C E **Center for**
Effective
Organizations

360-DEGREE ASSESSMENT:
TIME FOR REINVENTION

GINKA TOEGEL

London School of Economics

JAY A. CONGER

Center for Effective Organizations
Marshall School of Business
University of Southern California

May 2003

Abstract

One of the most popular management development tools in use today is the 360-degree assessment instrument. In recent years, however, its popularity has led to uses beyond its original application for management development. In particular, 360-degree assessment is now replacing the traditional performance appraisal. This trend towards multiple uses especially administrative ones should raise concerns. We discuss the implications of this trend. In particular, our focus will be on dilemmas created when a feedback tool is stretched to include potentially conflicting aims. The analysis is carried out on three different levels (individual, interpersonal and organizational) using three different frames (cognitive, psychometric and game-theoretical). It leads to the conclusion that 360-degree assessment is in danger of losing its efficacy as a process to deliver honest and constructive feedback if used for multiple purposes. We suggest that it is time for reinvention of the tool and its process methodology. In particular, we argue for the development of two distinct tools – one for management development and one for performance feedback. The management development tool would rely more heavily upon qualitative feedback and competences for development. The performance appraisal feedback tool would be designed around quantitative feedback and measuring performance outcomes and performance-related behaviors.

360-Degree Assessment: Time for Reinvention

A decade ago, few managers had ever heard of 360-degree feedback. After all, the boss was the traditional source of feedback. This individual made one's promotion and pay decisions, and it seemed only appropriate that he or she be the primary source for performance and development feedback. The notion that subordinates and peers should have a say would have seemed odd. Respect for the formal hierarchy mitigated against feedback from subordinates below. One's formal authority might be challenged if direct reports had an active voice in assessing the boss's capabilities. Work arrangements demanding extensive collaboration with peers were fewer in number, and so it appeared that one's co-workers were often not in a position to provide detailed feedback. It was not until the late 1980s and early 1990s that a handful of organizations began to experiment with a development tool called 360-degree feedback. This typically competence-based survey instrument solicited confidential evaluations from the full range of working relationships a manager possessed – subordinates, peers, and bosses - using a quantitatively-based multi-item questionnaire. Each targeted individual in turn received a report summarizing in numerical and descriptive assessments of their capability to effectively demonstrate specific competences based on the perceptions of those assessing them.

Over the last decade, several forces have conspired to radically expand interest in 360-degree feedback and turn it into one of the most popular tools in the history of management development. One particularly important force has been the dramatic rise in formal leadership development programs for managers and executives. Integral to many of these programs is 360-degree feedback around a set of leadership competences. Given the critical role that subordinates play in determining whether one possesses leadership capabilities or not, this group of individuals has now become an essential source for feedback .

A second force driving the popularity of 360-degree feedback is the new work arrangements. Specifically, as hierarchies have flattened and more work is performed across functions and in cross-functional teams, peer input has gained in importance. Today it can be argued that peers are better positioned to give insightful feedback than a decade ago. In addition, effective leadership is defined in part around one's capacity to build and sustain networks of relationships throughout an organization. Who better to assess this capacity than the very members of one's network – peers.

Finally, organizations, nowadays, try to measure almost everything and everyone. Performance measurement is particularly popular because it informs both employees and organizations of their effectiveness in getting results and achieving goals. The quantitative ranking associated with 360-degree assessment therefore has had great appeal as a potential performance measurement tool.

Despite the dramatic rise in its popularity, the 360-degree assessment is far from perfect. A meta-analysis suggests that over one third of the feedback interventions even decreased performance (Kluger & DeNisi, 1996). Similarly, half of the ratees studied by Atwater, Waldman, Atwater, and Cartier (2000) failed to respond to feedback with improvement. Consequently, it is sensible to question whether the current form and use of 360-degree assessment is the most effective. This is a critical time to step back and assess the efficacy of feedback - especially in light of a trend in organizations to substitute 360-degree assessment processes for traditional performance appraisals. Over the last several years, we have been witnessing the migration of 360-degree assessment from a developmental tool to a performance assessment tool (London & Smither, 1995; Bettenhausen & Fedor, 1997; Waldman, Atwater & Antonioni, 1998; Fletcher & Baldry, 2000;

Bracken, Timmreck, Fleenor, & Summers (2001). Feedback is no longer for the sole purpose of *informing* one's developmental needs. Rather it is for *assessing* one's performance in the current job and potential for promotion. This trend began in the 1990s when criticism began to mount about traditional performance appraisals. According to a Pulse Survey conducted at the Conference Board's 1999 Human Resource Conference in New York, ninety percent of human resource executives said that, if given the opportunity, they would modify, revise or even eliminate the performance appraisal system currently used in their companies (HRFocus, 2000).

By the 1990s, certain workplace trends had seriously eroded the meaningfulness of formal appraisals and caused a growing dissatisfaction with the review process. For example, flatter organizational structures had loosened the link between reviews and promotions (Kennedy, 1999). Outside hires in place of internal promotions had devalued the formal appraisal process, which was seen as the 'guarantee' for opportunities higher up in the organization. Some experts have even argued that traditional appraisals are so dysfunctional that they need to be abolished (Coens & Jenkins, 2000). In their view, using one process for feedback, compensation, staffing, and succession is impractical (Scholtes, 1999). There appears to be such a high degree of dissatisfaction with performance appraisal (PA) systems that "the very term PA has been virtually censored" from the organizational vocabulary and has been replaced with the term 'performance management' (Bernardin, Hagan, Kane & Villanova, 1998:4). Obviously, there are many unresolved issues in the theory and practice of performance appraisals. This dissatisfaction, however, did produce a salvage operation by organizations to 'co-opt' 360-degree assessment in order to make performance appraisal decisions more objective and acceptable.

There are clear advantages to addressing multiple objectives such as developmental feedback and performance appraisals using a single 360-degree assessment tool. For example, the 360-degree process becomes much more efficient. It offers a higher return on the investment in its development and use with multiple applications (Edwards & Ewen, 1996; Bracken, 1997; Jako, 1997; Fletcher 1998,1999). Given the time intensive demands of collecting 360-degree assessment data, there is a natural inducement for managers to employ these data for administrative decisions such as rewards, performance appraisals, and succession as well as for developmental guidance. When 360-degree assessment is used for appraisal decisions, it can be an empowering mechanism that gives direct reports and peers a real say in how effective their boss or peer is as a leader (Bracken, Timmreck, Fleenor & Summers, 2001). Finally, unlike the traditional boss-subordinate appraisal process, 360-degree data provide a more comprehensive picture of the manager's performance in contrast to the singular lens of the boss's viewpoint (Fletcher, 1999).

At the same time, there is growing concern about the migration of 360-degree assessments towards dual purposes – development and performance appraisal (Pollack & Pollack, 1996; Jones & Bearley, 1996; Dalton, 1997; Pollman, 1997; DeNisi & Kluger, 2000). For example, opponents of the migration towards appraisal argue that the goal of 360-degree assessment should be broader than simply assessing performance. It should foster continuous learning and personal development. To achieve this end, a psychologically safe environment where individuals can comfortably accept their ratings is a necessary condition – in other words, an environment where they are not motivated to "argue with, deny, reject the data, and fail to change" because of a perceived threat or personal vulnerability (Dalton, 1997:3). This perspective suggests that using 360-degree data for performance appraisal makes the developmental process potentially 'punitive' and one that is 'forcing' instead of 'enabling' change (Pollman, 1997:7). As a by-product, respondents may either inflate ratings so as not to hurt colleagues or else provide more negative ratings in retribution for poor leadership. Those who argue against the shift towards administrative use also point out that certain assumptions have to be made for multi-source assessments to

be used for performance appraisals. These assumptions may be largely invalid: e.g. organizations are rational entities, administrative systems can be highly reliable, and people can be trained and motivated to be unbiased and candid in assessments (McCauley, 1997).

