Assessing innovation: A 360-degree appraisal study

Digital Object Identifier (DOI):
10.1111/ijsa.12012

Running head: ASSESSING INNOVATION

Assessing innovation: A 360-degree appraisal study

Kristina Potočnik
University of Edinburgh
Kristina.Potocnik@ed.ac.uk

Neil Anderson
Brunel University
Neil.Anderson@brunel.ac.uk

Forthcoming in International Journal of Selection and Assessment

Author note

Requests for reprints should be addressed either to Kristina Potočnik, University of Edinburgh Business School, 29 Buccleuch Place, Edinburgh, EH8 8JS, UK (Kristina.Potocnik@ed.ac.uk) or to Neil Anderson, Brunel University, Business School, Uxbridge, UB8 3PH, UK (Neil.Anderson@brunel.ac.uk).

The authors wish to thank the consultancy organization concerned with this 360-degree appraisal instrument for their kind co-operation, in accordance with their wish to remain anonymous. We also thank the Action Editor for his handling of this manuscript and the reviewers for their constructive comments on earlier versions of this paper. We also thank Jo Silvester and Blanka Tacer for their valuable comments and suggestions that helped us improve the earlier versions of this paper. This research was supported by British Academy grant number SG110409 awarded to the first author, and by Leverhulme Trust grant number IN-2012-095 awarded to the second author.

Abstract

This study examines the evaluation of innovation in a wider competency framework and within a 360-degree rating procedure among managerial-level job holders. The total sample of 2979 individuals consisted of 296 target employees and their 318 bosses, 1208 peers, 828 direct reports and 329 others who provided ratings on a competency framework. The results showed significant differences in innovation-related competence ratings between different raters. Self-ratings were significantly lower compared to the overall observer ratings and were correlated only with peer ratings. Different patterns of results were found for the lower and upper quartiles based on self-ratings. For instance, no correlations were observed between self-ratings and the ratings of any observers in the group of best self-rated individuals. Implications for practice and future research in assessment and evaluation of innovation are discussed in conclusion.

Introduction

Innovative performance has been argued to be essential for organizational success and survival over several decades. Hence, not surprisingly a considerable body of research has built up on phenomena involving innovation and creativity during past decades at different levels of analysis – the individual, workplace teams, and at the organizational level-of-analysis (Anderson, De Dreu, & Nijstadt, 2004; West, Smith, Feng, & Lawthom, 1998). As for the individual level, the existing body of research has employed a wide variety of measures to assess employee innovative performance. Most commonly, these have consisted of a survey-based questionnaire measures to which employees have responded both for predictor variables and outcome measures of innovation. In other words, the most commonly applied approach to assess individual innovation so far has been self-reports, and this has extended not only to the measurement of variables deemed to be predictive of innovation but also of outcome measures of innovativeness in the workplace (Hülsheger, Anderson, & Salgado, 2009).

While the use of self-reports is a convenient way to collect data as they are quick and easy to administer and facilitate the data collection on large samples of respondents, this measurement method is subjected to different biases which might potentially distort the true score or feelings of the respondents about the phenomena under consideration (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Hence, the exclusive use of self-reports to assess innovative performance might be questionable and other more appropriate ways of assessing this type of performance should be explored. For instance, in their meta-analysis Hülsheger et al. (2009) found stronger correlations between predictors and innovation-related outcomes in studies that used exclusively self-reports compared to those that used independent measures of innovation. Although smaller correlations with the independent measures of innovation could be due to lower reliability of this type of measures these results could also support the concern that self-reports are inflated with a common-method variance.

Another issue that calls for more empirical investigation is differentiating innovative performance from other types of performance indicators. While research into individual innovation has witnessed an exponential growth, most of the models in the performance domain do not include innovative performance at all (e.g., Campbell, McCloy, Oppler, & Sager, 1993; Hunt, 1996). Most commonly, these models would cover job-specific and non-job specific task proficiency, maintaining personal discipline, supervision and leadership, schedule flexibility, among others. Perhaps innovative dimension of performance was not included in most of the performance frameworks, because scholars were concerned with studying generic (i.e. non-job specific) work behaviors, assuming that innovation is not part of all jobs. However, past research has shown that innovation is present in any type of job, ranging from shop-floor (Axtell et al., 2000) to top management (West & Anderson, 1996). Moreover, the fact that performance assessment has traditionally focused on job-related behaviors might have led to "*the omission of other important components of overall performance*" (Welbourne, Johnson, & Erez, 1998, p.541), such as innovation. Therefore, more empirical research is needed to demonstrate that innovative performance represents a conceptually distinct construct compared to other performance aspects.

Taking into account these arguments, the aim of the present study is twofold: (1) to evaluate the construct validity of innovative performance (in terms of innovation-related competence), and (2) to analyze the measurement of innovation-related competence within a multisource, 360-degree framework by examining the relationships between self-ratings and the ratings of independent observers, including boss, peers, direct reports, and others. To the best of our knowledge, no studies so far have addressed innovation within a 360-degree framework, and thus this type of measure of performance has remained curiously underresearched either in the performance management or innovation literatures (Moneta, Amabile, Schatzel, & Kramer, 2010). Examining the innovation competence in the context of multisource framework is important not only for assessing the individual innovation competence in a more reliable way but also because this method provides feedback from peers, managers, direct reports and others that might enhance employee's understanding of his or her needs to develop this competence further. Thus, our findings might have important contributions for the assessment of individual innovative performance for both future research and practice in the field of HRM.

Innovative performance in the workplace: Conceptual delimitation and measurement

Innovative performance in the workplace embraces both the generation of creative ideas in the first step and their implementation in the second step (Hülsheger et al., 2009; West & Farr, 1990). Noting these points, the present authors have recently put forward the following integrative definition of workplace creativity and innovation:

'Creativity and innovation at work are both the process and outcomes of attempts to develop and introduce new and improved ways of doing things. The creativity stage of this process refers to idea generation, and innovation to the subsequent stage of implementing ideas toward better procedures, practices, or products. Creativity and innovation can occur at the level of the individual, work team, organization, or at more than one of these levels combined, but will invariably result in identifiable benefits at one or more of these levels-of-analysis' (Anderson, Potočnik, & Zhou, submitted). This definition clearly distinguishes innovation from creativity which has been defined as the generation of novel and useful ideas (Amabile et al., 2002; Rank, Pace, & Frese, 2004) and thus represents the first part of the innovation process. Previous research has explored different antecedents of innovation at individual, team, and organizational levels, yielding a vast amount of findings regarding which factors enhance or inhibit innovation in the workplace. For instance, openness to experience (Baer, 2010; George & Zhou, 2001), personal initiative (Frese & Zapf, 1994), or autonomy at the workplace (Axtell et al., 2000; Axtell, Holman, & Wall, 2006) have been consistently positively linked to innovation at the individual level. At the team level, shared vision (Gillson & Shalley, 2004; Pearce & Ensley, 2004; West & Anderson, 1996), support for innovation (Hülsheger et al., 2009), diversity (Chi, Huang, & Lin, 2009) and task and goal interdependence (Gilson & Shalley, 2004; Zhang, Hempel, Han, & Tjosvold, 2007) were found to be conducive of innovation. At the organizational level, factors such as structure (West, Smith, Feng, & Lawthom, 1998), organizational and human capital (Subramaniam, & Youndt, 2005) and culture (Baer & Frese, 2003; Berson, Oreg, & Dvir, 2008; West & Anderson, 1992) have been found to influence innovation.

