-self-ratings research is interesting but not very useful when using 360° Feedback across multiple rates at one time. The main reason we see further research as not being particularly useful is that self-ratings are primarily beneficial as a reflection point for individuals receiving feedback. Self-scores can help highlight differences in perception (e.g., blind spots), but they are the least reliable of all rating sources for evaluating behavior and guiding decisions about development, placement, or fit (Eichinger & Lombardo, 2004; Greguras & Robie, 1998). Large numbers of “blind spot” or “hidden strengths” often have much less to do with self-perceptions than self-agendas. Self-ratings are primarily beneficial to (a) familiarize participants with the content, (b) show a sign of commitment, and (c) provide an opportunity to reflect on differences in perspective. Additionally, when the feedback is only for one person and that person has a coach, the self-evaluation can be a good basis for dialogue. But self-ratings across large populations have less value. Thus, although practically useful in some 360° Feedback processes, self-ratings are less valuable for research.

More accurate rating source research. Although there is considerable research available discussing trends in rater group differences (e.g., Church, 1997, 2000; Church & Waclawski, 2001c; Furnham & Stringfield, 1994; Harris & Schaubroeck, 1988; J. W. Johnson & Ferstl, 1999), these studies presume that these groups can be meaningfully combined. In our experience, the peer

group in particular is far from a homogeneous group, and the way individuals are selected into these groups can have a meaningful impact on scores. For instance, often “peers” are selected who are project team members, internal customers, or internal business partners but who do not report to a single manager. Although we don’t believe that it is particularly helpful from a research or practice standpoint to create “new combinations” of rater groups across predefined categories based on variability (as might be suggested by Mount et al.’s [1998], findings), any research purporting to compare ratings between groups across many individuals should (a) clearly operationalize group membership and (b) document efforts to assure accuracy when determining group membership.

More research on how to create sustainable change. What are the characteristics of 360° Feedback processes that result in measurable distal change after the “event” and/or coach have gone away (e.g., Walker & Smither, 1999)? What is the role that personality (Berr, Church, & Waclawski, 2000; Church & Waclawski, 1998b; Walker et al., 2010) plays, if any, and what about feedback climate (Steelman, Levy, & Snell, 2004)? What can we learn and borrow from other disciplines that also have to wrestle with the challenge of sustaining behavior change, such as recidivism for juvenile offender programs that are testing follow-up technologies (Bahr et al., 2015)?

For Practitioners

Be clear about purpose—and recognize the impact of each design choice on your intended uses and outcomes. The value of a purpose statement goes far beyond just design; it begins with engaging stakeholders and communicating to participants (raters, ratees, users) to create clarity and commitment. It is the first step in any effective feedback model (Church & Waclawski, 2001b). We have typically seen purpose described in terms of behavior change, and 360° Feedback creates a methodology for measuring achievement of that goal (or lack thereof). Creating sustainable behavior change, at the individual and organizational levels, will typically require some other “purpose” statements such as how the data will be used (e.g., integration into HR systems) that will have implications for design and implementation decisions. In some cases, these may be part of a broader talent strategy that includes 360° Feedback as a key component (e.g., Efron & Ort, 2010; Silzer & Dowell, 2010). Either way, whether stand alone or integrated, having a clearly defined and transparent purpose is critical.

Hold leaders accountable for change and feedback providers accountable for accuracy—doing so is an essential driver of value. When asked, “what is the most common reason for the failure of a 360° Feedback process?” we would have to say it is lack of follow-through or, in other words, lack of accountability. One of the classic symptoms of such a process is the “desk drop” of

feedback reports that are, in turn, dropped into the various shapes of files that we discussed earlier. This definition of accountability—that is, leader follow-through—is the typical response and the easy target for blame. But we encourage you to go back to the London et al. (1997) article that places equal emphasis on the role of the rater and the organization in fulfilling their own roles in the equation. Rater accountability, in particular, is a ripe area for discussion and solutions. To the extent possible, formal mechanisms and integration points in the broader HR system also support accountability as well. As we noted earlier, if we are able to acknowledge that 360° Feedback does not exist for “development only” (i.e., because others in the organization eventually see the data in some form, and decisions of some type are being made, even if “only” about development), then it may be easier to design and implement formal tracking and outcomes measures than in the past. In our opinion, fear of sharing the data with the wrong people has been an excuse, at times, for lack of accountability as well.