As we can see, the debate about dual purposes is ongoing and it is far from being resolved. In the meantime, companies continue to experiment. For example, a number of global firms have started using 360-degree assessment not only for development, but for administrative decisions as well. Even academics such as Waldman and Bowen (1998:117) predict that 360-degree will be "included as a routine part of future appraisal systems".

We believe that 360-degree assessment is indeed at a critical juncture in its history. The underlying premise of the 360-degree methodology - obtaining information from various sources - is sound whether for development or appraisal. The dilemma is that employing one tool and one data gathering process for such diverse purposes weakens the tool and its ability to deliver on its objectives. We believe that two different tools are necessary to achieve these differing goals. One tool, more qualitative in nature, should aim at providing richer data and in highlighting the differences in views of the participant's strengths and weaknesses. This tool should serve the sole purpose of management development. The second tool, with stronger psychometric properties (e.g. high inter-rater reliability and accurate measurement of performance outcomes based on standards), should be used for administrative decisions such as performance appraisals that drive promotions and salary increases. Both should differ in their content and format.

The discussion that follows will illustrate the potential conflicts, contradictions and dilemmas that the trend towards multiple purposes with 360-degree assessment is creating. Our arguments favoring two separate instruments will center on three different levels of analysis - individual, interpersonal and organizational - using three different frames - cognitive, psychometric and game-theoretical (see Table 1). At the core of multi-source feedback is a *cognitive* process of self-reflection, which increases our self-awareness (Tornow, 1993; Yammarino & Atwater, 1993; Church, 1994) – in other words, at the individual level of analysis. Multi-source feedback is also an *interactive* process that takes place at an interpersonal level. Others are providing their perceptions based upon their history of interpersonal relations and task accomplishments with the target individual receiving feedback. Finally, there is an organizational dimension that determines the way ratings are processed as *metrics* for administrative purposes. Later, we will describe the two tools and how they can best be deployed.

Contradictions, Dilemmas and Conflicts in Today's Use of 360-Degree Assessments

To illustrate the problems that arise from a dual purposes approach to 360-degree assessments (both development and appraisal combined), we will discuss some of the contradictions and dilemmas that arise from the way people process information and make decisions concerning their behavior. We will look at arguments from the cognitive, psychometric and game-theoretical domains that reflect the three different levels of analysis: individual, interpersonal and organizational. Consequently, we will discuss three issues: 1) how the development versus appraisal applications frame ownership and purpose and in turn shape the implications for learning and performance improvement, 2) the probability of rater-ratee collusion under the development versus appraisal applications and process implications, and 3) tradeoffs made around measurement depending upon the chosen application of either developmental or performance feedback.

The Individual Level: Cognitive Arguments

At the heart of 360-degree assessment is the giving and processing of evaluative ratings. These are cognitive activities that embrace attention, encoding, storage, retrieval, integration and evaluation (DeNisi & Williams, 1988) and are embedded in situational contexts (Judge & Ferris, 1993) with multiple internal and external influences. How these evaluative ratings are used in turn influences the motives and behavior of both the rater and those being rated. For example, what are the implications when the individual targeted for the feedback alone *owns* the feedback (development) versus when it is shared with others in their organization (appraisal)? Similarly, what are the implications when the *framing* of the purpose of data collection is for development versus for performance assessment? We believe that different frames lead to different outcomes.

Framing of the ownership of feedback. People remember and process more thoroughly information that pertains to themselves (Rogers, Kuiper & Kirker, 1977). However, if this information is negative and it has to be shared with relevant constituencies, e.g. in the context of performance appraisal, a defensive reaction is activated, which can lead to performance decline (Meyer, Kaye, & French, 1965; DeNisi & Kluger, 2000). Feedback will have a special status if the person receiving feedback owns it entirely, i.e. the results are confidential. Research suggests that information gains in value merely by being owned (Kahneman, Knetsch and Thaler, 1990), and this effect might extend to seemingly trivial things (Baumeister, 1998). Researchers at the Center for Creative Leadership, for example, argue that, in order to change and develop, individuals need to feel psychologically safe and to 'own' their evaluations (Dalton, 1997). Therefore, if the goal is development, ownership of 360-degree data by the recipient is crucial. When 360-degree data is used for administrative purposes like performance assessment, the data become the property of the organization, not the individual (Lepsinger & Lucia, 1997). This dilemma is one example of the problems caused by the coupling of development and appraisal aims.

Framing the purpose of the feedback. Decision making begins with a framing activity (Harrison & Phillips, 1991; Nutt, 1993). In other words, individuals make sense of their environment by attaching labels to the situations they are involved in (Daft & Weick, 1984). These labels are abbreviated meanings, which in turn affect a decision maker's subsequent actions (Mintzberg, Raisanghani & Theoret, 1976; Fredrickson, 1985; Dutton & Jackson, 1987). In other words, there is a link between the framing or labeling of an issue and a subsequent action. Consequently, we can expect that the labeling of the purpose of 360 data collection as either *personal development* or as *performance appraisal* will activate different motivations and therefore differences in behavior of those receiving the feedback. There is empirical evidence, for example, that raters change their ratings when 360-degree becomes a performance evaluation tool (Farh, Cannella & Bedeian, 1991; London & Wohlers, 1991, Murphy & Clevelend, 1995; Waldman, Atwater & Antonioni, 1998) in place of a development tool. More than 70% of managers in one study admitted to having inflated or deflated evaluations (Longenecker & Ludwig, 1990) in order to send a signal to protect a colleague or to shock a poor performer. Another reason for inflation is that high ratings of subordinates by managers make the supervisor look effective (Bowman, 1999).

Framing of the purpose therefore has an impact on respondents' attitudes towards 360-degree assessment. In the developmental context, raters appear to have an honesty mind-set, while in the appraisal framing it is an accuracy mind-set where 'accuracy of ratings and honesty of ratings are not necessarily the same thing' (McCauley, 1997:34). An example illustrates the latter situation. Raters may have observed on two occasions their boss making hasty decisions

without a thoughtful review of all the relevant information. In a developmental context, where the goal is constructive feedback and where the honesty mind set dominates, raters would be more inclined to share this observation with the focal individual. However, in the appraisal framing where the accuracy mind-set prevails and the individual tries to be as objective as possible, raters might decide not to share this information since it is based on a limited number of observations. In other words, they may have doubts about its accuracy.

The framing of the purpose of the 360-degree assessment influences not only raters, but those being rated as well. When the purpose is development, individuals receiving the feedback are more motivated to look for accurate feedback with an aim to make decisions about enhancing the effectiveness of their behavior (see our discussion on self-enhancement motives later in this section). When the purpose is performance appraisal, they are more inclined to seek favorable feedback and to increase their ratings by impression management. While it is true that we always engage in impression management, employees may act in certain situations with far greater deliberation to manipulate the impression that others have of them (London & Smither, 1995). Self-presentation refers to our tactics to convey information about ourselves to others. One of the motives that drives our inclination to manipulate others' opinion of ourselves is the *strategic self-presentation*. It is instrumental because "the task of impressing others is a strategy for achieving ulterior goals" (Baumeister, 1998:704), e.g. promotions, pay increase, etc. There are different forms of self-presentation (Jones & Pitman, 1982): *ingratiation* (emphasizing appealing traits in order to be liked), *self-promotion* (getting respect after convincing others in one's competence), or *exemplification* (showing one's moral virtues). When feedback is used for appraisal, the ratee can either aim at real behavioral change or at manipulating the impression of him-/herself through self-presentation. Research shows that "if the goal is to secure rewards from the other person, then one tries to present oneself as closely as possible to the other person's values and preferences" (Baumeister, 1998:705). Self-presentation is an easier way to achieve favorable feedback. Hence, employees may focus only on what needs to be done to get higher ratings. In other words, the focus of their attention will shift from behavior change, which requires serious effort, to self-presentation through well-designed manipulation of target raters. In other words, participants will seek *favorable* feedback over *accurate* feedback. For example, they may selectively demonstrate more timely responses or provide greater favors to those peers who are likely to be chosen as their raters.

