

---

# Relationships Among Developmental Competency Measures and Objective Work Outcomes in a New Zealand Retail Context

Duncan J. R. Jackson, Helena D. Cooper-Thomas,  
Marco van Gelderen, Jane Davis

*Competencies represent an important and popular topic in human resource development. Despite this popularity, a divide exists between practitioner approaches to developmental competency measures and the empirical scrutiny of such approaches. However, the scarce empirical studies on competency measures have begun to bridge this gap. In the present study, behavioral competency ratings and objective outcome measures were collected from 118 entry-level employees in a retail organization in New Zealand. A correlational design was applied to data in this study and, with the use of canonical correlation analyses, meaningful relationships were observed among competency measures and objective work outcomes. Such relationships are presented as being practically useful when making decisions about weighting certain competencies over others for developmental purposes.*

Competencies represent an important topic in human resource development (HRD) and have been applied in a variety of developmental manifestations, both internationally (Meriot, 2005) and in New Zealand (Markus, Cooper-Thomas, & Allpress, 2005; Ruth, 2006). Applications include the development of job performance (Bartram, Robertson, & Callinan, 2002; Spencer & Spencer, 1993) and more general job-related behaviors (Woodruffe, 1991, 1993), performance assessment (Bartram, 2005; Catano, Darr, & Campbell,

2007; Xiao, 2006), emotional and interpersonal development (Riggio & Lee, 2007), and job analysis (Lievens, Sanchez, & de Corte, 2004; Wilson, Levine, Morgeson, Thomsen, & Harvey, 2007). Given their broad application, it appears that competencies are nigh impossible to define as a unitary concept (Meriot, 2005; Schippmann et al., 2000). In an effort to capture the inherently multidimensional nature of this approach, Hoffmann (1999) proposed a tri-modal definition by suggesting that competencies can manifest as (a) observable performance, (b) standards or quality assessments of outcome performance, and/or (c) underlying attributes or traits (in the Boyatzis, 1982 tradition) that cause performance. The present article focuses on the first of these; that is, competencies as observable performance (Hoffmann, 1999). The competency movement has been largely driven by consultancies (Athey & Orth, 1999), and independent empirical research into their use as developmental measures is scarce (Barrett & Depinet, 1991). This research aims to establish the validity and utility of behavioral competencies for HRD. It does so by relating developmental competency measures to objective work outcomes.

## Competencies and Human Resource Development

The published literature tends to suggest that, under certain circumstances, competencies hold promise for providing guidance in HRD practices. For example, Bartram (2005) reports on a set of competency measures founded on the Great Eight model (see Kurz & Bartram, 2002). The Great Eight comprises a set of eight generic behavioral clusters that describe key areas of work-related performance (e.g., leading and deciding, supporting and cooperating). In particular, Bartram notes that competency approaches hold the capacity to differentiate subcomponents of job performance. This feature holds theoretical advantages over undifferentiated measures of overall job performance (OJP; see Bartram, 2005, p. 1185), which are commonly observed in the literature (Austin & Crespín, 2006; Murphy, 2008). That said, Bartram found a sizable average meta-analytic correlation (weighted  $r = 0.75$ ; unweighted  $r = 0.79$ ) between Great Eight competency measures and OJP. Questions arise, therefore, as to whether competencies are redundant with OJP and, thus, whether they add benefit for organizations investing in them. Conversely, it is possible that this sizable correlation was due, in part, to common-method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003) in that the competency and OJP measures in Bartram's study both took the form of judgmental ratings based on scale responses.

Notwithstanding these debates, the ability to separate different aspects of performance carries key advantages. First, differentiating performance characteristics provides the capacity to specify predictor-criterion relationships for a wider range of criteria. Thus, organizations can more specifically assess which human resource measures adequately predict performance criteria. Second, the differentiation of various components of performance carries clear

developmental opportunities, with employees being more receptive to specific feedback (Kinicki, Prussia, Wu, & McKee-Ryan, 2004). Receiving feedback, for example, on the basis of a set of *behaviorally* defined competencies (see Woodruffe, 1993, p. 67) is likely to be more meaningful than that received from an omnibus indication of OJP. Third, a broader array of criteria may be easier for managers conducting developmental discussions with direct reports, with the likelihood that negative feedback will only be part of a broader picture including positive elements. Research shows that managers giving negative feedback tend to be more specific with the aim of mitigating unfavorable reactions (Larson, 1986).

A recent empirical study on operational criterion competency measures found promising effects. Catano et al. (2007) summarized their findings from a performance appraisal system based on a set of generalized behavior-based core competencies (also see Clardy, 2008; Prahalad & Hamel, 1990). As performance indicators, competency measures were found to be reliable for the assessment of police officers. The mean intrarater reliability (see Schmidt & Hunter, 1996) in this study was estimated at 0.89 (range = 0.57–0.94). In keeping with previous studies, interrater reliability was estimated within a range of 0.46–0.59 (see Viswesvaran, Ones, & Schmidt, 1996). Catano et al. report that employees and supervisors supported the appraisal system, and that competency scores predicted career advancement.