Challenge the misuse of our research; herald its contributions. We hope that it has been useful to go back to some original source documents to track down the bases for some numbers that appear in the popular literature (e.g., Buckingham & Goodall, 2015). All too often we cite secondary or tertiary (or beyond) sources, and the actual findings lose some important qualifiers or caveats. Every field of research and practice has its list of those qualifiers, and we have tried to highlight some of the most critical. Also, consumers outside of our field are unlikely to even begin to understand just how hard it is to reliably account for even 30% of the variance in some performance measure, so let us not belittle that achievement but continue to try to make it even better!

For Researchers, Practitioners, and End Users

The name “multisource” (and its cousin “multirater”) has run its course. . . . Let’s call it “360° Feedback” when that’s what it is and be clear about what we mean. As mentioned earlier, when two of us (Bracken, Timmreck, & Church, 2001) worked with Carol Timmreck to pull together *The Handbook of Multi-source Feedback*, we chose not to use the term “360° Feedback” in deference to the copyright holder. Clearly the term has become widely used, perhaps too much so. The definition we provide is clear in requiring questionnaire-based anonymous feedback collected from multiple perspective groups in quantifiable form. We propose that other versions of feedback (qualitative, single perspective, etc.) use other labels such as AFF so as to not confuse users.

Establish a governing body/set of standards. Levy, Silverman, and Cavanaugh (2015) proposed that there could be value in the establishment of some group that, as one function, would be a clearinghouse for research

collected by organizations on performance management. Some modest attempts at creating a consortium for companies with 360° Feedback systems have been attempted in the past, such as the Upward Feedback Forum in the 1990s (Timmreck & Bracken, 1997) and, more recently, the Strategic 360 Forum. Consortium efforts have been challenging to sustain due to a lack of common core content, such as that used by the Mayflower Group (R. H. Johnson, 1996). As this article has emphasized, the ability to pool knowledge bases has challenges that go far beyond just common data due to the impact of methodological differences. But it may be research on those methodological factors themselves that creates the greatest opportunities.

Conclusion: If We Think We've Made Progress

In the conclusions section of their powerful article, London et al. (1997) provided a list of signs that multisource (their term) feedback has become part of an organization's culture. Given the date of the article, it is amazingly prescient, aspirational, and discouraging at the same time:

- a) Collect ratings at regular intervals; b) use feedback to evaluate individuals and make organizational decisions about them; c) provide feedback that is accompanied by (norms); d) encourage or require raters, as a group, to offer ratees constructive, specific suggestions for improvement; e) encourage ratees to share their feedback and development plans with others; f) provide ratees with resources . . . to promote behavior change; g) are integrated into a human resource system that selects, develops, sets goals, appraises, and rewards the same set of behaviors and performance dimensions . . . ; and h) track results over time. (London et al., 1997, p. 181)

Here, almost 20 years later, we would be hard pressed to come up with a list of any significant length of organizations that have accomplished these things for a sustainable period (though a few do exist). Maybe some of you will say that it is because it is just not a good idea or practical. Maybe there are success stories of which we are not aware. Maybe we are not trying hard enough.

References

- Adler, S., Campion, M., Grubb, A., Murphy, K., Ollander-Krane, R., & Pulakos, E. D. (2016). Getting rid of performance ratings: Genius or folly? A debate. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 9(2), 219–252.
- American National Standards Institute/Society of Human Resource Management. (2012). *Performance management*. (Report No. ANSI/SHRM-09001-2012). Alexandria, VA: SHRM.
- Antonioni, D. (1996). Designing an effective 360-degree appraisal feedback process. *Organizational Dynamics*, 25(2), 24–38. doi:10.1016/S0090-2616(96)90023-6
- Antonioni, D., & Woehr, D. J. (2001). Improving the quality of multisource rater performance. In D. W. Bracken, C. W. Timmreck, & A. H. Church (Eds.), *The handbook of multisource feedback* (pp. 114–129). San Francisco, CA: Jossey-Bass.
- Atwater, L., Waldman, D., Atwater, D., & Cartier, P. (2000). An upward feedback field experiment: Supervisors' cynicism, reactions, and commitment to subordinates. *Personnel Psychology*, 53, 275–297.