Despite this increasing interest in studying innovation in the workplace, and as noted above, only a few models of performance have included an “innovation” dimension within their taxonomies (e.g., Bartram, 2005; Welbourne et al., 1998). Thus, research to date has not provided sufficient evidence, either conceptual or empirical, about innovative performance being conceptually different from other dimensions of employee performance. Some studies, however, have addressed different individual performance indicators, including innovation, but they did not empirically examine their construct validity. For instance, Miron, Erez, and Naveh (2004) differentiated between three dimensions of individual performance: innovation, quality, and efficiency. Although they found significant correlations between these three dimensions, they did not examine whether they conceptually differ from each other.

There are, however, some studies that examined innovative dimension of performance. For instance, Welbourne et al. (1998) tested a theory-driven questionnaire of employee performance. Based on role and identity theories, their model includes five roles to measure the components of employee behavior that encompass different aspects of performance, namely job, organization, career, innovator and team roles. Their findings supported the theoretically predicted five-role model of employee performance, showing that the innovator role was a conceptually distinct aspect of employee performance. More recently, Bartram (2005) presented evidence for Great Eight competencies framework that cover the following competencies: leading and deciding, supporting and cooperating, interacting and presenting, analyzing and interpreting, creating and conceptualizing (innovation competence is included here), organizing and executing, adapting and coping, and enterprising and performing. The principle component analysis provided support for the model showing creating/conceptualizing competence as a distinct factor. Hence, there is some evidence that innovative performance represents a conceptually different aspect of performance, but clearly more research is needed to support this argument.

Along with the conceptual delimitation of innovative performance and demonstrating its construct validity, it is important to address how innovative performance can be measured in a reliable and valid way. The vast majority of studies into innovation has applied self-ratings (e.g. Axtell et al., 2000; Clegg, Unsworth, Epitropaki, & Parker, 2002; Ohly, Sonnentag, & Pluntke, 2006). Fewer studies have used independent or observer ratings, such as supervisor ratings (Oldham & Cummings, 1996; Yuan & Woodman, 2010), peer ratings (Alge, Ballinger, Tangirala, & Oakley, 2006; Amabile et al., 2002), or expert ratings (Shalley, 1995; Shalley & Perry-Smith, 2001). Yet, fewer studies assessed innovation based on objective criteria, such as counting the number of contributions to a suggestion system or the number of patents or new products (Cardinal, 2001; Frese, Teng, & Wijnen, 1999; Latham & Braun, 2009; Pirola-Merlo & Mann, 2004).

There are well-documented limitations to the sole use of self-reports to assess any phenomena, including innovative performance (Podsakoff et al., 2003). For instance, individuals might reply to the items in a way to show them in a more favorable fashion. Assuming that exhibiting high innovative performance is considered as a positive behavior by the organizations, employees might self-assess their innovative behaviors higher than they actually are. Further evidence for this is presented in the meta-analysis by Hülsheger et al. (2009). Here, the authors presented findings of substantial differences between correlations from innovation studies at the workgroup level that had used self-report outcome measures compared with those having used independent outcome measures of innovation. Studies using self-report measures throughout found higher correlations by some margin than those that had utilized more robust independent outcome measures of innovation, thus confirming earlier concerns over possible percept-percept bias in self-rating studies (e.g. Anderson et. al, 2004; West & Farr, 1990). In order to rule out any misperception of one's own innovative performance, studies should examine multiple sources simultaneously within the same study, including the employees themselves, to conclude whether self-ratings provide accurate information about employee innovative performance (Moneta et al., 2010). One way to approach this issue is the use of multisource or 360-degree framework in which individual innovative performance is assessed by multiple raters. To the best of our knowledge, the study by Moneta et al. (2010) is the only study so far that assessed innovative performance using multiple raters simultaneously. However, unlike in the present study in which we address innovative performance at the individual level, Moneta et al. (2010) explored creative contributions (thus, only the first step in the innovation process) in a team context in which each team member was both a target and an assessor of all other team members.

Multisource or 360-degree feedback in competency appraisals

Multisource or 360-degree feedback refers to the "*evaluations gathered about a target participant from two or more rating sources, including self, supervisor, peers, direct reports, internal customers, external customers, and vendors or suppliers*" (Dalessio, 1998, p. 278). It has been argued that 360-degree feedback from different sources is an important mechanism which provides employees with information about how they are perceived by others from their working environment (Eckert, Ekelund, Gentry, & Dawson, 2010). Thus, "*even if others' ratings are not objective or accurate, it is important for individuals to understand how they are perceived by others in order to navigate the political realities in organizations*" (Ostroff, Atwater, & Feinberg, 2004, p. 333). The feedback obtained from a 360-degree appraisal is then used to develop individual competencies where needed, thus contributing to employee personal and professional development. This is important given that previous research showed a significant association between personal qualities or competences and different aspects of organizational effectiveness (Salgado, 1997; Barrick, Mount, & Judge 2001; Arthur, Bennet, Edens, & Bell, 2003).

A competence has been defined as a set of knowledge, skills, attitudes, values and personal qualities that individuals should posses to successfully perform a group of related tasks (Hensel, Meijers, Van der Leeden, & Kessels, 2010). In line with the definition of innovation outlined previously (Anderson et al., submitted; West & Farr, 1990), innovation competence can be defined as a cluster of skills and knowledge that are required to generate innovative work behavior. Innovative performance can then be defined in terms of the outcomes of innovative behaviors and competence at the individual level, often as perceived by important others such as the job holders immediate supervisor. As other competences, also innovation competence can be trained and developed based on the multisource feedback. It has been argued that feedback improves performance in many areas (Hensel et al., 2010), and thus, we could expect that by providing feedback about innovation competence, the innovative performance could be enhanced. Further, because multisource data integrates perspectives from multiple sources on the competence of an individual in the workplace, this type of appraisal evaluation may be particularly useful to examine innovative behavior at work. Even so, our comprehensive review of the literatures across both workplace innovation and multisource appraisal processes revealed no published studies that had focused specifically upon innovation as a crucially important competence.