## A Behavioral Approach to Competencies

As noted above, our focus here is on behavioral competencies in line with Woodruffe (1993, p. 67), who suggests that “competencies are clusters of behavior. Their headings, such as self-confidence and incisiveness, aim to capture the essence of each cluster.” Such Woodruffe-defined competencies (WDCs) constitute lists of items that describe behaviors associated with essential elements of job performance. Behavioral descriptors are clustered under competency subheadings. For example, behavioral items for the competency tolerance could be expressed as the extent to which a manager (a) remains calm in the presence of a disgruntled employee and (b) conveys contentious employee information in a manner that is composed and controlled. We note that this behavioral focus is in contrast to the popular notion that competencies are akin to underlying traits (Boyatzis, 1982).

A feature of the Catano et al. (2007) study was the use of behaviorally anchored rating scales (see Borman, 1986) and, thus, their implied intention to conceptualize competencies as indicative of behavioral manifestations. Thus, in keeping with Woodruffe’s (1993) ideas, the results of the Catano et al. study provide evidence that competencies are most useful when treated as behavioral clusters. In this respect, competency titles should not be confused with or expected to perform in the same manner as underlying traits. The misconception that behavioral performance should manifest in a manner akin to stable

traits (i.e., strictly stable across situations) has led to problems in other areas, including multisource assessments (Darr & Catano, 2008; Mount, Judge, Scullen, Sytsma, & Hezlett, 1998) and assessment centers (ACs; Lance, Baranik, Lau, & Scharlau, 2009). Woodruffe's contention implies that behavioral clusters should not necessarily be expected to show signs of (strict) cross-situational consistency because behavioral performance reflects a multitude of situational and psychological characteristics. Yet, in spite of these constraints, we know that past behavior is the best predictor of future behavior (Barrick & Zimmerman, 2005; Wernimont & Campbell, 1968). According to behavioral consistency theory, there should be concordance among behaviors exhibited across similar situations (Wernimont & Campbell, 1968). However, unlike trait approaches, behavior-based competencies do not attribute this consistency to an alternative, underlying cause.

Arguments in favor of a behavioral competency approach in HRD particularly stem from the notion that, for developmental purposes, behavioral feedback is almost always considered superior to trait-based feedback (Hartmann, Roper, & Bradford, 1979; Kluger & DeNisi, 1996; Lance, 2008a, 2008b; Murphy & Cleveland, 1995). The behavioral approach carries fewer internal attributions than feedback based on (alleged) traits, and is thus likely to be perceived in a less threatening manner (Hartmann et al., 1979). Moreover, behavioral feedback is specified in more concrete terms than a trait-based approach, and is therefore more likely to lead to behavioral change (Jackson, Stillman, & Atkins, 2005; Lance, 2008a, 2008b). In this regard, Murphy and Cleveland (1991, p. 92) argue that “[f]eedback that is specific and behaviorally oriented is more likely to be useful than feedback that is general and vague.”

Although recent research by Catano and colleagues has furthered our knowledge of competency measures from psychometric perspectives (Catano et al., 2007; Darr & Catano, 2008), additional research is needed to establish the validity and utility of behavioral competencies used for HRD. In particular, investigations are warranted to explore multivariate relationships among “soft” behavioral competency measures used for development and “hard” objectively oriented work outcomes such as those based on target completion, tenure, and pay level (see Landy & Farr, 1983). Moreover, the use of target-based outcomes allows for an objectively oriented assessment that circumvents the potential common method issues in Bartram's (2005) study. Although these issues are pervasive in the use of competencies, we note that calls for such studies into competency measures have been made in the New Zealand context in particular, where the use of competency measures is popular, yet there is the common issue that empirical evidence is lacking (Markus et al., 2005; Ruth, 2006).

Our research aim is a general and exploratory one, in that we seek to investigate nomological relationships among a set of soft WDC measures and a set of hard objectively oriented work outcomes. An investigation into the interrelationships between WDCs and objective outcomes has potentially important implications for the manner in which WDCs are weighted and utilized for

assessment and development applications. In the sections that follow, we present results from a sample of entry-level employees who were rated on a set of WDCs constructed specifically for the organization under scrutiny. In terms of use, the focus of this study is on WDCs that are applied as tools for providing developmental feedback to employees as part of a performance appraisal system (see Murphy & Cleveland, 1991). The objective outcomes in our study constitute ratings of participants on the extent to which they had met financial targets, their tenure, and their pay level. The following sections describe the methodological approach employed in terms of the hard and soft measures used and how these were related using multivariate techniques. Results are then presented and discussed in terms of their relevance to research and practice in HRD.