- Atwater, L., Wang, M., Smither, J., & Fleenor, J. (2009). Are cultural characteristics associated with the relationship between self and others' rating of leadership? *Journal of Applied Psychology, 94*, 876–886.
- Bahr, S. J., Cherrington, D. J., & Erickson, L. D. (2015). The evaluation of the impact of goal setting and cell phone calls juvenile rearrests. *International Journal of Offender Therapy and Comparative Criminology*. Advance online publication. doi:10.1177/0306624X15588549
- Berr, S., Church, A. H., & Waclawski, J. (2000). The right relationship is everything: Linking personality preferences to managerial behaviors, *Human Resource Development Quarterly, 11*(2), 133–157.
- Bracken, D. W. (1996). Multisource (360-degree) feedback: Surveys for individual and organizational development. In A. I. Kraut (Ed.), *Organizational surveys: Tools for assessment and change* (pp. 117–143). San Francisco, CA: Jossey-Bass.
- Bracken, D. W., & Church, A. H. (2013). The “new” performance management paradigm: Capitalizing on the unrealized potential of 360 degree feedback. *People & Strategy, 36*(2), 34–40.
- Bracken, D. W., Dalton, M. A., Jako, R. A., McCauley, C. D., & Pollman, V. A. (1997). *Should 360-degree feedback be used only for developmental purposes?* Greensboro, NC: Center for Creative Leadership.
- Bracken, D. W., & Paul, K. B. (1993, May). *The effects of scale type and demographics on upward feedback*. Presented at the 8th Annual Conference of The Society for Industrial and Organizational Psychology, San Francisco, CA.
- Bracken, D. W., & Rose, D. S. (2011). When does 360-degree feedback create behavior change? And how would we know it when it does? *Journal of Business and Psychology, 26*(2), 183–192. doi:10.1007/s10869-011-9218-5
- Bracken, D. W., Summers, L., & Fleenor, J. W. (1998). High tech 360. *Training and Development, 52*(8), 42–45.
- Bracken, D. W., & Timmreck, C. W. (2001). Guidelines for multisource feedback when used for decision making. In D. W. Bracken, C. W. Timmreck, & A. H. Church (Eds.), *The handbook of multisource feedback* (pp. 495–510). San Francisco, CA: Jossey-Bass.
- Bracken, D. W., Timmreck, C. W., & Church, A. H. (Eds.). (2001). *The handbook of multisource feedback*. San Francisco, CA: Jossey-Bass.
- Bracken, D. W., Timmreck, C. W., Fleenor, J. W., & Summers, L. (2001). 360 feedback from another angle. *Human Resource Management Journal, 40*(1), 3–20. doi:10.1002/hrm.4012
- Buckingham, M., & Goodall, A. (2015). Reinventing performance management. *Harvard Business Review, 93*(4), 40–50.
- Burke, W. W. (1982). *Organization development: Principles and practices*. Glenview, IL: Scott, Foresman.
- Burke, W. W., & Jackson, P. (1991). Making the SmithKline Beecham merger work. *Human Resource Management, 30*, 69–87.
- Burke, W. W., & Litwin, G. H. (1992). A causal model of organizational performance and change. *Journal of Management, 18*, 523–545.
- Burke, W. W., Richley, E. A., & DeAngelis, L. (1985). Changing leadership and planning processes at the Lewis Research Center, National Aeronautics and Space Administration. *Human Resource Management, 24*, 81–90.
- Bynum, B. H., Hoffman, B. J., Meade, A. W., & Gentry, W. A. (2013). Reconsidering the equivalence of multisource performance ratings: Evidence for the importance and meaning of rating factors. *Journal of Business and Psychology, 28*(2), 203–219.
- Campion, M. C., Campion, E. D., & Campion, M. A. (2015). Improvements in performance management through the use of 360 feedback. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 8*(1), 85–93.
- Cappelli, P. (2008). *Talent on demand*. Boston, MA: Harvard Business Press.
- Church, A. H. (1995). First-rate multirater feedback. *Training & Development, 49*(8), 42–43.
- Church, A. H. (1997). Do you see what I see? An exploration of congruence in ratings from multiple perspectives. *Journal of Applied Social Psychology, 27*(11), 983–1020.