Research suggests that the construction of self-knowledge is a cognitive process that can be driven by three different motives – the appraisal, self-enhancement, and consistency motives (Baumeister, 1998). The *appraisal motive*, for example, reflects the need of an individual for accurate feedback from others. People are curious how others perceive their abilities, traits, etc. The *self-enhancement motive* encourages people to seek favorable information about themselves. It suggests that we have strong preferences to learn that we are good. Finally, the *consistency motive* is a quest for evidence that confirms what people already know about themselves. It is reflected in a tendency to reject contrary evidence and to be resistant to change. Common sense would suggest that accurate information is more useful than favorable or confirming information. However, research shows that self-enhancement is the strongest motive when we pursue knowledge about ourselves. This is followed by consistency, and unfortunately, the appraisal motive comes last when we process information referring to ourselves (Sedikides, 1993). These findings point out that there is a strong and deeply rooted self-enhancement orientation in our behavior (Baumeister, 1998).

When the goal of 360-degree assessment is for appraisal decisions, we can expect that the self-enhancement motive will be further strengthened because it often leads to rewards in terms

of promotions and pay increases. In this case, thinking well of oneself and inflating one's view of their selves will become even more desired. As a result, 'positive illusions' about the self may be fostered (Taylor & Brown, 1988; Taylor, 1989; Baumeister, 1998). The impact that self-deception processes have on our behavior is well documented. People try to minimize the time spent on processing critical feedback (Baumeister & Cairns, 1992). They may selectively forget negative feedback (Crary, 1966). They compare themselves against less successful others under ego threats (Crocker & Major, 1989). They create a sense of uniqueness of their abilities (Marks, 1984). To make matters worse, the organization itself ends up with invalid data. The only mitigating factor may involve receiving accurate information from people whose perceptions matter for real consequences such as supervisors. In contrast, research shows that an individual's tendency to seek constructive feedback about shortcomings is positively related to ratings of overall effectiveness and therefore is a sound developmental strategy for managers (Ashford & Tsui, 1991).

As we have seen, there are benefits from the positive illusions we have about ourselves, because they promote well-being, may improve interpersonal relations and lead to higher motivation and persistence (Taylor and Brown, 1989). On the other side, there is a danger of having inflated or false beliefs about the self. Interestingly enough, there is a mechanism to manage this balance between benefits and dangers. Research shows that we turn off this self-enhancement orientation and become quite accurate in our self- assessments when we are in the mode of decision-making (Gollwitzer & Kinney, 1989). For example, if a manager has a serious problem dealing with a poor performer, s/he would have to make a decision whether to attribute the problem to his/her management style or to the employee alone. In this situation, s/he will accurately analyze feedback related to issues like 'does not delegate', 'does not help people develop as professionals', and 'does not give feedback and recognition'. In the developmental context when the focal individual is the actor who has to make decisions about possible changes, the self-enhancement motive will be weaker compared to the appraisal context when others make decisions and the ratee is the recipient. Receiving accurate information from people whose opinion has genuine consequences for the focal individual does not guarantee behavior change in and of itself. For example, negative information from supervisors even if accurate, can lead to dissatisfaction and frustration (Podsakoff & Fahr, 1989). Under such circumstances, high discrepancies between self- and supervisor's evaluations may evoke reactions of anger and discouragement and motivate ratees to argue that the negative information provided is inaccurate (Brett & Atwater, 2001). In general, empirical evidence shows that poorer than expected feedback can lead to negative emotions, which is followed by a reduction in the motivation to change, denial of the usefulness of multisource feedback and questioning about its accuracy (Atwater, Waldman, Atwater, & Cartier, 2000; Brett & Atwater, 2001; Atwater, Waldman & Brett, 2002). The appraisal context can further solidify negative and defensive reactions no matter how accurate the 360-degree assessment information is.

In sum, how the *purpose* of 360-degree assessment is framed creates different motivational responses from participants. In the developmental frame, the motivation on the part of those being rated is 'to *be* more effective', while in the administrative one it is 'to be *rated* as more effective' (Pollman, 1997:7). This contradiction calls into question the combining of purposes in 360-degree assessment settings.

The Interpersonal Level: A Game-Theoretical Argument

In relationships, much of behavior is based on the notion of reciprocity. Some authors argue that modern society is like a dry-stone wall, which does not need cement since it holds its stones in place by the mechanism of reciprocity: the 'I'll-scratch-your-back-if-you'll-scratch-mine' principle (Binmore, 1994). How an individual 'helps' a colleague through feedback will differ depending upon the end purpose – whether for development or performance appraisal. For example, if 360-degree assessment is used for performance appraisal, it may actually encourage collusive behavior. In this case, collusion would imply that employees have an explicit or implicit agreement to cooperate - giving each other a positive evaluation. This might appear to be a very instrumental view of organizations, but research supports the idea that people are guided by the principle of reciprocity and that they will and do collude. Social-psychological studies, for example, provide evidence that people will produce mutually advantageous interactions even without deliberately planning to do so and without explicitly communicating their desire to do so (Rabinowitz, Kelley & Rosenblatt, 1966).

In any situation requiring interdependence between individuals, mutually advantageous solutions are quite likely (Frank, 1988). The coupling of purposes in 360-degree assessment gives employees a context where they can 'game' the system. Since there is an interdependence among the participants in the evaluation process, they may collude because of "self-interested calculation of the benefits and losses that may accrue from 'polite' behavior" (Kreps, 1990: 505). In other words, individuals are more likely to give a positive rating of each other with the implicit threat that they will change their strategy towards more punitive responses if the other side deviates from positive feedback. Needless to say that collusion is counterproductive in performance appraisal because it does not lead to objective ratings and therefore undermines fairness.

If 360-degree assessment is used for performance appraisal purposes, collusion makes perfect sense. The 360-degree assessment process meets several requirements for effective collusion between individuals: 1) it is repeated, 2) the players face the prospect of having to evaluate each other again and again, 3) the way they have evaluated each other can often be inferred, and 4) punishments are usually not delayed because of the relatively short intervals between evaluations (six months to one year on average). Moreover, to facilitate the sense making process, the data is made as rich as possible. With this goal in mind, first the different constituencies are separated, for example, direct reports, peers, bosses, clients, etc. Second, ratees are given a respondent summary, which provides information about the response patterns of their individual raters (though it does not disclose individual identities). Finally, there is often rich information supplied by the written feedback. Our experience shows that ratees put these pieces of information together and develop hypotheses how certain people might have evaluated them. This is not surprising having in mind that ratees themselves choose their raters. In other words, in order to facilitate development, data can be aggregated only to a certain level. This process, however, might make the inference about the identity of the rater much easier and more of a concern under an appraisal situation.

As a result of the above factors, a 'tit-for-tat' strategy (beginning with cooperation, or evaluating high the other individual, and then copying whatever move (s)he made in the previous stage) could shape the behavior when 360-degree assessment is used for performance appraisal. As a result, accurate feedback on individual behavior is difficult to obtain in an appraisal situa-

tion. Cooperation wins out over objectivity. However, there are some practices that can help to reduce the probability of collusion.