## Method

**Participants.** Participants consisted of 118 entry-level employees working within the head office of a retail chain located nationwide throughout New Zealand. This sample constituted the entire data set available from this work group. The New Zealand context represents a developed Western-style country where competency use is prevalent, though underresearched (Markus et al., 2005; Ruth, 2006). Duties performed by participants included administrative and clerical tasks. The organization under scrutiny provides a range of consumer products such as housewares, hardware, music, and food items. Note that the organization here did not collect demographic information throughout the entire course of this study. As such, this information was only available for portions (i.e., around 50%) of the sample, and available data are reported below. Around 67% of the sample were female (33% male) and the mean age was 31.10 years ( $SD = 8.77$ ). In terms of ethnicity, the available information stated that most participants described themselves as New Zealand European (21%). Other individuals identified themselves as Asian (5%) and of Māori and Pacific Island origin (5%). With regard to education, 8% of the sample held a secondary school certificate, 7% held a university degree, 5% held a professional or industry qualification, and 2% held a postgraduate degree.

**Measures.** As part of their ongoing performance development, participants were rated on a WDC measure that included seven competency sub-headings (see Table 1). This measure was developed specifically for the organization, and assessment items were developed on the basis of an inductive job analysis (see Williams & Crafts, 1997). This process involved interviewing subject matter experts with a view to identifying critical and challenging work behaviors for assessment purposes (see the next section for details). In line with suggestions from the competency literature (Schippmann et al., 2000), and, thus, aligning these measures closely with the general competency approach, company strategy also guided item development (Catano et al., 2007; Lievens et al., 2004). Specifically, this organization valued a team

**Table 1. Definitions for Competency Measures**

<i>Team Members</i>		
<i>No.</i>	<i>Competency</i>	<i>Definition</i>
1.	Service action	At an operational level, works toward managing customer relations in an efficient and effective manner.
2.	Promotability	Is motivated to take on more responsibility and move upwards through the ranks of the organization.
3.	Team focus	Demonstrates a strong capacity to work effectively with other members of the work team.
4.	Commitment	Is demonstrably proud of working for the organization and supports its core purpose.
5.	Personal leadership	Is aware of his or her own strengths and weaknesses and works well under challenging circumstances.
6.	Driving business performance	Understands operational goals and works hard in his or her role to assist in achieving these goals.
7.	Focus on results	Reliably delivers expected results in a timely manner and maintains attention to detail.

focus, self-awareness, and the motivation to perform effectively among employees. For each assessment item that was developed, scale anchors ranged from 1 (*this behavior is rarely true for the employee*) to 4 (*this behavior is always true for the employee*).

*Prior Scale Development.* Woodruffe-defined competencies subscale development was completed in advance and involved the use of an oblique exploratory and, in turn, a confirmatory factor analysis (CFA; see Brown, 2006) with correlated latent variables on a sample of 225 entry-level participants. Note that these participants were from the same organization as those in the primary study here. The exploratory analysis returned a relatively clean seven-factor structure. For brevity, however, only the results of the CFA are presented here in detail. The resulting seven-factor CFA structure returned a solution that approached limits for acceptability (see Hu & Bentler, 1998, 1999) on a range of goodness-of-fit indexes ( $\chi^2 = 770.78$ ,  $df = 539$ ; squared root-mean residual = 0.05; root-mean-squared error of approximation = 0.04; Tucker-Lewis index = 0.93; comparative fit index = 0.94). The solution was also admissible (nonnegative unique variances, standardized estimates  $<|1|$ ; see Marsh, 1994) and submeasures showed evidence for internal consistency (see Table 2) according to commonly cited guidelines (Nunnally & Bernstein,

**Table 2. Descriptive Statistics and Correlations Among Competencies and Objectively Oriented Outcomes**

Criteria	M	SD	i	1	2	3	4	5	6	7	8	9	10
Service action	1.96	0.46	7	<b>0.85</b>									
Promotability	1.56	0.65	5	0.26**	<b>0.87</b>								
Team focus	2.18	0.48	6	0.57**	0.28**	<b>0.85</b>							
Commitment	2.18	0.60	2	0.43**	0.32**	0.53**	<b>0.85</b>						
Personal leadership	1.86	0.52	4	0.63**	0.28**	0.50**	0.36**	<b>0.79</b>					
Driving business performance	1.98	0.53	3	0.59**	0.31**	0.45**	0.39**	0.56**	<b>0.77</b>				
Focus on results	2.06	0.51	5	0.68**	0.22**	0.61**	0.44**	0.60**	0.54**	<b>0.85</b>			
Financial targets	1.01	0.05	1	0.35**	0.21*	0.22*	0.26**	0.22*	0.30**	0.42**	–		
Tenure	4.58	4.18	1	0.02**	0.02	–0.10	0.21*	0.03	0.07	0.01	0.05	–	
Pay level	2.41	0.83	1	–0.16**	0.17	–0.11	–0.05	–0.23*	–0.18	–0.23*	–0.13	0.14	–

Note: i = number of items; coefficients alpha as estimates of intrarater reliability (see Schmidt & Hunter, 1996) for competency measures appear in bold on the diagonal. The results of additional nonparametric analyses did not alter conclusions surrounding the patterns suggested above.