- Church, A. H. (2000). Do better managers actually receive better ratings? A validation of multi-rater assessment methodology. *Consulting Psychology Journal: Practice & Research*, 52(2), 99–116.
- Church, A. H. (2014). What do we know about developing leadership potential? The role of OD in strategic talent management. *OD Practitioner*, 46(3), 52–61.
- Church, A. H. (2016). Is engagement overrated? Five questions to consider before doing your next engagement survey. *Talent Quarterly*, 9, 1–7.
- Church, A. H., & Bracken, D. W. (1997). Advancing the state of the art of 360-degree feedback. *Group & Organization Management*, 22(2), 149–161. doi:10.1177/1059601197222002
- Church, A. H., Golay, L. M., Rotolo, C. T., Tuller, M. D., Shull, A. C., & Desrosiers, E. I. (2012). Without effort there can be no change: Reexamining the impact of survey feedback and action planning on employee attitudes. In A. B. Shani, W. A. Pasmore, & R. W. Woodman (Eds.), *Research in organizational change and development* (Vol. 20, pp. 223–264). Bingley, UK: Emerald Group.
- Church, A. H., Javitch, M., & Burke, W. W. (1995). Enhancing professional service quality: Feedback is the way to go. *Managing Service Quality*, 5(3), 29–33.
- Church, A. H., & Oliver, D. H. (2006). The importance of taking action, not just sharing survey feedback. In A. Kraut (Ed.), *Getting action from organizational surveys: New concepts, technologies, and applications* (pp. 102–130). San Francisco, CA: Jossey-Bass.
- Church, A. H., Rogelberg, S. G., & Waclawski, J. (2000). Since when is no news good news? The relationship between performance and response rates in multirater feedback. *Personnel Psychology*, 53(2), 435–451.
- Church, A. H., & Rotolo, C. T. (2013). How are top companies assessing their high potentials and senior executives? A talent management benchmark study. *Consulting Psychology Journal: Practice and Research*, 65(3), 199–223.
- Church, A. H., Rotolo, C. T., Ginther, N. M., & Levine, R. (2015). How are top companies designing and managing their high-potential programs? A follow-up talent management benchmark study. *Consulting Psychology Journal: Practice and Research*, 67(1), 17–47.
- Church, A. H., Rotolo, C. T., Shull, A. C., & Tuller, M. D. (2014a). Inclusive organization development: An integration of two disciplines. In B. M. Ferdman & B. Deane (Eds.), *Diversity at work: The practice of inclusion* (pp. 260–295). San Francisco, CA: Jossey-Bass.
- Church, A. H., Rotolo, C. T., Shull, A. C., & Tuller, M. D. (2014b). Understanding the role of organizational culture and workgroup climate in core people development processes at PepsiCo. In B. Schneider & K. M. Barbera (Eds.), *The Oxford handbook of organizational climate and culture* (pp. 584–602). New York, NY: Oxford University Press.
- Church, A. H., & Waclawski, J. (1998a). Making multirater feedback systems work. *Quality Progress*, 31(4), 81–89.
- Church, A. H., & Waclawski, J. (1998b). The relationship between individual personality orientation and executive leadership behavior. *Journal of Occupational & Organizational Psychology*, 71, 99–125.
- Church, A. H., & Waclawski, J. (2001a). *Designing and using organizational surveys: A seven-step process*. San Francisco, CA: Jossey-Bass.
- Church, A. H., & Waclawski, J. (2001b). A five-phase framework for designing a successful multisource feedback system. *Consulting Psychology Journal: Practice & Research*, 53(2), 82–95.
- Church, A. H., & Waclawski, J. (2001c). Hold the line: An examination of line vs. staff differences. *Human Resource Management*, 40(1), 21–34.
- Church, A. H., Waclawski, J., & Burke, W. W. (2001). Multisource feedback for organization development and change. In D. W. Bracken, C. W. Timmreck, & A. H. Church (Eds.), *The handbook of multisource feedback* (pp. 301–317). San Francisco, CA: Jossey-Bass.
- Church, A. H., Walker, A. G., & Brockner, J. (2002). Multisource feedback for organization development and change. In A. H. Church & J. Waclawski (Eds.), *Organization development: A data-driven approach to organizational change* (pp. 27–51). San Francisco, CA: Jossey-Bass.
- Companies use “360-degree feedback” to evaluate employees. (1993, October). *Communication World*, 10, p. 9.