The first one is a random assignment of raters. Freedom to choose who assesses, which is the usual practice in a 360-degree assessment, creates conditions for collusion, because it enhances the possibility that employees will orchestrate an outcome that is favorable to themselves (Ward, 1997). Common sense dictates that if the goal is appraisal and not development, the only individuals that will be selected by the person being evaluated are those he or she expects to be favorable for an evaluation. Therefore, in order to make collusion more difficult, employees should be randomly assigned to assess their co-workers. However, this approach contradicts the current practice of 360-degree assessment where the ratee has discretion over the appointment of raters. Coming from a developmental perspective, ratees would presumably choose raters who could provide critical but mostly constructive feedback. However, if ratings are used for appraisal, those being evaluated have no incentive to look for the most critical feedback and instead will most likely select people who will provide favorable ratings.

The second option is for less frequent use of 360-degree evaluation in order to counteract collusion when appraisals are being conducted. If an evaluation is repeated often, it is far easier to ascertain the gains from collusion. That said, it could be argued that collusion should be a concern even if the assessment is carried out annually. On the other hand, when deploying multi-source feedback for development, follow-up surveys after 6 months are far more desirable to measure an individual's progress. Again, there is a contradiction between what is best for the process in order to foster development and what is needed for appraisal assessment decisions.

The third factor that has to be considered is the degree of interdependence among the evaluators. Having work relationships, which span a long period of time may lead to stronger interdependence. While this interdependence is helpful from the perspective of developmental feedback (the evaluators have had more opportunities to know and assess the progress of each other), it is bad for performance assessment - the higher the interdependence, the higher the probability that evaluators will cooperate in evaluating each other more positively. This suggests that coupling development and performance assessment into one tool creates inherent contradictions, which hinder the optimal outcomes desired from each application.

The Organizational Level: A Psychometric Argument

One of the great appeals of 360-degree assessment is its numerical scoring. These convey the impression of objectivity and fairness. But there may be important tradeoffs made with these assessments when 360-degree assessment is employed for either development or appraisal. These include the desire for 'accuracy' over variation of information, the wish to aggregate data versus to discern differences, and the need to emphasize performance results over leadership competencies.

Ideographic versus nomothetic approach: When used for performance appraisal, the goal of 360-degree assessment is accuracy, while for development we will argue that it is honest perspectives even if they vary among evaluators or even contradict one another. In both cases the information is valid. However, our requirements towards its content are different. The strength of 360-degree assessment when used for development is that information comes from multiple perspectives since different constituencies provide data from their point of view. Multiple sources invoke different value systems as a result of their roles in the organizational hierarchy (Salam, Cox and Sims, 1997). Borman (1997) suggests that the organizational perspective of raters has a

significant effect on performance ratings. In other words, peers, direct reports and bosses observe different aspects of the working situation; therefore they focus their attention on different facets of employee's performance and attach differing weights to them. However, the different sources may have differing goals or expectations. For example, superiors often rate 'challenge to the status quo' and 'encouraging independent action' as negatively related to performance, while subordinates find a positive link between them (Salam et al., 1997). Different perceptions inevitably lead to limited agreement among raters (Cardy & Dobbins, 1994). This is not, however, a problem, if 360-degree assessment is used for developmental purposes. In this case, the intent is to capture rater variance and not to solicit like-minded opinions, because those receiving feedback will have more variation to learn from. The rationale is that differences in rater views reflect legitimate differences in the perceptions of the ratee's various roles. Also, significant discrepancies between the ratings of respondents might lead the individual receiving feedback to an awareness that he or she interacts with employees in different ways.

On the other hand, if 360-degree assessment is employed as an appraisal, the approach may well be driven by the desire to *minimize disagreements* – to seek a uniform measurement of the individual's overall performance. Consequently, there could be a strong desire to boost its 'validity' by using mechanistic approaches like discounting outliers and dropping respondents who produce ratings inconsistent with others (see Edwards & Ewen, 1996). The notion that inconsistency is completely valid and represents 'accuracy' is acceptable in the developmental paradigm, but may not be in the appraisal one. Under the appraisal view, the greater the agreement there is among raters the greater the accuracy of the overall feedback. From this point of view, assuming that the ratings are honest, any disagreement is considered to be a rater error and therefore less desirable.

In sum, while the appraisal approach has a nomothetic connotation and focuses on measurement, the developmental one is more concerned with the ideographic value of increasing variation.

Aggregation versus patterns: If organizations use 360-degree assessment for performance appraisal decisions, data across different sources has to be pulled together. This brings us to another caveat of using 360-degree assessment for performance appraisal concerning the aggregating of feedback from different raters. This process of averaging may conceal important variations (Fletcher & Baldry, 1999).

Differences in the raters' levels cause *method effects* (Conway, 1996) - in other words, bosses evaluate differently from peers, who assess differently from direct reports and so on. Moreover, research shows that inter-rater reliability for overall performance is not the same for the different reference groups. The consistency among bosses' ratings is the highest (.52), followed by that of peers (.42) and that of direct reports (.26) (Viswesvaran, Schmidt & Ones, 1996). Tsui & Ohlott (1988) had similar results. It seems that supervisors, peers and direct reports ground their assessments in different criteria. The different socialization paths of organizational members strongly influence their expectations.

In addition to the method effects, there is an *idiosyncratic rating tendency* of individual raters (Mount, Judge, Scullen, Sytsman & Hezlett, 1998). This means that assessment from each rater, regardless of level, captures unique variance, and these are different enough to constitute a separate method. The conclusion is that the widespread practice of aggregating ratings within or across rating levels is inappropriate because it reduces the construct validity of the ratings (Mount et al., 1998). The only exception is the level of supervisors where validity does not decrease as a result of aggregation. Bosses appear to share a more common frame of reference as a result of a more structured training and experience.

All these effects worsen the psychometric properties of 360-degree assessment. In addition, its accuracy is considerably restricted by cognitive limitations or intentional manipulations, leniency and harshness errors, central tendency, range restriction, halo effect, friendship bias, etc. Some of these errors may be caused by the desire of raters to maintain good working relationships or to show empathy for a colleague who has personal problems. All these examples should dim any enthusiasm to use 360-degree assessment as a psychometric instrument that accurately reflects performance or behavior. Given such dilemmas with the quantitative dimension of 360-degree assessment, it is worthwhile exploring the underutilized qualitative dimension. In order to draw richer conclusions on the developmental needs of managers, data analysis should move from a quantitative statistical comparison of averages and standard deviations to the discussion of response patterns to the written comments and examples that are cited. This is supported by focus groups, which have found that rating scales are not very helpful to understand others' expectations and instead recommend descriptive comments (Antonioni, 1996). Likewise, managers from medium-sized organizations studied by Waldman et al. (1998) found verbatim comments to be more valuable than numerical ratings.

Results versus competencies: Organizations have strong incentives to focus principally on an individual's performance outcomes, and as a byproduct they may fail to differentiate between competences and results (Coates, 1998). What 360-degree survey items capture are rarely the objective outcomes of an employee's activity and behavior, but *judgments* about their competencies. These judgments often incorporate non-performance factors, which violates the principle that appraisals should evaluate outcomes and not the personal style or personality (Bernardin et al., 1998), or as Bowman (1999:569) puts it, "verisimilitude trumps veracity". Moreover, research suggests that 360-degree assessment and performance evaluations reflect different types of performance assessments – 360-degree assessment reflecting what is termed 'contextual performance' and performance evaluations reflecting 'task performance' (Beehr, Ivanitskaya, Hansen, Erofeev & Gudanowki, 2001). In this case, contextual performance consists of non-job-specific behaviors such as dedication to the job or team or interpersonal skills or good citizenship (Borman & Motowidlo, 1997; Conway, 1999). In contrast, task performance consists of core task-specific behaviors. Consequently if 360-degree assessment is employed for performance evaluations, we cannot claim that it is only outcome-based, because contextual elements like good citizenship behavior often enter into the evaluation. Yet under conditions of an appraisal, people may wish to be evaluated for what they have accomplished and less for how they did it (Bookman, 1999).