\* $p < 0.05$ .

\*\* $p < 0.01$ .

1994). As has been reported in previous studies (see Catano et al., 2007, for a summary), the competency submeasures were found to be intercorrelated (minimum  $r = 0.31$ , maximum  $r = 0.80$ ).

*Competency Ratings.* Participants in the study sample were rated on all seven WDCs by managers ranked one level senior to the ratees. Raters held a supervisory role over employees, which could have biased the WDC ratings to some degree. However, such effects were minimized through the use of clear behavioral descriptors that focused on elements of work performance (Landy & Farr, 1983). Moreover, raters were provided with a full-day course of frame-of-reference training (Bernardin & Buckley, 1982), which assists assessors to gain similar standards by which to judge performance. In this procedure, raters were initially familiarized with the instrument set. Then, and importantly, practice assessments took place on mock participants. In turn, ratings were shared among the rater group and any scores that deviated were discussed. This process was repeated to promote consistency among the rater group (see Pulakos, 1986).

*Objectively Oriented Work Outcomes.* Approximately 6 months later, participants were again rated by the same managers with respect to their financial targets for this period. In this regard, financial targets reflected the extent to which participants had met their individual budgetary expectations (e.g., the extent to which HRD coordinators met their budget in terms of funds spent on employee development). As such, the financial outcomes here are related to developmental aims because of the fact that employees had control over such outcomes. Scale anchors for this measure ranged from being well below target (with the lower anchor set at having achieved around 85% of their financial target) to exceptional. (The upper anchor is having overachieved the financial target, at 115%. Strictly speaking, percentages cannot, of course, be greater than 100, and 115% is merely represented here as an indication of being greater than the preset financial target.) Data were also obtained from organizational records on the length of tenure (number of years in the organization as a proxy for experience) along with their pay level (ranging from 1 = *low* to 7 = *high*). Note that pay level was determined by the focal organization based on a job evaluation, and reflects such issues as increasing job responsibility and job complexity.

*Analyses.* The summarization of data-driven relationships among constructs of interest was pertinent to this study. Given that this was an exploratory goal aimed at investigating patterns of covariation, an exploratory correlational technique was selected as the analytic approach of choice. Moreover, two sets of variables can be clearly partitioned in this study; specifically, WDCs and objectively oriented work outcomes. In this respect, canonical correlation analysis (CCA; see Fan, 1996; Sherry & Henson, 2005; Thompson, 1984) is well suited to these aims and was used to explore multivariate patterning among the variables under scrutiny. CCA holds the advantage of being capable of summarizing multivariate relationships between variable subsets



whilst acknowledging relationships within such subsets (Thompson, 1984). SPSS Statistics (version 17) was used to generate all of the analyses herein, with the use of the MANOVA and CANCELL syntax procedures.

## Results

Table 2 shows descriptive statistics for the sample, along with intercorrelations and reliability estimates for the competency measures. Note that reliability information was unavailable for the objective measures because they were taken at a single time and with a single item. Intrarater reliability was estimated by coefficient alpha (see Schmidt & Hunter, 1996) to determine the consistency with which raters assessed within each competency category. Among the seven competency measures for this group, all returned estimates of intrarater reliability that were within guidelines commonly cited for coefficient alpha (ranging from 0.77 to 0.87; see Kline, 1999; Nunnally & Bernstein, 1994). Also of note were the positive correlations observed between meeting financial targets and all seven WDC measures. The highest of these was with focus on results ( $r = 0.42, p < 0.01$ ). Among the other small intercorrelations, tenure was found to relate to commitment ( $r = 0.21, p < 0.05$ ) and pay level displayed small negative correlations with personal leadership ( $r = -0.23, p < 0.05$ ) and focus on results ( $r = -0.23, p < 0.05$ ). On the subject of objective outcomes, we assuage potential concerns about the nonnormality observed in the tenure variable (see Table 2) with the assertion of Hair, Anderson, Tatham, and Black (1998, p. 448) that CCA “can accommodate any metric variable without the strict assumption of normality.”

Following from the bivariate analysis, a CCA (see Fan, 1996; Hair et al., 1998; Tabachnick & Fidell, 2001; Thompson, 1984) was conducted and included two variable subsets comprising (a) competency measures and (b) objectively oriented outcome measures (see Table 3). The CCA generated three successive functions with squared canonical correlations ( $R_c^2$ ) of 0.26, 0.12, and 0.07, respectively. The full model, including all functions (1–3), was statistically significant, Wilks’s  $\lambda = 0.608, F(21, 310.67) = 2.798, p < 0.01$ . From a multivariate perspective,  $1 - \text{Wilks’s } \lambda$  revealed that the variable subsets shared around 39% of their variance. Note, however, that Function 3 was non-significant. Thus, only Functions 1 and 2 were worthy of consideration given their relative  $R_c^2$  estimates and the no significance observed in Function 3.