Previous research has also shown the discrepancy between self- and observer ratings on the same dimensions of 360-degree feedback tools (Eckert et al., 2010; Morgeson, Mumford, & Campion, 2005). Such discrepancy or incongruence was argued to have a negative effect on the target individual because it might imply low self-awareness, lack or deficiency in skills or inaccurate self-perceptions (Eckert et al., 2010). On one hand, self-ratings on certain dimensions can be higher than the ratings of the observers and this type of discrepancy is called overrating. On the other hand, self-ratings can be lower compared to the ratings of the observers and in this case, we talk about underrating (Atwater & Yammarino, 1997). There is also a situation of agreement when target individuals rate themselves similarly as the observers, either high or low. It is important to identify the cases of discrepancy as previous research found that higher self-observer discrepancies are related with lower performance (Ostroff et al., 2004) and lower motivation to improve future behavior and less actual improvement (Atwater & Brett, 2005).

Study aims and objectives

To summarize, in the present study our overall aim is to address the congruency of self and observer ratings of innovation-related competence within a 360-degree framework. Specifically, first, we aim to explore the construct validity of innovation-related competence by examining the psychometric properties of the applied measure. Second, we aim to examine the congruency between self and observer ratings of innovation-related competence. Third, we aim to evaluate the issue of overrating and underrating further on by specifically exploring the ratings of different observers in a lower and upper quartile groups based on the self-ratings of innovation-related competence, respectively.

Method

Sample and procedure

Data for this study were provided by an international consultancy and were obtained via a multi-rater assessment tool – i.e. via 360 degree appraisal ratings. This is an online instrument aimed at providing organizations with a structured feedback regarding the strengths and development needs of their employees. The target sample comprised 296 target individuals (69.26% male). In terms of management level, 30.40% of target individuals were 1st level managers, 37.80% were 2nd level managers, 4.40% of target individuals did not have any managerial responsibilities and 27.40% did not indicate their management level. Regarding the functional area, the majority were in finance (14%), followed by marketing (13%), management and operation (11%), HR (7%), and IT (5%).

Target individuals provided self-ratings and selected observers to provide ratings on the same tool. Target individuals were evaluated by boss, peers, direct reports and others. Overall, 318 bosses, 1208 peers, 828 direct reports and 329 others assessed the individuals' competences. The overall total sample for this study was therefore 2979 individuals.

Measure

Data from the online multi-rater assessment instrument were examined. The 360-degree appraisal tool measured twelve competences. For the confidentiality reasons, we label the competencies as competencies A, B, C, D, E, F, G, H, I, J, K, and L (this one was a competence related to innovation). Each competence was measured with 5 items, using a 5-point response scale (1 – least effective; 5 – most effective). For the present paper, the competence L was the main focus of our analytical efforts.

Pre-analysis checks and analytical approach

We first carried out a series of reliability and exploratory factor analyses with principal components method to examine the internal consistency and the factor structure of the 360-degree measure. Afterwards, we carried out confirmatory factor analysis to examine alternative factor structures in order to retain the best fitting model (based on chi-square difference between models). Following the recommendations by Hu and Bentler (1998), we use χ^2, RMSEA (Root Mean Square Error of Approximation), CFI (Comparative Fit Index), NNFI (Non-Normed Fit Index) and SRMR (Standardized Root Mean Residual) to evaluate the goodness of fit of the examined models. The ratio between χ^2 and *df* is considered as an appropriate indicator of goodness-of-fit because χ^2 is sample size dependent. Values of this normed χ^2 that fall below 3.0 are indicative of acceptable fit. The cut-off values close to.06 for RMSEA, .08 for SRMR, and .95 for CFI and NFI respectively are considered to indicate good fit (Hu & Bentler, 1999).

In the second part of this pre-analysis, we examined the agreement between different raters specifically in innovation-related competence (competence L). To this end, we calculated the average deviation index ($AD_{M(J)}$, Burke, Finkelstein, & Dusig, 1999), intra-class correlations (ICC(1), ICC(2), James, 1982; Bliese, 2000) and the interrrater agreement index ($r_{WG(J)}$, James, Demaree, & Wolf, 1984; Lindell, Brandt, & Whitney, 1999). Next, we examined the differences in innovation-related competence (L) between different raters by means of ANOVA (these are presented in full below).

To assess the issue of overrating and underrating, we compared the mean values of self-ratings and the ratings of all observers merged together by means of *t*-test. In addition, we computed standardized effect sizes (*d* values) between the means of both groups (i.e. target individuals vs. independent observers) which indicate the standardized mean difference between the two groups being compared (Cohen, 1998; see also Anderson & Ones, 2003 and Ones & Anderson, 2002). Specifically, an effect size is calculated by subtracting means of both groups divided by the pooled standard deviation. It was important to calculate *d* values in this study because the sample sizes of the groups being compared were different.

Standardized effect sizes are appropriate in this case because they show the magnitude of the difference between two groups regardless of the sample size (Cohen, 1998). According to Cohen (1998), effect sizes of .20 are considered small, around .50 are medium and the values above .80 are considered to be large. We also explored the correlations between different observers' ratings of innovation-related competence (L). Finally, we computed an upper and lower quartile split of the sample according to the self-ratings of innovation-related competence (competence L) in order to examine for possible differences between high-performing and low-performing individuals with regard specifically to the competence L. We then examined the relationships between self and observer ratings in each group separately following the same analytical procedure as the one just explained for the overall sample.

Results

Validity and reliability of the 360-degree measure

To assess whether different competencies of 360-degree measure were distinct from each other, we first carried out exploratory factor analysis on the sample of target individuals and each group of observers, separately. In line with previous research on 360-degree feedback, the analyses on the observers were done with scores aggregated to the target-individual level (Beehr, Ivanitskaya, Hansen, Erofeev, & Gudanowski, 2001). The results showed rather unambiguous support for the 12-factor structure of the measure, although some items cross-loaded on different factors. The reliability analysis showed acceptable values for each competence (see Table 1).

PLEASE INSERT TABLE 1 ABOUT HERE

A series of CFAs were conducted next to compare alternative factor structures of the applied instrument (e.g., first order one-factor model, first order 12-factor model, and a series of other alternative first and second order models with different number of factors). Our results showed that a second-order factor model in which each of the 12 competences represents a distinct factor, but at the same time they are dependent on a single underlying factor exhibited acceptable fit in all sources (see Table 2). The fit of this model was significantly better compared to the goodness-o-fit of other models, such as first order single factor model or as good as the fit of other higher-order models with more underlying second order factors. Thus, based on the goodness-of-fit comparison tests with alternative models and considering parsimony principle, we accepted the single higher order factor model with twelve first order factors as the most preferred model. We can conclude that the twelve competences represent conceptually distinct concepts, one of them being innovation-related competence (L) and relatively high correlations between them (as shown in Table 3 for the sample on target individuals[1]) imply that they are dependent on one overall performance factor.