- Costar, D. M., & Bracken, D. W. (2016, April). Improving the feedback experience within performance management using 360 feedback. In D. W. Bracken (Chair), *Feedback effectiveness within and without performance management*. Symposium conducted at the 31st Annual Conference of the Society for Industrial and Organizational Psychology, Anaheim, CA.
- Dalton, M. (1997). When the purpose of using multi-rater feedback is behavior change. In D. W. Bracken, M. A. Dalton, R. A. Jako, C. D. McCauley, & V. A. Pollman (Eds.), *Should 360-degree feedback be used only for developmental purposes?* (pp. 1–6). Greensboro, NC: Center for Creative Leadership.
- DeNisi, A., & Kluger, A. (2000). Feedback effectiveness: Can 360-degree appraisal be improved? *Academy of Management Executive*, 14(1), 129–139.
- Donnon, T., Al Ansari, A., Al Alawi, S., & Violato, C. (2014). The reliability, validity, and feasibility of multisource feedback physician assessment: A systematic review. *Academic Medicine*, 89, 511–516.
- Edwards, M. R., & Ewen, A. J. (1995). Assessment beyond development: The linkage between 360-degree feedback, performance appraisal and pay. *ACA Journal*, 23, 2–13.
- Edwards, M. R., & Ewen, A. J. (1996). *360° feedback: The powerful new model for employee assessment & performance improvement*. New York, NY: AMACOM.
- Effron, M. (2013, November). The hard truth. *Leadership Excellence*, 30(11), 4–5.
- Effron, M., & Ort, M. (2010). *One page talent management: Eliminating complexity, adding value*. Boston, MA: Harvard Business School Publishing.
- Eichinger, R. W., & Lombardo, M. M. (2004). Patterns of rater accuracy in 360-degree feedback. *Human Resource Management*, 27(4), 23–25.
- English, A., Rose, D. S., & McLellan, J. (2009, April). *Rating scale label effects on leniency bias in 360-degree feedback*. Paper presented at the 24th Annual Conference of the Society for Industrial Organizational Psychology, New Orleans, LA.
- Facteau, J. D., & Craig, S. B. (2001). Are performance ratings from different sources comparable? *Journal of Applied Psychology*, 86(2), 215–227.
- Farr, J. L., & Newman, D. A. (2001). Rater selection: Sources of feedback. In D. W. Bracken, C. W. Timmreck, & A. H. Church (Eds.), *The handbook of multisource feedback* (pp. 96–113). San Francisco, CA: Jossey-Bass.
- Ferguson, J., Wakeling, J., & Bowie, P. (2014). Factors influencing the effectiveness of multisource feedback in improving the professional practice of medical doctors: A systematic review. *MBC Medical Education*, 14(76), 1–12.
- Fleener, J., Taylor, S., & Chappelow, C. (2008). *Leveraging the impact of 360-degree feedback*. New York, NY: Wiley.
- Furnham, A., & Stringfield, P. (1994). Congruence of self and subordinate ratings of managerial practices as a correlate of supervisor evaluation. *Journal of Occupational and Organizational Psychology*, 67, 57–67.
- Ghorpade, J. (2000). Managing five paradoxes of 360-degree feedback. *Academy of Management Executive*, 14(1), 140–150.
- Goldsmith, M., & Morgan, H. (2004). Leadership is a contact sport: The “follow-up” factor in management development. *Strategy + Business*, 36, 71–79.
- Greguras, G. J., & Robie, C. (1998). A new look at within-source interrater reliability of 360-degree feedback ratings. *Journal of Applied Psychology*, 83, 960–968. doi:10.1037/0021-9010.83.6.960
- Hardison, C. M., Zaydman, M., Oluwatola, T., Saavedra, A. R., Bush, T., Peterson, H., & Straus, S. G. (2015). *360-degree assessments: Are they the right tool for the U.S. military?* Santa Monica, CA: Rand.
- Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41, 43–62. doi:10.1111/j.1744-6570.1988.tb00631.x
- Hazucha, J., Hezlett, S. A., & Schneider, R. J. (1993). The impact of 360-degree feedback on management skills development. *Human Resources Management*, 32(2–3), 325–351. doi:10.1002/hrm.3930320210