Table 3 presents the standardized canonical function coefficients and structure coefficients ( $r_s$ ) for Functions 1 and 2. Sherry and Henson (2005) suggest that  $r_s$  estimates  $> |0.45|$  indicate noteworthy contributions. In keeping with this guideline, Table 3 shows that Function 1 summarized the relationships among competency measures and the extent to which participants met financial targets. The CCA suggested that the competency and outcome variates shared around 26% of their variance within Function 1 ( $R_c^2 = 0.26$ ). All but two competency measures contributed notably in this regard and the most substantial

Table 3. Canonical Solution Including Competency and Objectively Oriented Outcome Measures

	Function 1			Function 2			$h^2$ (%)
	Coef	$r_s$	$r_s^2$ (%)	Coef	$r_s$	$r_s^2$ (%)	
<b>Competency measures</b>							
Service action	-0.26	<b>-0.71</b>	50.41	0.27	0.07	0.49	<b>50.90</b>
Promotability	-0.11	-0.27	7.29	0.94	<b>0.70</b>	49.00	<b>56.29</b>
Team focus	0.42	-0.39	15.21	0.35	0.21	4.41	19.62
Commitment	-0.23	<b>-0.57</b>	32.49	-0.44	-0.08	0.64	33.13
Personal leadership	0.07	<b>-0.54</b>	29.16	-0.61	-0.29	8.41	37.57
Driving business performance	-0.26	<b>-0.66</b>	43.56	-0.32	-0.10	1.00	44.56
Focus on results	-0.78	<b>-0.88</b>	77.44	0.22	0.02	0.04	<b>77.48</b>
$R_c^2$			25.88			11.68	
<b>Outcome measures</b>							
Financial targets	-0.88	<b>-0.93</b>	86.49	0.49	0.36	12.96	<b>99.45</b>
Tenure	-0.24	-0.23	5.29	-0.55	-0.41	16.81	22.10
Pay level	0.31	0.40	16.00	0.84	<b>0.70</b>	49.00	<b>65.00</b>

Note: Structure coefficients ( $r_s$ ) > |0.45| and communality estimates ( $h^2$ ) > 45% appear in bold (see Sherry & Henson, 2005). Coef = standardized canonical function coefficient;  $r_s^2$  = squared structure coefficient;  $R_c^2$  = squared canonical correlation.

impacts were associated with focus on results ( $r_s = -0.88$ ), service action ( $r_s = -0.71$ ), and driving business performance ( $r_s = -0.66$ ). Note that because the signs of the  $r_s$  coefficients in the competency variate were the same as those in the other criteria variate, directionwise, these relationships are interpreted as being positive (see Table 3). Function 2 summarized relationships between competency measures and pay level. Within Function 2, the CCA suggested that the competency and outcome variates shared around 12% of their variation ( $R_c^2 = 0.12$ ). In terms of relative impact, promotability was the only competency that held a noteworthy relationship with pay level ( $r_s = 0.70$ ).

## Discussion

Competencies represent an approach that is popular in practice (Boyatzis, 2008; Schippmann et al., 2000). However, although they have been successfully utilized for employee development (Riggio & Lee, 2007), little is known about how developmental competency measures relate to objective work outcomes. The results of the present study suggest that WDCs, when set up according to appropriate psychometric principles, can reflect meaningful multivariate relationships with objectively oriented work outcomes. From a multivariate perspective, it was found that five out of the seven WDCs under scrutiny displayed a degree of correlation with the extent to which participants met financial objectives ( $R_c^2 = 0.26$ ). The strongest relationships were associated with focus on results, service action, and driving business performance. One WDC, promotability, predicted pay level ( $R_c^2 = 0.12$ ). Thus, intuitive links were evident between WDCs and work outcomes in the present sample. For example, it seems reasonable that the competencies focus on results and service action should show relatively greater strength to financial outcomes. It is also intuitive that the competency promotability should relate positively to an outcome associated with pay level. Such links may be useful for the purposes of weighting certain WDCs over others for developmental purposes. For example, in Table 3, if the intention was to provide developmental feedback to help employees meet financial targets, then the  $r_s$  estimates indicate that development activities should primarily focus on the behaviors underlying the competencies of focus on results, service action, and driving business performance.

Previous research has implied that, where competency measures strongly relate to OJP, they may be redundant (Bartram, 2005). However, this result may have been due to common-method bias. In the present study, we assessed WDCs against objectively oriented criteria to assist in avoiding such bias. Although meaningful relationships between competency measures and other work outcomes emerged in this study, these were not extreme to the extent that they suggested redundancy. The results here revealed that relationships among WDCs and work criteria remained when such measures differed in form, although the relationships were smaller than those found in Bartram's study.