PLEASE INSERT TABLE 2 ABOUT HERE

PLEASE INSERT TABLE 3 ABOUT HERE

Innovation-related competence: Differences between self and others' ratings

To examine the discrepancy between self and observer ratings in innovation-related competence, we first aggregated the ratings of direct reports, peers, and others to the target-

[1] Similar patterns of correlations were observed on the samples of independent observers.

individual level. As outlined previously, we calculated different indexes to justify the aggregation across observers in each group. As can be seen in Table 4, the values of these indexes showed acceptable agreement between raters in each group. Apart from the ICC(2) values which were slightly lower compared to those found in previous research in the field of 360-degree framework (e.g. Ostroff et al., 2004), the $r_{wg(j)}$ values are comparable to those found in other studies with multiple raters (Eckert et al., 2010).

PLEASE INSERT TABLE 4 ABOUT HERE

Table 4 also shows the mean values on innovation-related competence for self and observer ratings. The ANOVA results showed significant differences in this competence between different raters ($F(4, 2974) = 33.472$; $p < .01$). Specifically, we found that direct reports rated individual innovation-related competence significantly higher compared to peers, target individuals themselves, others and boss. Also peers rated innovation-related competence higher than the boss. Moreover, self-ratings were significantly lower compared to the ratings of all observers together ($t(398.75) = 2.48$; $p < .05$; $d = .13$). These results indicate that individuals in general underrated their innovation-related competence, although we should bear in mind that the effect size is small.

The correlations between self-ratings and the ratings of different observers are shown in Table 5. We can observe that self-ratings are significantly correlated only with peer ratings, whereas the boss ratings correlate significantly with the ratings of others, peers and direct reports. We also note a significant correlation between other and peer ratings and between direct reports and peer ratings.

PLEASE INSERT TABLE 5 ABOUT HERE

To explore these results further in more detail, we performed the upper-lower quartile split of the sample in self-rated innovation-related competence. In Table 6 we show the descriptive and aggregation statistics for each quartile group.

PLEASE INSERT TABLE 6 ABOUT HERE

Compared to the results for the overall sample, the ICC(1) and ICC(2) values in the group of the lowest self-rated individuals were lower, indicating less agreement across observers. However, the $r_{WG(J)}$ and $ADI_{M(J)}$ showed acceptable values. Furthermore, the ANOVA results showed significant differences between different sources of innovation-related competence ratings ($F(4, 802) = 30.633$; $p < .01$). We found that direct reports rated target individuals' innovation-related competence significantly higher compared to peers, others, boss, and target individuals. Furthermore, we found that target individuals self-rated their innovation-related competence lower then bosses, others, peers, and direct reports. In the same line, the combined ratings of all observers were significantly higher than self-ratings ($t(165.31) = 15.31$; $p < .01$; $d = 1.01$). These results indicate that observers rated innovation-related competence of the worse self-rated individuals higher compared to the target individuals themselves.

We present correlations between ratings from different sources in the lower quartile group in Table 7 (above the diagonal). As can be seen, in this group boss ratings were correlated with self-ratings and the ratings from others and peers.

PLEASE INSERT TABLE 7 ABOUT HERE

Finally, we conducted the same analyses on the upper quartile of self-ratings of innovation-related competence. Also in this group the level of agreement across different observers was acceptable, with the only exception being the ratings of others that exhibited very low ICC(1) and ICC(2) values (see Table 6).

As on the whole sample and in a lower quartile group, the ANOVA results showed significant differences in innovation-related competence ratings between different sources ($F(4, 1102) = 24.748$; $p < .01$). We found that target individuals in the upper quartile self-rated their innovation-related competence significantly higher compared to peers, others, and boss. We also found that direct reports rated innovation-related competence significantly higher than peers, others and boss. Finally, target individuals rated their innovation-related competence significantly higher compared to the rest of the observers as a whole ($t(312.16) = -11.71$; $p <.01$; $d = -.59$). Thus, individuals in the upper quartile overrated their innovation-related competence in relation to the rest of the observers. We also found that boss ratings correlated with the ratings of others, peers, and direct reports (see Table 7, below the diagonal). Also the correlation between the ratings of peers and direct reports was significant.

Discussion

The main aims of the present study were to examine the construct validity of innovation competence and to explore its measurement within a multisource framework. To the best of our knowledge, this study is the first to address innovative performance evaluations within a 360-degree appraisal framework. Our findings showed that innovation is conceptually and psychometrically different from other types of competencies. Moreover, interesting results were observed when examining the congruencies between self and observer ratings of innovation-related competence. These findings suggest important implications for research in the field of innovation at the workplace which to date has largely relied exclusively on self-ratings of individual innovative performance.

Implications for future research

In accord with the employee performance model proposed by Welbourne et al. (1998) and Bartram's (2005) Great Eight competency framework, our findings show that innovation competence is conceptually different from a wide range of other competencies, even when measured within the same 360-degree appraisal procedure. Thus, the performance management literature should address innovation as aspect of employee performance within their models as a distinct dimension, especially given that this type of performance is pursued by many organizations in their attempts to improve their effectiveness (Anderson et al., 2004; Anderson et al., submitted).

As for the measurement of innovative performance, our findings challenge the vast majority of previous research that assessed this type of performance with self-reports. The present study adds additional evidence to previous research that showed self-ratings to be "*biased, inaccurate, and generally suspect when compared to the ratings of others*" (Yammarino & Atwater, 1993, p. 231). We found that target individuals underrated their innovation-related competence compared to the global ratings of other observers. The largest difference between self and observer ratings were found in case of direct reports which could imply the tendency that subordinates were afraid of assessing their boss in a harsh or strict way. Our findings are in line with previous research that showed how self-ratings frequently differ from the ratings of observers, such as supervisors, peers or subordinates regardless of the rated traits or behaviors (Ostroff et al., 2004). Such incongruence between self and observer ratings is an important issue as it has been argued that disagreement between the self and observer ratings is associated with poorer performance (Ostroff et al., 2004). These findings also add more weight to earlier meta-analytical findings in innovation that have suggested that self-report designs may inflate cause-effect relations partly by measuring the dependent variable of innovation either unreliably or inappropriately (Hulsheger et al., 2009). Indeed, our findings suggest that individuals' self-evaluation of their own innovation-related competence likely differs from other work colleagues evaluations, and more pointedly, from that of their immediate supervisor or boss.

Our correlational analysis of ratings from different sources showed that self-ratings were only correlated with the ratings of peers, whereas the ratings of observers were significantly correlated. Interestingly, previous research in the field of 360-degree feedback reports significant correlations between self and independent ratings of other competencies (Beehr et al., 2001; Ostroff et al., 2004). Moneta et al. (2010) also report significant correlations between self and supervisory ratings and self and peer ratings of creativity, respectively. One possible explanation for this is that peers and self-ratings covaried because peers spend more time with target individuals compared to the rest of the observers and they start to view target individuals similarly as target individuals view themselves. Moreover, the lack of correlation between self and boss ratings could be due to bosses not having enough opportunity to observe their employees and hence, provide unreliable ratings (Rothstein, 1990)[2]. However, in the present study we did not have information regarding the length of observation period to address this possibility directly. We also have to bear in mind that observer ratings were aggregated to the target individual-level and thus might represent higher intercorrelations because of higher reliability of the aggregated ratings (Scullen, 1997).