In conclusion, the potential of 360-degree assessment as a development tool will be strongly restricted if 'accuracy' (the minimizing of disagreements) wins over variation. On the other hand, there are clear expectations for greater accuracy and objectivity of the instrument if 360-degree assessment is used for appraisal purposes. This dilemma suggests two different paths for the tools. Enhancing the qualitative properties of such multi-source feedback would increase its developmental potential and make it more effective in capturing variation. Similarly, for appraisal purposes, administrators would need to improve the psychometric properties of the feedback instrument.

Guidelines for Creating Two Distinct Tools

Our discussion has been aimed at illustrating some of the potential conflicts, contradictions and dilemmas that occur when the same 360-degree assessment tool and process are used

for multiple purposes. Based on this analysis, different conclusions could be drawn. One possibility is to infer that 360-degree assessment should *never* be used for appraisal decisions. Another one is to look for a *third way* between 'never' and 'always' (McCauley, 1997). We suggest exploring the latter. It may be that further research still leads us to 360-degree assessment for 'development-only' scenario, but before making that final judgment, we would suggest trying some important modifications.

The main asset of 360-degree assessment is its methodology: getting information from different individuals at different levels. Therefore, the strategy should be to keep the basic principle of multiple evaluation perspectives, but to 'tailor' the tool and the process according to the specific needs of either development or appraisal. This would imply that two tools should be developed: one, a more qualitatively based format where the 360-degree survey that is used solely for development, and another solely for performance appraisal. We need two instruments, because their content will be different. One of them will focus on competencies to be developed and the second one will reflect performance outcomes and performance-related behaviors to be measured. The two feedback tools would be used at separate times during a year to avoid the perception that both are part of the appraisal process. The development tool could have multiple applications over a year depending upon the desire of those being evaluated to have periodic assessments of their progress. The appraisal tool would be a once-a-year application given the problem of collusion. Ideally, application of the two would be separated by one to several months so as not to confound development with performance outcomes and to minimize the dilemmas described earlier around receptivity to learning versus the need to appear highly effective in one's role.

360-Degree Assessment for Development

We would suggest the following guidelines for designing 360-degree assessment for the developmental needs of an individual:

- The 360-degree assessment instrument is organized around competencies related to the strategic and talent needs of the company or the operating unit – especially those that have high value for the future and require extensive development across the organization. Items would be continually reviewed and updated to ensure their relevance in the face of changes in organizational demands. Rating of performance would not appear as part of the feedback. Rather this information would be provided through a second 360-degree assessment devoted solely to performance outcomes.
- Since raters observe different facets of the individual they are evaluating, raters should be subgrouped according to their level in the organization and asked to provide information only on dimensions that are relevant to their work with the individual being rated (Furnham & Stringfield, 1998). In other words, the 360-degree assessment instrument should be customized for direct reports, peers and bosses using different 'master lists' with behavioral descriptions that are relevant for the particular referent group. This approach suggests that raters can select items and individualize the survey choosing items for which they are best positioned to make reliable assessments (Kaplan, 1993; Westerman & Rosse, 1997). They do not have to respond to items and categories for which they are poorly positioned to provide useful assessments.

- Feedback is presented separately for each rater group (e.g. peers, direct reports, etc.) avoiding aggregation of data across sources (Yukl & Lepsinger, 1995). There should, however, be more than three respondents in each category, so that anonymity is guaranteed (London & Smither, 1995). Research suggests that normative feedback, which draws comparisons with the performance of others is not very effective and can even lead to performance decline (Butler, 1987). However, some ratees have a strong need to compare their feedback with that of others. Comparing averages of one manager with the averages of another will not be of particular help (see DeNisi & Kluger, 2000), unless there are reliable norms about specific industries and managerial levels, which would make the comparison meaningful.
- The potential of 360-degree assessment as a *qualitative* instrument has been largely neglected. It is not that qualitative feedback is better than quantitative, but the combination of both adds to the richness of the process. While both the development and the appraisal tools would have a qualitative dimension, the former would have a greater proportion of qualitative questions. It could significantly enhance the properties of the tool, especially for developmental purposes to have open-ended questions. Stronger emphasis on detailed written comments will help managers to learn in what specific ways they are different from the others and therefore to capitalize on what's unique about themselves in leadership roles (Goffee & Jones, 2000). Rarely can items on 360-feedback surveys reveal those differences that are more subtle and not immediately apparent, but qualitative feedback can. In other words, it can add richness of information to numeric feedback. Given issues of respondent fatigue, open-ended qualitative questions could precede the quantitative items. In addition, for each item requiring a quantitative ranking on the questionnaire, space should be provided for qualitative comments that explain and justify the numerical evaluation (Ghorpade, 2000). A request should be made that qualitative comments be supported by recent, highly illustrative examples rather than general statements. This would provide opportunities to link the quantitative with the qualitative sections.
- Qualitative questions should be of different types. Some of them can be the conventional ones, for example: "What specific behaviors can your manager/peer perform to make you more productive in your job?"; "How would you describe the capabilities requiring development for your manager/peer?", "What in his or her behavior makes it difficult to work with your manager/peer?", or "What strength of this individual has the potential to become a liability in the near future?". Other qualitative questions, however, might be more creative: "Please give three positive and three negative adjectives that best describe your manager/peer, explain why and give specific examples to support your choice." Some 360-degree programs have excellent experience with projective questions (e.g. "If you had to characterize your manager/peer as a verb/color/shape, which one would you choose and why?"). Talking about a person in more metaphorical terms can make a problem more salient and can ease its discussion. Projective techniques have long traditions in psychology and are successfully used in marketing research (Crumbaugh, 1990; Donoghue, 2000).
- Those individuals receiving feedback should be trained in how to select respondents in order to leverage diversity of viewpoints about their various competences, how to analyze structured feedback and qualitative comments, and how to develop a personal development plan. In addition to training those being rated, evaluators should also be trained in how best to provide informative qualitative feedback for development. For example, they could be taught how to write brief but highly illustrative case examples which highlight a particular strength

or developmental area. They could also instructed in how best to provide suggestions for action steps.

- The ratees should have full ownership of the information. The process should be frequently conducted, and the target individual can request a 360-degree assessment with a specific focus depending on the circumstances. For example, if the department is to undergo restructuring, he/she may ask for an emphasis of the survey feedback on leading change.
- Ratee's accountability to raters, top management, and also themselves has often been cited as one of the key factors that can enhance the effectiveness of 360-degree feedback and thus lead to significant behavior change (London, Smither, & Adsit, 1997). Therefore, follow-up activities and sharing the developmental plan with mentors can have a positive impact on the process of multisource feedback.
-

In sum, we suggest a more qualitative approach when multi-source feedback is used for development because it gets us much closer to the goal of personal growth. In addition, the recipients of feedback need to 'own' the data themselves rather than the organization.