In our study, competencies were treated as though they were clusters of work-related behaviors, in a similar manner to the measures reported in Catano et al. (2007). Treating competency data in this way led to a reasonable CFA structure and meaningful patterns of covariation with objectively oriented work outcomes. In this respect, our results suggest that when differentiated competency measures are used along with objective work outcomes, CCA (or other multivariate techniques) can provide guidance in terms of how particular competency measures may be ranked over others.

As we built on the empirical work on competency measures (Bartram, 2005; Bartram et al., 2002; Catano et al., 2007; Tett, Guterman, Bleier, & Murphy, 2000; Xiao, 2006), we also found that WDCs can produce reliable, valid, and meaningful information for developmental purposes. In this regard, we agree with the assertion of Catano et al. (2007) that when “competency modeling is done right, it can provide useful information” (p. 226). This reprises an issue that has been raised by other authors around the necessity for competencies to be defined in exacting terms and in such a way that lends them to empirical scrutiny (Barrett & Depinet, 1991; Meriot, 2005). We suggest that the optimal way forward for competencies is to define them in the manner suggested by such authors as Woodruffe and Catano et al. and to treat them as clusters of behavioral descriptors of effective performance on the job, rather than as measures of traits (in the Boyatzis, 1982, tradition).

**Limitations.** The generalizability of our findings is limited by a focus on a particular competency approach in a single organization. The relatively small sample may also restrict the generalizability of our results. Nonetheless, we note that much competency research, since it is practice-based, will present this issue. Related to this is the idea, presented earlier, that competencies and their applications are, by their very nature, broadly defined. However, we anticipate that the competencies presented here will have clear overlaps with other competency measures in use. Thus, our results may serve as a comparison or act as a starting point for human resource practitioners and researchers.

We also found that the WDC set in this study did not relate strongly to tenure. However, Table 3 reveals a possible borderline negative multivariate relationship between promotability and tenure (within Function 2,  $r_s$  [promotability] = 0.70,  $r_s$  [tenure] = -0.41). This could present an area for further scrutiny. For example, studies could investigate whether increased tenure is associated with diminishing potential for promotion. Such effects could, however, be explained by a ceiling placed on promotion opportunities as one progresses up the ranks in an organization. Another possibility is that newcomers to an organization may be more readily promoted. Further research into such patterns may assist in terms of finding factors, legitimate, extraneous, or otherwise, that influence promotion decisions.

Note also that the specific WDC team focus did not predict any of the objectively oriented work outcomes included here. This does not necessarily mean that team focus is unimportant. Rather, it may be that this interpersonally

oriented WDC may be related to other criteria, such as team goal achievement or social network characteristics (Morrison, 2002). Alternatively, it may represent participants' decisions to trade off efforts for their own benefit versus those oriented toward helping the team. Moreover, we acknowledge criticisms in the literature vis-à-vis the potentially weak relationship between the concept job performance versus measures of job performance (Murphy, 2008). Our contention is that job performance is likely to be multifaceted and broadly defined. As such, a single measure of this concept will likely fail to cover the necessary content domain. To provide reassurance in this regard, we triangulated our performance measures to the greatest degree possible. As such, three objective measures were included that took quite different perspectives on the job performance domain.

Our results take a step toward answering the calls of Lievens et al. (2004) for investigations into the psychometric properties of competency measures. However, only information on the intrarater reliability of our competency measures was available from the organization under scrutiny. Although comprehensive studies have been performed on the reliability of such measures as job performance ratings (Viswesvaran et al., 1996), further studies on the interrater reliability of competency measures are necessary, to add to the scarce literature in this regard (e.g., Catano et al., 2007). Sample-specific reliability estimates were also unavailable for our objectively oriented job performance measures, and future studies could incorporate such information into analyses that build upon the findings reported here. In a similar vein, future studies could expand upon our research by investigating relationships among competencies and additional work outcomes (e.g., training criteria and production data).

We also attempted to mitigate criterion contamination concerns because, although competencies were rated on judgmental scales, financial target setting was measured against preset targets that were not intended to be influenced by competency scores. On this note, the competency measures in this study were used by the organization for developmental purposes only and were not incorporated into decisions about promotion or pay. This said, our efforts here do not completely eliminate the possibility of criterion contamination, and future studies could look toward removing such potential influences.

Also, and as an aside, we purposely did not include redundancy analyses and the resulting redundancy coefficients that can be generated in CCA (see Thompson, 1984). This is because redundancy coefficients, unlike others under CCA, are not multivariate in nature (Enders, 2003). For this reason and others, a myriad of problems arise in their interpretation, leading Roberts (1999) to conclude that redundancy coefficients should (almost) never be used or interpreted.

**Implications for Human Resource Development Practice and Future Research.** The criteria upon which to provide developmental feedback to employees has long presented a quandary to practitioners and researchers in HRD. Competencies represent a vastly popular approach in this respect (Boyatzis,

2008), yet the validity of competency measures has been disputed (Barrett & Depinet, 1991). Our argument is that part of the problem surrounds the popular idea that competencies are underlying and relatively stable traits that drive job performance (see Boyatzis, 1982).