A different pattern of results was observed in the lower and upper quartile groups based on self-ratings. In the lower quartile group, individuals were found to underrate their

[2]We appreciate an action editor's suggestion regarding this potential explanation for our results.

innovation-related competence compared to all other observers. Interestingly, however, we found that although bosses rated this competence of lower quartile individuals significantly higher than individuals themselves, self-ratings in this group were significantly correlated with boss ratings. In contrast, target individuals in the upper quartile were found to overrate their innovation-related competence. According to the existing literature in the field of 360-degree feedback, such overrating could result from an inflated self-view or a self-enhancement bias and has the most damaging impact on performance (Eckert et al., 2010; Ostroff et al., 2004). Moreover, this argument can be supported with the lack of significant correlations between self and observers' ratings. Interestingly, boss ratings of the target individuals' innovation-related competence in the upper quartile were correlated with all the rest of the observer ratings. This finding might imply that observers were not harsh in rating individuals with respect to their innovation-related competence but rather that individuals lack self-awareness and have an inappropriate view of their innovating ability.

A lack of correlations between self and observer ratings could have important implications for future research regarding the measurement of innovation in the workplace. Our findings add additional evidence to previous research that reports problems with using self-reports for measuring innovation (Hülsheger et al., 2009). The fact that self-ratings do not correlate with independent ratings implies response bias and suggests that future studies in innovation should avoid this type of innovation measurement. Nevertheless, we would like to highlight that some might argue the opposite, namely that self-ratings could be more accurate than the observer ratings. For instance, it could be that innovative performance is not easily observed and that only individuals themselves can provide the true ratings of this type of performance. The existing research about true and accurate scores, however, does not support this argument. Becker and Miller (2002) suggest that measurement accuracy can be determined by comparing the (self-) ratings against a true score which is computed by averaging the ratings of multiple raters who possess expert knowledge about the skills being evaluated and who provide their ratings in optimal conditions (see also Smither, Barry, & Reilly, 1989). The less discrepancy between the rating (provided by the individual) and a true score (provided by independent raters), the more the rating is accurate.

In the present study we cannot firmly conclude that the observers were expert raters, however their ratings could be considered as approaching true scores as they knew what behaviors were required for innovation-related competence and they frequently interacted with target individuals and thus could observe their innovative performance on multiple occasions. Even though observer ratings cannot be claimed to be perfectly accurate, true or reliable measures, we did find that independent reports were significantly correlated, with the only exception being “other-direct reports” correlation (which was in fact marginally significant). These significant correlations imply a more reliable and valid measurement of innovation-related competence by independent observers. Finally, we would like to suggest that not only it is important to use independent ratings to more accurately assess innovation, but also to avoid common method bias and overestimation of effects in correlational designs (Hülsheger et al., 2009; Podsakoff et al., 2003).

Implications for practice

Our findings also have a number of important and novel implications for practice. First, the finding that individuals’ self-ratings and those of others including their immediate supervisors differ suggests that organizations need to apply the outcomes of 360-degree appraisal with some caution. As such multisource evaluation schemes have become more popular in organizations as a means to performance appraisal of employees (e.g., Eckert et al., 2010; Morgeson et al., 2005), patterns of either under-evaluation, or alternatively, over-evaluation by the target appraisee compared to the ratings of others need to be considered with some care. The feedback given to individuals going through such appraisal processes will need careful planning, but perhaps more so in the case where individuals have typically over-evaluated their own competence compared to the ratings of others, especially their boss (Eckert et al., 2010; Ostroff et al., 2004). In the case of our findings, the consistent finding that target individuals under-evaluated themselves on innovation-related competence compared to the ratings of others suggests that self-insight into one's own innovation abilities in the workplace may err on the side of false modesty. Second, our findings unambiguously suggest that organizations should design intervention techniques in multisource appraisal by paying attention to individuals who have substantially incongruent perceptions of their innovation-related competence. In this case multisource feedback could be used to help these individuals develop a more accurate and balanced perception of their innovation-related competence. Moreover, those who are perceived as having low competence to innovate could be offered specific and tailored training to enhance their skills and knowledge and develop this type of competence (Morgeson et al., 2005). A fourth and final practical implication of our results regards the management of innovation in the workplace, and in particular the way in which innovation is measured as an outcome. Again, our findings indicate clearly that organizations should not merely rely upon the self-ratings of individuals but are well advised to incorporate multiple ratings from different sources in order to gain an accurate picture of individuals' innovation competence. Whether such ratings are gleaned for appraisal or for other purposes, the implication of our findings is that basing this upon a multiple sources and perspectives is preferable to relying just upon the individuals' own self-assessment.

Strengths and limitations of the present study

Several strengths to the design of the present study can be noted. This research is the first to the knowledge of the authors to investigate innovation competence within an operational 360-degree performance evaluation framework. Although some could argue that the sample size of target individuals was limited to provide sound evidence about the validity of the 360-degree measure, the overall sample for our study is large – almost 3000 in total centered around the ratings of just under 300 mid- to senior managers in different occupations across different countries. In all of these job functions, arguably, innovation competence was an important determinant of overall job performance. A further strength is thus that we had an overall ratio of around 10:1 of ratings of others versus the target individuals' self-rating on the competency framework. A final strength is that our analyses and pre-checks of the psychometric properties of the 12-dimension framework confirmed this factor structure, with all factors displaying reliability and patterns of inter-rater consistency.

Some limitations of the present study also need to be acknowledged. First, these findings of course relate to a single competency framework to which the authors were kindly granted research access. Whether these findings generalize to other such frameworks is an open question and one that will require further research to establish. However, our study at least raises the prospect that such differences do exist, and we would also argue that the competency framework used for this study resembles others used more generally for managerial performance evaluation. Second, the ICC values were slightly lower compared to the ones found in previous research. Nevertheless, other indexes supported the aggregation of observer ratings to the target-individual level. Third, we did not have any independent criteria variable available to examine the predictive validity of innovation-related competence, although the availability of multisource data does ameliorate this concern somewhat.

Conclusion

To conclude, our findings contribute both to the literatures and research on innovation but also on multisource performance appraisal in the workplace. A notable shortcoming in our understanding has arguably been that there has been a dearth of published studies that have examined innovation competence, and have thus attempted to integrate across these two largely disparate areas of organizational psychology and HRM. Our findings have important implications for both areas, and for both future research and practice of performance management particularly with regard to 360-degree appraisal procedures. It is our hope that these initial findings provoke and stimulate more interest in these areas, most notably at the interface of performance appraisal and innovation which represents a critical interface for both areas of research and practice.