360-Degree Assessment for Appraisal Decisions

In order to be acceptable for performance appraisal decisions, administrators need to enhance the psychometric properties of 360-degree assessments. With this in mind, we suggest the following guidelines:

- The 360-degree instrument should be designed to record performance outcomes and performance related behaviors. A meta-analysis involving 50 studies and 8,341 individuals highlighted that objective measures and subjective indicators of performance are only moderately related and are not interchangeable (Bommer, Johnson, Rich, Podsakoff, & Mackenzie, 1995). Therefore, the goal should be to achieve precision in defining performance dimensions and to establish performance standards that focus on outcomes that can be measured in terms of relative frequency of a behavior (Bernardin et al., 1998). Research shows that clear performance standards increase inter-rater reliability (Viswesvaran et al., 1996), which is beneficial for the impression of objectivity and consequently for the acceptance of the ratings by the individual being evaluated. Measures should be related to values such as quantity, quality, or cost (Bernardin et al., 1998).
- Before an individual is assessed, the individual receiving feedback should participate actively with his/her supervisors and human resource managers in the selection of the performance dimensions. The latter have to be explained and if necessary, should be highly specified to fit the particular job. Items could be customized to the particular performance demands of an individual manager. In order to minimize impression management, performance dimensions that can be easily manipulated, like 'friendliness' for example, should be eliminated.
- Multiple feedback sources such as direct reports, peers, clients and supervisors should participate as raters. These individuals should be carefully chosen in conjunction with supervisors because of the intensive interaction with the individual receiving feedback and the possibilities that raters have had to observe the individual receiving feedback in working situations. The aim here is twofold: one is to minimize collusion, and the other is to get feedback

from those best positioned to provide useful assessments. Earlier, we discussed the possibility of random selection. This is a feasible alternative if the pool of evaluators is sufficiently large and diverse. One further possibility to reduce the amount of collusion is to select as ratees people who are in a sense customers of the ratee's leadership behavior.[1]

- The performance appraisal instrument can benefit from written comments that explain and illustrate the evaluations. This practice would be especially valuable around dimensions that require attention on the focal individual's part. In such cases, illustrations of the existing behavior as well as detailed explanations of the consequences or outcomes of this behavior would be most helpful. Respondents would also be encouraged to provide concrete suggestions for improvement – again with illustrations of more performance-oriented behaviors or actions.

- When used for performance appraisal or other administrative decisions, people expect a high degree of accuracy and training of their raters. As such, developing a culture where individuals can reach a relative consensus on performance evaluation issues becomes extremely important. Open discussions on disagreements on performance standards and training in overcoming judgement biases can help create this type of culture. Self-ratings therefore should be included in the assessment to gauge the degree of divergence of perspectives. In addition, the process should be able to reflect unusual circumstances. For example, a manager may be placed in an extremely difficult turnaround situation. Their overall performance may be less a reflection of their capabilities and more a reflection of external circumstances, which are beyond the manager's control. In other cases, the focal individual's peers may have little knowledge of the assignment or role yet due to administrative requirements their assessments must be included in the appraisal. In both of these circumstances, the evaluation must be flexible enough to reflect these unique conditions and not penalize the focal individual unfairly for performance. Objective and measurable performance standards counteract collusive behavior. In order to guarantee the integrity of the performance appraisal system, raters and their evaluators should be accountable to formal review by superiors when there are contested assessments. Ideally, a review group could include one's superior plus this individual's superior and several superiors of the evaluators. In addition, the raters have an obligation to provide feedback on an individual's performance deficiencies along with opportunities on how to correct them on a regular basis (Bernardin et al., 1998). In other words, continuous feedback on an informal basis is required along with the occasional application of 360-degree assessment.

Performance appraisal is such a complex system that the introduction of multiple sources of information may increase its validity, but it could also create new problems. Involving more people in the process of performance evaluation means more organizational time for filling out questionnaires, more individuals feeling uncomfortable with their role as evaluators, and new pressure for employees to keep records of the performance of their colleagues in order to be precise in their assessments. Organizations have to decide to what extent the benefits of these assessments outweigh the downsides.

That said, it is clear that 360-degree feedback can play an important role in a manager's learning and development. Managers do change as a result of feedback interventions. Kluger & DeNisi's (1996:275) meta-analysis suggests that "feedback interventions improve performance

---

[1] We thank our anonymous reviewer for this idea.

by approximately .4 of a SD", but there is a large variability of effects. Their feedback intervention theory explains what drives the change process following 360-degree feedback and argues that improvement depends on the level at which the feedback intervention focuses ratee's attention. In their view, our behavior is regulated by a comparison of the received feedback with a goal or a standard. The latter are arranged hierarchically. The highest one is the *self-level* where gaps make us question our self-concept. The second one is the *task-level* where discrepancies between feedback and goals make us work harder. The third one is the *task learning level* where attention is distracted to specific details of our performance at the expense of actual accomplishments. Since attention is a scarce resource, behavior will be regulated only by those feedback-standard gaps that receive sufficient attention. Therefore, feedback interventions change the locus of attention. The best strategy is to focus feedback on the task, because in this case, people are concerned with narrowing the gap between actual and goal performance. If efforts at this level fail, attention will be focused on the task-learning level. However, a serious problem occurs when attention is shifted to the level of the self, and focal leaders start questioning who they really are. In this case, subsequent performance may well suffer because of strong affective reactions, like disappointment or despair that could be produced.

Managers improve their management skills especially when 360-degree feedback is applied at regular intervals over a longer period of time. A longitudinal study of five annual administrations of upward feedback confirms the generalizability of its positive effects (Walker & Smither, 1999). Managers with initially low or moderate ratings improved even more compared to those who were rated high at the beginning of the study, which was probably due to ceiling effects. But the most interesting finding was about the role of the post-rating phase. It clearly demonstrated that what managers do with the evaluations really matters. Those of them who held feedback sessions with their direct reports to discuss their results improved more than those who did not. This conclusion was further verified by a within-subject analysis. The latter confirmed that managers improved more in those years when they held feedback meetings with their direct reports compared to those years when they did not have those meetings. Holding meetings seems to increase accountability, which leads to significant improvement. The results of this longitudinal study also suggest that organizations have to be more patient anticipating results from multisource feedback. They should not give up when the first follow-up shows no significant improvement, because scores improve gradually over a longer period of time.

Managers' attitudes have a strong impact on the learning and improvement that follow multisource feedback. Cynicism about the feedback process, for example, is negatively related to personal change (Atwater et al., 2000). In other words, before introducing a multisource feedback system, organizations should train ratees and should try to create positive attitudes towards the process (Antonioni, 1996). In situations where the feedback is used for punitive purposes, cynicism is likely to be far higher, and such uses may raise important ethical concerns.

In sum, for learning and improvement of management performance to occur, ratees should know in advance that they are held accountable for the use of feedback. Skillful facilitation of the feedback intervention, follow-up activities that enhance the perception of accountability, development of action plans for personal growth, tracking results over time, commitment to behavior change in public and participating in training that is closely linked to feedback results have positive effects on learning from multisource feedback (London et al., 1997). In the case of training, it is ideal to build the topics directly around the actual feedback dimensions - particularly the dimensions most in need of developmental attention. Our personal experience with feedback groups also shows that managers who have mentors have greater opportunities to learn

from feedback, which might be due to the accountability context that is created. A special case that works particularly well is when the mentor is a retired manager from the same company.

Learning from 360-degree feedback is a promising area for research. Field studies designed as longitudinal projects seem to have the highest potential to test the effectiveness of multisource feedback. Accountability of ratees has emerged as an important factor that influences the effectiveness of feedback. However, there are still many unanswered questions about it. For example, is process accountability more suitable than outcome accountability, or which constituency represents the best accountability context? Further questions of interest are whether written comments are more useful for learning than ratings, and which follow-up activities are most effective for personal growth.