- Healy, M. C., & Rose, D. S. (2003, April). *Validation of a 360-degree feedback instrument against sales: Content matters* (Technical Report No. 8202). Presented at the 18th Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- Hedge, J. W., Borman, W. C., & Birkeland, S. A. (2001). History and development of multisource feedback as a methodology. In D. W. Bracken, C. W. Timmreck, & A. H. Church (Eds.), *The handbook of multisource feedback* (pp. 15–32). San Francisco, CA: Jossey-Bass.
- Heidermeier, H., & Moser, K. (2009). Self–other agreement in job performance ratings: A meta-analytic test of a process model. *Journal of Applied Psychology, 94*, 353–370.
- Hofstede, G., & McRae, R. R. (2004). Personality and culture revisited: Linking traits and dimensions of culture. *Cross-Cultural Research, 38*, 52–88.
- Johnson, J. W., & Ferstl, K. L. (1999). The effects of interrater and self–other agreement on performance improvement following upward feedback. *Personnel Psychology, 52*(2), 271–303.
- Johnson, R. H. (1996). Life in the consortium: The Mayflower Group. In A. I. Kraut (Ed.), *Organizational surveys: Tools for assessment and change* (pp. 285–309). San Francisco, CA: Jossey-Bass.
- Jones, J. E., & Bearley, W. L. (1996). *360 feedback: Strategies, tactics, and techniques for developing leaders*. Valley Center, CA: Organizational Universe Systems.
- Kaiser, R. B., & Craig, S. B. (2005). Building a better mousetrap: Item characteristics associated with rating discrepancies in 360-degree feedback. *Consulting Psychology Journal: Practice and Research, 57*(4), 235–245.
- Kaplan, R. E. (1993). 360-degree feedback PLUS: Boosting the power of co-worker ratings for executives. *Human Resource Management, 32*(2–3), 299–314.
- Kirk, J., & Miller, M. L. (1986). *Qualitative Research Methods Series: Vol. 1. Reliability and validity in qualitative research*. Newbury Park, Sage.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback theory. *Psychological Bulletin, 119*, 254–284.
- Kraiger, K., & Ford, J. K. (1985). A meta-analysis of rater race effects on performance ratings. *Journal of Applied Psychology, 70*, 56–65.
- Lawler, E. E. (1967). The multitrait-multirater approach to measuring managerial job performance. *Journal of Applied Psychology, 51*(5), 369–381.
- Lepsinger, R., & Lucia, A. D. (1997). *The art and science of 360° feedback*. San Francisco, CA: Jossey-Bass.
- Lepsinger, R., & Lucia, A. D. (2009). *The art and science of 360° feedback* (2nd ed.). San Francisco, CA: Jossey-Bass.
- Leslie, J. B. (2013). *Feedback to managers: A guide to reviewing and selecting multirater instruments for leadership development* (4th ed.). Greensboro, NC: Center for Creative Leadership.
- Levy, P. E., Silverman, S. B., & Cavanaugh, C. M. (2015). The performance management fix is in: How practice can build on the research. *Industrial and Organizational Psychology, 8*(1), 80–85.
- Likert, R. (1967). *The human organization*. New York, NY: McGraw Hill.
- Lockyer, J. (2013). Multisource feedback: Can it meet criteria for good assessment? *Journal of Continuing Education in the Health Professions, 33*(2), 89–98.
- London, M. (2001). The great debate: Should multisource feedback be used for administration or development only? In D. W. Bracken, C. W. Timmreck, & A. H. Church (Eds.), *The handbook of multisource feedback* (pp. 368–388). San Francisco, CA: Jossey-Bass.
- London, M., & Beatty, R. W. (1993). 360-degree feedback as a competitive advantage. *Human Resource Management, 32*(2–3), 353–372. doi:10.1002/hrm.3930320211
- London, M., Smither, J. W., & Adsit, D. L. (1997). Accountability: The Achilles' heel of multisource feedback. *Group and Organization Management, 22*, 162–184.
- MacRae, I., & Furnham, A. (2014). *High potential: How to spot, manage and develop talent people at work*. London, UK: Bloomsbury.
- Massie, J., & Ali, J. M. (2016). Workplace-based assessment: A review of user perceptions and strategies to address the identified shortcomings. *Advances in Health Sciences Education, 21*(2), 455–473. doi:10.1007/s10459-015-9614-0