References

Alge, B. J., Ballinger, G. A., Tangirala, S., & Oakley, J. L. (2006). Information privacy in organizations: Empowering creative and extra role performance. *Journal of Applied Psychology, 91*, 221-232. doi:10.1037/0021-9010.91.1.221

Amabile, T. M., Mueller, J. S., Simpson, W. B., Hadley, C. N., Kramer, S. J., & Fleming, L. (2002). *Time pressure and creativity in organizations: A longitudinal field study.* (HBS Working Paper 02–073). Cambridge, MA: Harvard University, Harvard Business School.

Anderson, N., & Ones, D. S. (2003). The construct validity of three entry-level personality inventories: Cautionary findings from a multiple inventory investigation. *European Journal of Personality, 17*, 1-28. doi: 10.1002/per.484

Anderson, N., De Dreu, C. K. W., & Nijstad, B. A. (2004). The routinization of innovation research: A constructively critical review of the state-of-the-science. *Journal of Organizational Behavior, 25*, 147-173. doi:10.1002/job.236

Anderson, N., Potočnik, K. & Zhou, J. (submitted). *Innovation and Creativity in Organizations: A State-of-the-Science Review and Prospective Commentary*. Unpublished Manuscript

Arthur, W. J., Bennett, W. J., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology, 88*, 234-245. doi:10.1037/0021-9010.88.2.234

Atwater, L. E., & Brett, J. F. (2005). Antecedents and consequences of reactions to developmental 360° feedback. *Journal of Vocational Behavior, 66*, 532-548. doi:10.1016/j.jvb.2004.05.003

Atwater, L. E., & Yammarino, F. J. (1997). Self-other rating agreement: A review and model. In G. R. Ferris, & G. R. Ferris (Eds.), *Research in personnel and human resources management, vol. 15.* (pp. 121-174). US: Elsevier Science/JAI Press.

Axtell, C. M., Holman, D. J., Unsworth, K. L., Wall, T. D., Waterson, P. E., & Harrington, E. (2000). Shopfloor innovation: Facilitating the suggestion and implementation of ideas. *Journal of Occupational and Organizational Psychology, 73*, 265-285. doi:10.1348/096317900167029

Axtell, C., Holman, D., & Wall, T. (2006). Promoting innovation: A change study. *Journal of Occupational and Organizational Psychology, 79*, 509-516. doi:10.1348/096317905X68240

Baer, M. (2010). The strength-of-weak-ties perspective on creativity: A comprehensive examination and extension. *Journal of Applied Psychology, 95*, 592-601. doi:10.1037/a0018761

Baer, M., & Frese, M. (2003). Innovation is not enough: Climates for initiative and psychological safety, process innovations, and firm performance. *Journal of Organizational Behavior, 24*, 45-68. doi:10.1002/job.179

Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*, 9-30. doi:10.1111/1468-2389.00160

Bartram, D. (2005). The great eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology, 9.* 1185-1203. doi:10.1037/0021-9010.90.6.1185

Becker, G. A., & Miller, C. E. (2002). Examining contrast effects in performance appraisals: Using appropriate controls and assessing accuracy. *Journal of Psychology, 136*, 667-683. doi:10.1080/00223980209604827

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. doi:10.1002/job.113

Berson, Y., Oreg, S., & Dvir, T. (2008). CEO values, organizational culture and firm outcomes. *Journal of Organizational Behavior, 29*, 615-633. doi:10.1002/job.499

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analyses. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). San Francisco: Jossey-Bass.

Burke, M. J., Finkelstein, L. M., & Dusig, M. S. (1999). On average deviation indices for estimating interrater agreement. *Organizational Research Methods, 2*, 49-68. doi:10.1177/109442819921004

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35–70). San Francisco: Jossey-Bass.

Cardinal, L. B. (2001). Technological innovation in the pharmaceutical industry: The use of organizational control in managing research and development. *Organization Science, 12,* 19–36. doi:10.1287/orsc.12.1.19.10119

Chi, N., Huang, Y., & Lin, S. (2009). A double-edged sword? Exploring the curvilinear relationship between organizational tenure diversity and team innovation: The moderating role of team-oriented HR practices. *Group & Organization Management, 34*, 698-726. doi:10.1177/1059601109350985

Clegg, C., Unsworth, K., Epitropaki, O., & Parker, G. (2002). Implicating trust in the innovation process. *Journal of Occupational & Organizational Psychology, 75*, 409-422. doi:10.1348/096317902321119574

Cohen, J. (1998). *Standard power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Elburn.

Dalessio, A. (1998). Using multisource feedback for employee development and personnel decisions. In J. W. Smither (Ed.), *Performance Appraisal: State of the art in practice* (pp. 278–330). San Francisco, CA: Jossey-Bass.

Eckert, R., Ekelund, B. Z., Gentry, W. A., & Dawson, J. F. (2010). 'I don't see me like you see me, but is that a problem?' cultural influences on rating discrepancy in 360-degree feedback instruments. *European Journal of Work and Organizational Psychology, 19*, 259-278. doi:10.1080/13594320802678414

Frese, M., Teng, E., & Wijnen, C. J. D. (1999). Helping to improve suggestion systems: Predictors of making suggestions in companies. *Journal of Organizational Behavior, 2.* 1139-1155. doi:10.1002/(SICI)1099-1379(199912)20:7<1139::AID-JOB946>3.0.CO;2-I

Frese, M., & Zapf, D. (1994). Action as the core of work psychology: A German approach. In H. C. Triandis, M. D. Dunette, & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 4, pp. 271–340). Palo Alto, CA: Consulting Psychologists Press.

George, J. M., & Zhou, J. (2001). When openness to experience and conscientiousness are related to creative behavior: an interactional approach. *Journal of Applied Psychology, 86*, 513–524. doi:10.1037/0021-9010.86.3.513

Gilson, L. L., & Shalley, C. E. (2004). A little creativity goes a long way: An examination of teams' engagement in creative processes. *Journal of Management, 3.* 453-470. doi:10.1016/j.jm.2003.07.001

Hensel, R., Meijers, F., Van der Leeden, R., & Kessels, J. (2010). 360 degree feedback: How many raters are needed for reliable ratings on the capacity to develop competences, with personal qualities as developmental goals? *The International Journal of Human Resource Management, 21*, 2813-2830. doi:10.1080/09585192.2010.528664

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterization model misspecification. *Psychological Methods, 3,* 424-453. doi:10.1037/1082-989X.3.4.424

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55. doi:10.1080/10705519909540118

Hunt, S. T. (1996). Generic work behavior: An investigation into the dimensions of entry-level, hourly job performance. *Personnel Psychology, 49*, 51-83.