In conclusion, the idea that feedback for development should be separated from performance appraisal is not new and goes back to Meyer, Kaye, & French (1965). A significant amount of evidence follows this tradition and suggests that when organizations try to blend development and appraisal, ratings will be less reliable and less valid (Farh, Canella & Bedeian, 1991; Murphy & Cleveland, 1995; London & Smither, 1996). However, recent publications (Bracken et al., 2001:4) argue that "multisource feedback for decision making *can* work under the right conditions". Since the debate is far from resolved, we have tried to systematize the arguments in favor of the separation thesis. We believe that two 360-degree assessment tools and two different processes are necessary for development versus administrative decisions such as performance appraisal or succession decisions. One would complement the other, and they would be deployed at separate times during a year. The more qualitative one is competence-oriented and targets solely development. The more quantitative instrument is performance outcome-oriented measuring specific outcomes and performance-related behaviors, and it is used for administrative decisions. What should be preserved is the methodology of obtaining multiple points of view. Content and process, however, need to differ given the dilemmas we have highlighted throughout this article. Given the trend towards using 360-degree assessment for multiple purposes, now is indeed the time to reinvent a tool whose efficacy is increasingly in question.

References

Antonioni, D. 1996. Designing an effective 360-degree appraisal feedback process. *Organizational Dynamics,* 25(2):24-38.

Ashford, L., & Tsui, A. 1991. Self-regulation for managerial effectiveness: The role of active feedback seeking. *Academy of Management Journal, 34:* 251-280.

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., Waldman, D. & Brett, J. 2002. Understanding and optimizing multisource feedback. *Human Resource Management Journal*, 41: 193-208.

Baumeister, R. 1998. The self. In D. Gilbert, S. Fiske & G. Lindzey (Eds.), *The handbook of social psychology,* vol. 1: 680-740. Boston, MA: McGraw-Hill.

Baumeister, R. & Cairns, K. 1992. Repression and self-presentation: When audiences interfere with self-deceptive strategies. *Journal of Personality and Social Psychology,* 62:851-862.

Beehr, T. A., Ivanitskaya, L., Hansen, C. P., Erofeev, D., & Gudanowski, D. M. 2001. Evaluation of 360 degree feedback ratings: Relationships with each other and with performance and selection predictors. *Journal of Organizational Behavior*, 22: 775-788.

Bernardin, H., & Beatty, R. 1984. *Performance appraisal: Assessing human behavior at work*. Boston: Kent.

Bernardin, J., Hagan, C., Kane, J., & Villanova, P. 1998. Effective performance management: A focus on precision, customers, and situational constraints. In J. Smither (Ed.), *Performance appraisal. State of the art in practice:3-48.* San Francisco: Jossey-Bass Publishers.

Bettenhausen, K. & Fedor, D. 1997. Peer and upward appraisals: A comparison of their benefits and problems. *Group & Organization Management,* 22:236-263.

Binmore, K. 1994. *Playing fair.* Cambridge, MA & London: The MIT Press.

Bommer, W. H., Johnson, J. L., Rich, G. A., Podsakoff, P. M., & Mackenzie, S. B. 1995. On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology*, 48: 587-605.

Bookman, R. 1999. Tools for cultivating constructive feedback. *Association Management,* 51(2):73-79.

Borman, W. 1997. 360 degree ratings: An analysis of assumptions and research agenda for evaluating their validity. *Human Resource Management Review,* 7:299-315.

Borman, W. & Motowidlo, S. 1997. Task performance and contextual performance: The meaning of personnel selection research. *Human Performance*, 10: 99-109.

Bowman, J. 1999. Performance appraisal: verisimilitude trumps veracity. *Public Personnel Management,* 28:557-576.

Bracken, D. 1997. Maximizing the use of multi-rater feedback. In D. Bracken, M. Dalton, R. Jako, C. McCauley, & V. Pollman (Eds.), *Should 360-degree feedback be used only for developmental purposes?:11-22.* Greensboro, NC: Center for Creative Leadership.

Bracken, D., Timmreck, C., Fleenor, J., & Summers, L. 2001. 360 feedback from another angle. *Human Resource Management*, *40*: 3-20.

Brett, J. & Atwater, L. 2001. 360 degree feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology*, 86: 930-942.

Butler, R. 1987. Task-evolving and ego-evolving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, 79: 474-482.

Cardy, R. & Dobbins, G. 1994. *Performance appraisal: Alternative perspectives.* Cincinnati: South-Western Publishing.

Church, A. 1994. Managerial self-awareness in high performing individuals and organizations. *Dissertation Abstracts International,* 55-OSB:2028.

Coates, D. 1998. Don't tie 360 feedback to pay. *Training,* 35(9):68-78.

Coens, T. & Jenkins, M. 2000. *Abolishing performance appraisals: Why they backfire and what to do instead.* Berrett-Koehler Publishers.

Conway, J. 1999. Distinguishing contextual performance from task performance for managerial jobs. *Journal of Applied Psychology*, 84:3-13.

Conway, J. 1996. Analysis and design of multitrait-multirater performance appraisal studies. *Journal of Management,* 22:139-162.

Crary, W. 1966. Reactions to incongruent self-experiences. *Journal of Consulting Psychology,* 30:246-252.

Crocker, J. & Major, B. 1989. Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review,* 96:608-630.

Crumbaugh, J. 1990. The meaning of projective techniques. In J. Crumbaugh, J. Graca, R. Hutzell, M. Whiddon, & E. Cooper (Eds.), *A primer of projective techniques of psychological assessment*. California: Libra.

Daft, R. & Weick, K. 1984. Toward a model of organizations as interpretive systems. *Academy of Management Review,* 9:284-295.

Dalton, M. 1997. When the purpose of using multi-rater feedback is behavior change. In D. Bracken, M. Dalton, R. Jako, C. McCauley, & V. Pollman (Eds.), *Should 360-degree feedback be used only for developmental purposes?:1-6.* Greensboro, NC: Center for Creative Leadership.

DeNisi, A. & Kluger, A. 2000. Feedback effectiveness: Can 360-degree appraisals be improved? *Academy of Management Executive,* 14(1):129-139.

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

DeNisi, A. & Williams, K. 1988. Cognitive approaches to performance appraisal. In G. Ferris & K. Rowland (Eds.), *Research in personnel and human resource management*. Greenwich, CT: JAI Press.

Donoghue, S. 2000. Projective techniques in consumer research. *Journal of Family Ecology and Consumer Sciences*, 28: 47-53.

Dutton, J. &   Jackson, S. 1987. Categorizing strategic issues: Links to organizational action. *Academy of Management Review,* 12:76-90.

Edwards, M. & Ewen, A. 1996. How to manage performance and pay with 360-degree feedback. *Compensation and Benefit Review,* 28(3):41.

Farh, J., Cannella, A., & Bedeian, A. 1991. Peer ratings: The impact of purpose on rating quality and user acceptance. *Group and Organization Studies*, *16*:367-386.

Fletcher, C. 1998. Circular argument. *People Management,* 4:346-349.

Fletcher, C. 1999. The implication of research on gender differences in self-assessment and 360-degree appraisal. *Human Resource Management Journal,* 9:39-46.

Fletcher, C. & Baldry, C. 1999. Multi-source feedback systems: a research perspective. In C. Cooper & I. Robertson (Eds.), *International Review of Industrial and Organisational Psychology, vol. : .* Chichester: Wiley.

Fletcher, C. & Baldry, C. 2000. A study of individual differences and self-awareness in the context of multi-source feedback. *Journal of Occupational and Organizational Psychology,* 73:303-319.

Frank, R. 1988. *Passion with reason.* New York: Norton.

Fredrickson, J. 1985. *Effects of decision motive and organizational performance level on strategic decision processes. Academy of Management Journal*, 28:821-843.

Furnham, A., & Stringfield, P. 1998. Congruence in job-performance ratings: A study of 360 degree feedback examining self, manager, peers, and consulting ratings. *Human Relations*, *51*: 517-530.

Ghorpade, J. 2000. Managing five paradoxes of 360-degree feedback. *Academy of Management Executive*, *14*(1): 140-150.