- McCauley, C. D., & McCall, M. M. (2014). *Using experience to develop leadership talent: How organizations leverage on-the-job development*. San Francisco, CA: Jossey-Bass.
- Mount, M. K., Judge, T. A., Scullen, S. E., Systma, M. R., & Hezlett, S. A. (1998). Trait, rater and level effects in 360-degree performance ratings. *Personnel Psychology*, 51(3), 557–576.
- Murphy, K. R., & Cleveland, J. N. (1995). *Understanding performance appraisal*. Thousand Oaks, CA: Sage.
- Nadler, D. A. (1977). *Feedback and organization development: Using data-based methods*. Reading, MA: Addison-Wesley.
- Nowack, K. M. (1993, January). 360-degree feedback: The whole story. *Training & Development*, 47, 69–72.
- Nowack, K. (2009). Leveraging multirater feedback to facilitate successful behavioral change. *Consulting Psychology Journal: Practice and Research*, 61, 280–297.
- Nowack, K. M., & Mashihi, S. (2012). Evidence-based answers to 15 questions about leveraging 360-degree feedback. *Consulting Psychology Journal: Practice and Research*, 64(3), 157–182.
- Oliver, D. H., Church, A. H., Lewis, R., & Desrosiers, E. I. (2009). An integrated framework for assessing, coaching and developing global leaders. In J. Osland (Ed.), *Advances in global leadership* (pp. 195–224). Bingley, UK: Emerald Group.
- O'Reilly, B. (1994, October 17). 360 feedback can change your life. *Fortune*, 130(8), 93–96.
- Pfau, B., & Kay, I. (2002). Does 360-degree feedback negatively affect company performance? *HRMagazine*, 47(6), 5–60.
- Pulakos, E. D., Hanson, R. M., Arad, S., & Moya, N. (2015). Performance management can be fixed: An on-the-job experiential learning approach for complex behavioral change. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 8(1), 51–76.
- Rogelberg, S. G., & Waclawski, J. (2001). Instrumentation design. In D. W. Bracken, C. W. Timmreck, & A. H. Church (Eds.), *The handbook of multisource feedback* (pp. 79–95). San Francisco, CA: Jossey-Bass.
- Rose, D. S. (2011, April). *Using strategically aligned 360-degree feedback content to drive organizational change*. Paper presented at the 26th Annual Conference of the Society for Industrial Organizational Psychology, Chicago IL.
- Rose, D. S., English, A., & Thomas, C. (2011, April). Taming the cyber tooth tiger: Our technology is good, but our science is better. *The Industrial-Organizational Psychologist*, 48(3), 21–27.
- Scott, J. C., & London, M. (2003). The evaluation of 360-degree feedback programs. In J. E. Edwards, J. C. Scott, & N. S. Raju (Eds.), *The human resources program-evaluation handbook* (pp. 177–197). Thousand Oaks, CA: Sage
- Scott, J. C., & Reynolds, D. H. (Eds.). (2010). *The handbook of workplace assessment: Evidenced based practices for selecting and developing organizational talent*. San Francisco, CA: Jossey-Bass.
- Scullen, S. E., Mount, M. K., & Goff, M. (2000). Understanding the latent structure of job performance ratings. *Journal of Applied Psychology*, 85, 956–970.
- Silzer, R., & Church, A. H. (2010). Identifying and assessing high potential talent: Current organizational practices. In R. Silzer & B. E. Dowell (Eds.), *Professional Practice Series: Strategy-driven talent management: A leadership imperative* (pp. 213–279). San Francisco, CA: Jossey-Bass.
- Silzer, R. F., & Dowell, B. E. (Eds.). (2010). *Strategy-driven talent management: A leadership imperative*. San Francisco, CA: Jossey-Bass.
- Smither, J., London, M., Flautt, R., Vargas, Y., & Kucine, I. (2003). Can working with an executive coach improve multisource feedback ratings over time? A quasi-experimental field study. *Personnel Psychology*, 56, 23–44.
- Smither, J., London, M., & Reilly, R. (2005). Does performance improve following multisource feedback? A theoretical model, meta-analysis, and review of empirical findings. *Personnel Psychology*, 58, 33–66. doi:10.1111/j.1744-6570.2005.514_1.x