Hülsheger, U. R., Anderson, N., & Salgado, J. F. (2009). Team-level predictors of innovation at work: A comprehensive meta-analysis spanning three decades of research. *Journal of Applied Psychology, 94*, 1128-1145. doi:10.1037/a0015978

James, L. (1982). Aggregation bias in estimates of perceptual measures. *Journal of Applied Psychology, 67,* 219-229. doi:10.1037/0021-9010.67.2.219

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69*, 85-98. doi:10.1037/0021-9010.69.1.85

Latham, S. F., & Braun, M. (2009). Managerial risk, innovation, and organizational decline. *Journal of Management, 35*, 258-281. doi:10.1177/0149206308321549

Lindell, M. K., Brandt, C. J., & Whitney, D. J. (1999). A revised index of interrater agreement for multi-item ratings of a single target. *Applied Psychological Measurement, 23*, 127-135. doi:10.1177/01466219922031257

Miron, E., Erez, M., & Naveh, E. (2004). Do personal characteristics and cultural values that promote innovation, quality, and efficiency compete or complement each other? *Journal of Organizational Behavior, 25*, 175-199. doi:10.1002/job.237

Moneta, G. B., Amabile, T. M., Schatzel, E. A., & Kramer, S. J. (2010). Multirater assessment of creative contributions to team projects in organizations. *European Journal of Work and Organizational Psychology, 19*, 150-176. doi:10.1080/13594320902815312

Morgeson, F. P., Mumford, T. V., & Campion, M. A. (2005). Coming full circle: Using research and practice to address 27 questions about 360-degree feedback programs. *Consulting Psychology Journal: Practice and Research, 57*, 196-209. doi:10.1037/1065-9293.57.3.196

Ohly, S., Sonnentag, S., & Pluntke, F. (2006). Routinization, work characteristics and their relationships with creative and proactive behaviors. *Journal of Organizational Behavior, 27*, 257-279. doi:10.1002/job.376

Oldham, G. R., & Cummings, A. (1996). Employee creativity: Personal and contextual factors at work. *Academy of Management Journal, 39,* 607–634. doi:10.2307/256657

Ones, D. S., & Anderson, N. (2002). Gender and ethnic group differences on personality scales in selection: Some British data. *Journal of Occupational and Organizational Psychology, 75,* 255-276. doi:10.1348/096317902320369703

Ostroff, C., Atwater, L. E., & Feinberg, B. J. (2004). Understanding self-other agreement: A look at rater and ratee characteristics, context, and outcomes. *Personnel Psychology, 57*, 333-375. doi:10.1111/j.1744-6570.2004.tb02494.x

Pearce, C. L., & Ensley, M. D. (2004). A reciprocal and longitudinal investigation of the innovation process: The central role of shared vision in product and process innovation teams (PPITs). *Journal of Organizational Behavior, 25*, 259-278. doi:10.1002/job.235

Pirola-Merlo, A., & Mann, L. (2004). The relationship between individual creativity and team creativity: Aggregating across people and time. *Journal of Organizational Behavior, 25*, 235-257. doi:10.1002/job.240

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88,* 879–903. doi:10.1037/0021-9010.88.5.879

Potočnik, K., & Anderson, N. (April, 2012). Applying a 360-degree Framework to Innovative Performance Measurement. In N. Anderson (Chair), *International Advances in Innovation and Creativity in the Workplace*. Invited symposium conducted at the Annual Conference of Society of Industrial and Organizational Psychology (SIOP), San Diego, USA.

Potočnik, K., & Anderson, N. (July, 2012). Assessment of Innovative Performance. In N. Anderson & K. Potočnik (Chairs), *Advances in Testing and Measurement: Job Performance, Change and Innovation*. Invited symposium conducted at the 8th Conference of the International Test Commission, Amsterdam, The Netherlands.

Rank, J., Pace, V. L., & Frese, M. (2004). Three avenues for future research on creativity, innovation, and initiative. *Applied Psychology: An International Review, 53*, 518-528. doi:10.1111/j.1464-0597.2004.00185.x

Rothstein, H. (1990). Interrater reliability of job performance ratings: Growth to asymptote level with increasing opportunity to observe. *Journal of Applied Psychology, 75,* 322-327. doi:10.1037/0021-9010.75.3.322

Salgado, J. F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology, 82*, 30-43. doi:10.1037/0021-9010.82.1.30

Scullen, S. E. (1997). When ratings from one source have been averaged, but ratings from another source have not: Problems and solutions. *Journal of Applied Psychology, 82,* 880-888. doi: 10.1037/0021-9010.82.6.880

Shalley, C. E. (1995). Effects of coaction, expected evaluation, and goal setting on creativity and productivity. *Academy of Management Journal, 38,* 483–503. doi:10.2307/256689

Shalley, C. E., & Perry-Smith, J. E. (2001). Effects of social–psychological factors on creative performance: The role of informational and controlling expected evaluation and modeling experience. *Organizational Behavior and Human Decision Processes, 84,* 1–22. doi:10.1006/obhd.2000.2918

Smither, J. W., Barry, S. R., & Reilly, R. R. (1989). An investigation of the validity of expert true score estimates in appraisal research. *Journal of Applied Psychology, 74*, 143-151. doi:10.1037/0021-9010.74.1.143

Subramaniam, M., & Youndt, M. A. (2005). The influence of intellectual capital on the types of innovative capabilities. *Academy of Management Journal, 48*, 450-463. doi:10.5465/AMJ.2005.17407911

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

Yuan, F., & Woodman, R. W. (2010). Innovative behavior in the workplace: The role of performance and image outcome expectations. *Academy of Management Journal, 53,* 323-342. doi:10.5465/AMJ.2010.49388995

Welbourne, T. M., Johnson, D. E., & Erez, A. (1998). The role-based performance scale: Validity analysis of a theory-based measure. *Academy of Management Journal, 41*, 540-555. doi:10.2307/256941

West, M. A. & Anderson, N. R. (1996). Innovation in top management teams. *Journal of Applied Psychology, 81,* 680–693. doi:10.1037/0021-9010.81.6.680

West, M. A., & Anderson, N. (1992). Innovation, cultural values and the management of change in British hospitals. *Work and Stress, 6*, 293–310. doi:10.1080/02678379208259959

West, M. A., & Farr, J. L. (1990). Innovation at work. In M. A. West & J. L. Farr (Eds.), *Innovation and creativity at work: Psychological and organizational strategies* (pp. 3–13). Chichester: Wiley.

West, M. A., Smith, H., Feng, W. L., & Lawthom, R. (1998). Research excellence and departmental climate in British universities. *Journal of Occupational and Organizational Psychology, 71*, 261–281.