Goffee, R., & Jones, G. 2000. Why should anyone be led by you? *Harvard Business Review*, *78*(5): 62-70.

Gollwitzer, P., & Kinney, R. 1989. Effects of deliberative and implemental mind-sets of illusion of control. *Journal of Personality and Social Psychology*, 56:531-542.

Harrison, M. & Phillips, B. 1991. Strategic decision making: An integrative explanation. *Research in the Sociology of Organizations,* 9:319-358.

HRFocus. 2000. HR update: HR execs dissatisfied with their performance appraisal systems. *HRFocus,* January:2.

Jako, R. 1997. Fitting multi-rater feedback into organizational strategy. In D. Bracken, M. Dalton, R. Jako, C. McCauley, & V. Pollman (Eds.), *Should 360-degree feedback be used only for developmental purposes?: 19-22.* Greensboro, NC: Center for Creative Leadership.

Jones, J. & Bearley, W. 1996. *360deg feedback: strategies, tactics and techniques for developing leaders.* Amherst, MA: HRD Press.

Jones, E. & Pittman, T. 1982. Toward a general theory of strategic self-presentation. In J. Suls (Ed.) *Psychological perspectives on the self*: 231-262. Hillsdale, NJ: Erlbaum.

Judge, T. & Ferris, G. 1993. Social context of performance evaluation decisions. *Academy of Management Journal,* 36:80-105.

Kahneman, D., Knetsch, J. & Thaler, R. 1990. Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy,* 98:1325-1348.

Kaplan, R. 1993. 360-degree feedback PLUS: Boosting the power of co-worker ratings for executives. *Human Resource Management*, *32*: 299-314.

Kennedy, M. 1999. The case against performance appraisal. *Across the Board,* January:51-52

Kreps, D. M. 1990. *A course in microeconomic theory.* London: Harvester Wheatsheaf.

Lepsinger, R. & Lucia, A. 1997. 360 degree feedback and performance appraisal. *Training,* 34(9):62-70.

London, M. & Smither, J. 1995. Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. *Personnel Psychology,* 48:803-839.

London, M., Smither, J. & Adsit, D. 1997. Accountability: The Achilles' heel of multisource feedback. *Group & Organization Management*, 22(2): 162-184.

London, M., & Wohlers, A. 1991. Agreement between subordinate and self-ratings in upward feedback. *Personnel Psychology*, *44*: 375-390.

Longenecker, C., & Ludwig, D. 1990. Ethical dilemmas in performance appraisals revisited. *Journal of Business Ethics*, *9*: 961-969.

Malos, S. 1998. Current legal issues in performance appraisal. In J. Smither (Ed.), *Performance appraisal. state of the art in practice:49- .* San Francisco: Jossey-Bass Publishers.

Marks, G. 1984. Thinking one's abilities are unique and one's opinions are common. *Personality and Social Psychology Bulletin,* 10:203-208.

Markus, H., & Zajonc, R. 1985. The cognitive perspective in social psychology. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology,* vol. 1: 137-230. New York: Random House.

McCauley, C. 1997. On choosing sides: seeing the good in both. In D. Bracken, M. Dalton, R. Jako, C. McCauley, & V. Pollman (Eds.), *Should 360-degree feedback be used only for developmental purposes?:23-36.* Greensboro, NC: Center for Creative Leadership.

Meyer, H., Kay, E. & French, J. 1965. Split roles in performance appraisal. *Harvard Business Review*, 43(January/February): 123-129.

Mintzberg, H., Raisanghani, D. & Theoret, A. 1976. The structure of the 'unstructured' decision processes. *Administrative Science Quarterly,* 21:246-275.

Mount, M., Judge, T., Scullen, S., Sytsman, M. & Hezlett, S. 1998. Trait, rater and level effects in 360-degree performance ratings. *Personnel Psychology,* 51:557-576.

Murphy, K., & Cleveland, J. 1995. *Understanding performance appraisal: Social, organizational, and goal-based perspectives.* Thousand Oaks, CA: Sage.

Nutt, P. 1993. The formulation processes and tactics used in organizational decision making. *Organization Science,* 4:226-251.

Pollack, D. & Pollack, L. 1996. Using 360-degree feedback in performance appraisal. *Public Personnel Management,* 25:507-528.

Pollman, V. 1997. Some faulty assumptions that support using multi-rater feedback for performance appraisal. In D. Bracken, M. Dalton, R. Jako, C. McCauley, & V. Pollman (Eds.), *Should 360-degree feedback be used only for developmental purposes?:7-9.* Greensboro, NC: Center for Creative Leadership.

Podsakoff, P. & Fahr, J. 1989. Effects of feedback sign and credibility on goal setting and task performance. *Organizational Behavior and Human Decision Processes*, 44:45-67.

Rabinowitz, L., Kelley, H. & Rosenblatt, R. 1966. Effects of different types of interdependence and response conditions in the minimal social situation. *Journal of Experimental Social Psychology,* 2:169-197.

Rogers, T., Kuiper, N. & Kirker, W. 1977. Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology,* 35:677-688.

Salam, S., Cox, J. & Sims, H. 1997. In the eye of the beholder: How leadership relates to 360-degree performance ratings. *Group & Organization Management,* 22:185-209.

Scholtes, P. 1999. Performance appraisal: State of the art in practice. *Personnel Psychology,* 52:177-181.

Sedikides, C. 1993. Assessment, enhancement, and verification determinants of the self-evaluation process. *Journal of Personality and Social Psychology,* 65:317-338.

Taylor, S. 1989. *Positive illusions: Creative self-deception and the healthy mind.* New York: Basic Books.

Taylor, S. & Brown, J. 1988. Illusion and well-being; A social psychological perspective on mental health. *Psychological Bulletin,* 103:193-210.

Tornow, W. 1993. Perceptions or reality: Is multi-perspective measurement a means or an end? *Human Resource Management,* 32:221-229.

Tsui, A. & Ohlott, P. 1988. Multiple assessment of managerial effectiveness: Interrater agreement and consensus in effectiveness models. *Personnel Psychology,* 41:779-803.

van Damme, E. 1991. *Stability and perfection of Nash equilibria.* Berlin: Springer Verlag.

Viswesvaran, C., Schmidt, F. & Ones, D. 1996. Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology,* 81:557-574.

Waldman, D., Atwater, L. & Antonioni, D. 1998. Has 360 feedback gone amok? *The Academy of Management Executive,* 12(2):86-94.

Waldman, D. & Bowen, D. 1998. The acceptability of 360 degree appraisals: A customer-supplier relationship perspective. *Human Resource Management*, 37: 117-129.

Walker, A. G. & Smither, J. W. 1999. A five-year study of upward feedback: What managers do with their results matters. *Personnel Psychology*, 52(2): 393-423.

Ward, P. 1997. *360-degree feedback.* London: Institute of Personnel Development.

Westerman, J., & Rosse, J. 1997. Reducing the threat of rater nonparticipation in 360-degree feedback systems: An exploratory examination of antecedents to participation in upward ratings. *Group & Organization Management*, *22*: 288-309.

Yammarino, F. & Atwater, L. 1993. Understanding self-perception accuracy: Implications for human resource management. *Human Resource Management,* 32:231-247.

Yukl, G., & Lepsinger, R. 1995. How to get the most out of 360 degree feedback. *Training*, *32*(12):45.

TABLE 1

**Relationship Between the Level of Analysis, the Type of Arguments
and the Resulting Contradictions and Dilemmas**

Level of Analysis	Type of Arguments	Dilemmas & Contradictions
Individual	Cognitive	'To be effective' vs. 'To be rated as effective'
Interpersonal	Game-theoretical	Accuracy and objectivity vs. reciprocity and collusion
Organizational	Psychometric	Psychometric tool vs. qualitative instrument