- Steelman, L. A., Levy, P. A., & Snell, A. F. (2004). The feedback environment scale: Construct definition, measurement and, and validation. *Educational and Psychological Measurement*, 64(1), 165–184.
- Summers, L. (2001). Web technologies for administering multisource feedback programs. In D. W. Bracken, C. W. Timmreck, & A. H. Church (Eds.), *The handbook of multisource feedback* (pp. 165–180). San Francisco, CA: Jossey-Bass.
- Swoboda, F. (1993, March 30). “360-degree feedback” offers new angles in job evaluation. *Star Tribune*, p. 2D.
- 3D Group. (2003). *Benchmark study of North American 360-degree feedback practices* (Technical Report No. 8214). Berkeley, CA: Data Driven Decisions.
- 3D Group. (2004). *Benchmark study of North American 360-degree feedback practices* (3D Group Technical Report No. 8251). Berkeley, CA: Data Driven Decisions.
- 3D Group. (2009). *Current practices in 360-degree feedback: A benchmark study of North American companies* (Report No. 8326). Berkeley, CA: Data Driven Decisions.
- 3D Group. (2013). *Current practices in 360-degree feedback: A benchmark study of North American companies*. Emeryville, CA: Data Driven Decisions.
- Timmreck, C. W., & Bracken, D. W. (1997). Multisource feedback: A study of its use in decision making. *Employment Relations Today*, 24(1), 21–27. doi:10.1002/ert.3910240104
- Toegel, G., & Conger, J. (2003). *360-degree assessment: Time for reinvention* (Report No. G03-17[445]). Los Angeles, CA: Center for Effective Organizations. Retrieved from <http://ceo.usc.edu/360-degree-assessment-time-for-reinvention/>
- Tornow, W. W. (Ed.). (1993). Special issue on 360-degree feedback. *Human Resource Management*, 32(2–3). Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/hrm.v32:2/3/issuetoc>
- Tornow, W. W., London, M., & CCL Associates. (1998). *Maximizing the value of 360-degree feedback: A process for successful individual and organizational development*. San Francisco, CA: Jossey-Bass.
- United States Office of Personnel Management. (2012). *Executive development best practices guide*. Washington, DC: Author.
- Van Velsor, E. (1998). Designing 360-feedback to enhance involvement, self-determination, and commitment. In W. W. Tornow & M. London (Eds.), *Maximizing the value of 360-degree feedback* (pp. 149–195). San Francisco, CA: Jossey-Bass.
- Van Velsor, E., Leslie, J. B., & Fleenor, J. W. (1997). *Choosing 360: A guide to evaluating multi-rater feedback instruments for management development*. Greensboro, NC: Center for Creative Leadership.
- Viswesvaran, C., Schmidt, F., & Ones, D. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81, 557–574.
- Walker, A. G., & Smither, J. W. (1999). A five-year study of upward feedback: What managers do with their results matter. *Personnel Psychology*, 52(2), 393–423.
- Walker, A. G., Smither, J. W., Atwater, L. E., Dominick, P. G., Brett, J. F., & Reilly, R. R. (2010). Personality and multisource feedback improvement: A longitudinal investigation. *Journal of Behavioral and Applied Management*, 11(2), 175–204.
- Ward, P. (1997). *360-degree feedback*. London, UK: Institute of Personnel and Development.
- Woehr, D. J., & Huffcutt, A. I. (1994). Rater training for performance appraisal: A quantitative review. *Journal of Occupational and Organizational Psychology*, 67, 189–206.
- Wood, L., Hassell, A., Whitehouse, A., Bullock, A., & Wall, D. (2006). A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. *Medical Teacher*, 28(7), 185–191.