Job performance ratings are routinely provided on the basis of a rating scale from a single source (Murphy, 2008; Murphy & Cleveland, 1991). In such circumstances, we suggest that competencies should be defined as clusters of similar behaviors, rather than traits, in a manner akin to the approach defined by Woodruffe (1993). So-called Woodruffe-defined competencies (WDCs) work on a behavioral basis, and make some assumptions about behavioral consistency across similar situations (Wernimont & Campbell, 1968). In this regard, WDCs allow for behavioral feedback to be provided in developmental scenarios. Such forms of feedback have been found to be superior to trait-based feedback in terms of improving employee performance (Kluger & DeNisi, 1996).

Our contention is that in order for competencies to be helpful to the practitioner, they need to be developed according to psychometric principles (cf. Brannick, 2008). Thus, in the present study, CFA was used to develop a set of seven competency categories. The behavioral items within each category were used as developmental prompts. To take this further, we also related WDC aggregate scores to objective measures of performance and found meaningful multivariate relationships (through CCA) between particular WDC clusters and objective outcomes (e.g., the WDC focus on results and financial targets). This provided a set of weights that suggested which WDCs were important in terms of predicting particular outcomes. Circumstances are likely to arise where an organization wishes to foster development around a certain objective outcome. In such cases, the weights provided in CCA (or other multivariate techniques) could be used to guide the practitioner in terms of which competency clusters should be emphasized over others.

In terms of future research, an obvious area for further investigation relates to the use of single methods or multiple methods for competency development and improvements in employee performance. We argue for a definition of competencies as clusters of behavioral items. Thus, the competency labels, under this definition, are nothing more than titles used, out of convenience, to categorize similar behavioral items. Within such a definition, trait notions are not relevant. The fundamental idea that competencies constitute traits has caused problems in techniques involving multiple methods, as exemplified in the AC literature. Thus, empirical research indicates that behavioral indicators cluster within AC exercises rather than within competencies (Lance, Lambert, Gewin, Lievens, & Conway, 2004), suggesting consistency of behaviors within situations. Aligning with such findings, we argue that competency titles are far less important than a realistic acknowledgment of the manner in which behavioral indicators tend to cluster. Thus, if behavioral items are found to cluster around competency titles, then the titles can be used, but should not

be interpreted to suggest underlying traits. If, however, behavioral items cluster around exercises in an AC, then the resulting factors should, in our view, be interpreted as exercise-specific performance clusters rather than method bias (Lance, 2008a, 2008b). Here, the use of competency titles might be misleading, and rather, exercise titles could be used for aggregate scores to emphasize the meaning of such scores (Jackson, Stillman, & Englert, 2010). To this end, empirical research is needed on whether a competency approach based on clusters of behavioral items (i.e., a WDC approach) leads to greater improvement in HRD scenarios than the traditional approach to competencies, which is to treat them as underlying traits (Boyatzis, 1982). Such research could assist to develop a course of action for competencies that are reliable, valid, and of assistance to the practice of HRD.

## References

- Athey, T. R., & Orth, M. S. (1999). Emerging competency methods for the future. *Human Resource Management, 38*, 215–226.
- Austin, J. T., & Crespin, T. R. (2006). From “criterion problem” to problems of criteria in industrial and organizational psychology: Progress, pitfalls, and prospects. In W. Bennett, Jr., C. E. Lance, & D. J. Woehr (Eds.), *Performance measurement: Current perspectives and future challenges*. Mahwah, NJ: Erlbaum.
- Barrett, G. V., & Depinet, R. L. (1991). A reconsideration of testing for competence rather than for intelligence. *American Psychologist, 46*, 1012–1024.
- Barrick, M. R., & Zimmerman, R. D. (2005). Reducing voluntary, avoidable turnover through selection. *Journal of Applied Psychology, 90*, 159–166.
- Bartram, D. (2005). The great eight competencies: A criterion-centric approach to validation. *Journal of Applied Psychology, 90*, 1185–1203.
- Bartram, D., Robertson, I. T., & Callinan, M. (2002). Introduction: A framework for examining organizational effectiveness. In I. T. Robertson, M. Callinan, & D. Bartram (Eds.), *Organizational effectiveness: The role of psychology* (pp. 1–10). Chichester, United Kingdom: Wiley.
- Bernardin, H. J., & Buckley, M. R. (1982). Strategies in rater training. *Academy of Management Review, 6*, 205–212.
- Borman, W. C. (1986). Behavior-based rating scales. In R. A. Berk (Ed.), *Performance assessment: Methods and applications* (pp. 100–120). Baltimore: Johns Hopkins University Press.
- Boyatzis, R. E. (1982). *The competent manager: A model for effective performance*. New York: Wiley.
- Boyatzis, R. E. (2008). Competencies in the 21st century. *Journal of Management Development, 27*, 5–12.
- Brannick, M. T. (2008). Back to basics of test construction and scoring. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 131–133.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.
- Catano, V. M., Darr, W., & Campbell, C. A. (2007). Performance appraisal of behavior-based competencies: A reliable and valid procedure. *Personnel Psychology, 60*, 201–230.
- Clardy, A. (2008). Human resource development and the resource-based model of core competencies: Methods for diagnosis and assessment. *Human Resource Development Review, 7*, 387–407.
- Darr, W., & Catano, V. M. (2008). Multisource assessments of behavioral competencies and selection interview performance. *International Journal of Selection and Assessment, 16*, 68–72.
- Enders, C. K. (2003). Performing multivariate group comparisons following a statistically significant MANOVA. *Measurement and Evaluation in Counseling and Development, 36*, 40–56.