Zhang, Z., Hempel, P. S., Han, Y., & Tjosvold, D. (2007). Transactive memory system links work team characteristics and performance. *Journal of Applied Psychology, 92*, 1722-1730. doi:10.1037/0021-9010.92.6.1722

Table 1

Descriptives and reliabilities of the measured competencies for each source of data

	Target individuals			Boss			Other			Peer			Report		
	M	*SD*	α	*M*	*SD*	α	*M*	*SD*	α	*M*	*SD*	α	*M*	*SD*	α
Competence A	3.27	.58	.79	3.27	.63	.80	3.37	.59	.81	3.39	.38	.83	3.68	.53	.84
Competence B	3.20	.61	.78	3.32	.65	.80	3.32	.63	.82	3.31	.44	.84	3.55	.51	.84
Competence C	3.31	.54	.68	3.28	.62	.75	3.33	.55	.81	3.33	.39	.79	3.57	.51	.79
Competence D	3.20	.57	.77	3.21	.59	.76	3.30	.53	.77	3.29	.39	.83	3.55	.54	.83
Competence E	3.22	.56	.73	3.04	.58	.70	3.20	.57	.78	3.21	.39	.77	3.47	.52	.77
Competence F	3.32	.51	.61	3.22	.55	.64	3.38	.53	.74	3.32	.39	.75	3.58	.56	.73
Competence G	3.19	.54	.71	3.16	.56	.74	3.22	.58	.80	3.21	.43	.80	3.43	.60	.85
Competence H	3.36	.51	.68	3.29	.60	.79	3.32	.52	.74	3.37	.38	.80	3.59	.57	.84
Competence I	3.32	.53	.68	3.29	.57	.67	3.40	.47	.65	3.40	.37	.74	3.67	.50	.78
Competence J	3.31	.54	.70	3.29	.60	.75	3.38	.47	.77	3.39	.37	.77	3.63	.56	.78
Competence K	3.22	.52	.71	3.32	.60	.72	3.36	.50	.74	3.37	.37	.73	3.61	.55	.78
Competence L*	3.29	.57	.77	3.20	.65	.81	3.22	.76	.83	3.33	.65	.83	3.60	.71	.84

Note. * - competence L in the present framework was a competence related to innovation and the central competence in the present paper

Table 2

Fit indices for the one-factor higher order model with 12 first order factors in each source

Source	χ^2	*df*	$\chi2/df$	RMSEA (90% CI)	SRMR	NNFI	CFI
Target individuals	3661.82	1698	2.16	.06 (.06-.06)	.06	.96	.97
Boss	4477.28	1698	2.64	.07 (.07-.07)	.06	.97	.97
Other	3505.31	1698	2.06	.10 (.09-.10)	.09	.92	.92
Peer	5455.96	1698	3.21	.09 (.08-.09)	.07	.97	.97
Report	4771.65	1698	2.81	.08 (.08-.09)	.07	.97	.97

Notes. The analyses on each group of observers were done with scores aggregated at the target individual level.

Table 3

Correlations between 360-degree competences

		1	2	3	4	5	6	7	8	9	10	11
1.	Competence A											
2.	Competence B	.61										
3.	Competence C	.52	.60									
4.	Competence D	.67	.58	.59								
5.	Competence E	.44	.51	.65	.65							
6.	Competence F	.52	.51	.51	.62	.66						
7.	Competence G	.51	.52	.59	.61	.66	.67					
8.	Competence H	.66	.65	.66	.71	.67	.64	.65				
9.	Competence I	.65	.59	.54	.70	.57	.59	.60	.72			
10.	Competence J	.55	.61	.59	.63	.53	.54	.58	.72	.65		
11.	Competence K	.64	.55	.69	.63	.63	.58	.67	.68	.63	.61	
12.	Competence L*	.66	.58	.56	.68	.53	.56	.54	.65	.60	.59	.65

Note. $N = 296$ target individuals. * - innovation-related competence. All correlation coefficients were significant at $p < .01$.

Table 4

Means, standard deviations and aggregation statistics of innovation-related competence (L) ratings from different sources

Source	*M*	*SD*	Range	$ADI_{M(J)}$		$r_{WG(j)}$		ICC(1)	ICC(2)
				M	*SD*	*M*	*SD*		
Target individual	3.29	.57	1-5						
Boss	3.20	.62	1-5						
Peers	3.31	.40	2.01-4.16	.56	.27	.88	.16	.16	.43
Direct reports	3.56	.52	1.70-4.87	.46	.27	.87	.13	.22	.48
Others	3.25	.60	0-4.60	.48	.36	.83	.26	.14	.33
Overall observers	3.38	.34	2.15-4.22	.66	.20	.86	.10	.14	.60

Note. 5-point response scale, ranging from 1 to 5. All *Fs* for ICC(1) were significant.

Table 5

Correlations between innovation-related competence (L) ratings from different sources

	1	2	3	4
1. Competence L-target individual				
2. Competence L -boss	.10			
3. Competence L -others	.03	.20*		
4. Competence L -peers	.20**	.30**	.26**	
5. Competence L -reports	.10	.14*	.16	.31**

Note. The correlations were computed at the target-individual level (*N*=296).

*p < .05; **p < .01

Table 6

Means, standard deviations and aggregation statistics of innovation-related competence (L) ratings from different sources the lower and upper quartile split

Source	*M*	*SD*	Range	$ADI_{M(J)}$		$r_{WG(J)}$		ICC(1)	ICC(2)
				M	*SD*	*M*	*SD*		
Lower quartile split									
Target individual	2.55	.32	1-2.8						
Boss	3.13	.67	1-5						
Peers	3.20	.72	0-5	.65	.31	.85	.19	.11	.33
Direct reports	3.53	.67	1.4-5	.43	.27	.86	.13	.05	.15
Others	3.16	.73	0-4.7	.52	.32	.81	.30	.10	.27
Overall observers	3.38	.70	0-5	.71	.20	.85	.11	.08	.46
Upper quartile split									
Target individual	3.85	.27	3.6-5						
Boss	3.26	.58	1.3-4.4						
Peers	3.39	.63	0-5	.53	.27	.89	.15	.18	.46
Direct reports	3.67	.73	0-5	.48	.26	.86	.14	.34	.63
Others	3.29	.66	1.5-5	.47	.33	.86	.17	.02	.06
Overall observers	3.46	.68	0-5	.63	.18	.87	.09	.19	.67

Note. 5-point response scale. In the lower quartile split the *Fs* for ICC(1) were significant in case of peers and overall observer rating. In the upper quartile split the *F* for ICC(1) was not significant in case of others.

Table 7

Correlations between innovation-related competence (L) ratings from different sources for the lower and upper quartile split

	1	2	3	4	5
1. Competence L-target individual	-	.23*	-.11	.16	.15
2. Competence L -boss	-.07	-	.34*	.29*	.06
3. Competence L -others	-.19	.34*	-	.01	-.17
4. Competence L -peers	.10	.24*	.13	-	.14
5. Competence L -reports	-.08	.32**	-.16	.43**	-

Note. The correlations were computed at the target-individual level (*N*=113 in the upper quartile group and *N*=76 in the lower quartile group). The correlations for the lower quartile group are presented above the diagonal.

*p < .05; **p < .01