- Fan, X. (1996). Canonical correlation analysis as a general analytic model. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 4, pp. 71–94). Greenwich, CT: JAI.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Hartmann, D. P., Roper, B. L., & Bradford, D. C. (1979). Some relationships between behavioral and traditional assessment. *Journal of Behavioral Assessment, 1*, 3–21.
- Hoffmann, T. (1999). The meanings of competency. *Journal of European Industrial Training, 23*, 275–285.
- Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424–453.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. *Structural Equation Modeling, 6*, 1–55.
- Jackson, D. J. R., Stillman, J. A., & Atkins, S. G. (2005). Rating tasks versus dimensions in assessment centers: A psychometric comparison. *Human Performance, 18*, 213–241.
- Jackson, D. J. R., Stillman, J. A., & Englert, P. (2010). Task-based assessment centers: Empirical support for a systems model. *International Journal of Selection and Assessment, 18*, 141–154.
- Kinicki, A. J., Prussia, G. E., Wu, B., & McKee-Ryan, F. M. (2004). A covariance structure analysis of employees' response to performance feedback. *Journal of Applied Psychology, 89*, 1057–1069.
- Kline, P. (1999). *Handbook of psychological testing* (2nd ed.). New York, NY: Routledge.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and preliminary feedback theory. *Psychological Bulletin, 119*, 254–284.
- Kurz, R., & Bartram, D. (2002). Competency and individual performance: Modeling the world of work. In I. T. Robertson, M. Callinan, & D. Bartram (Eds.), *Organizational effectiveness: The role of psychology* (pp. 227–255). Chichester, United Kingdom: Wiley.
- Lance, C. E. (2008a). Why assessment centers do not work the way they are supposed to. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 84–97.
- Lance, C. E. (2008b). Where have we been, how did we get there, and where shall we go? *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 140–146.
- Lance, C. E., Baranik, L. E., Lau, A. R., & Scharlau, E. A. (2009). If it ain't trait it must be method: (Mis)application of the multitrait–multimethod methodology in organizational research. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in organizational and social sciences*. New York: Routledge.
- Lance, C. E., Lambert, T. A., Gewin, A. G., Lievens, F., & Conway, J. M. (2004). Revised estimates of dimension and exercise variance components in assessment center postexercise dimension ratings. *Journal of Applied Psychology, 89*, 377–385.
- Landy, F. J., & Farr, J. L. (1983). *The measurement of work performance: Methods, theory, and applications*. New York: Academic Press.
- Larson, J. R. (1986). Supervisors' performance feedback to subordinates: The impact of subordinate performance valence and outcome dependence. *Organizational Behavior and Human Decision Processes, 37*, 391–408.
- Lievens, F., Sanchez, J. I., & de Corte, W. (2004). Easing the inferential leap in competency modeling: The effects of task-related information and subject matter expertise. *Personnel Psychology, 57*, 881–904.
- Markus, L. H., Cooper-Thomas, H. D., & Allpress, K. N. (2005). Confounded by competencies? An evaluation of the evolution and use of competency models. *New Zealand Journal of Psychology, 34*, 117–126.
- Marsh, H. W. (1994). Confirmatory factor analyses of factorial invariance: A multifaceted approach. *Structural Equation Modeling, 1*, 5–34.
- Meriot, S. (2005). One or several models for competence descriptions: Does it matter? *Human Resource Development Quarterly, 16*, 285–292.



- Morrison, E. W. (2002). Newcomers' relationships: The role of social network ties during socialization. *Academy of Management Journal*, *45*, 1149–1160.
- Mount, M. K., Judge, T. A., Scullen, S. C., Sytsma, M. R., & Hezlett, S. A. (1998). Trait, rater, and level effects in 360-degree performance ratings. *Personnel Psychology*, *51*, 557–575.
- Murphy, K. (2008). Explaining the weak relationship between job performance and ratings of job performance. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *1*, 148–160.
- Murphy, K., & Cleveland, J. N. (1991). *Performance appraisal: An organizational perspective*. Needham Heights, MA: Allyn and Bacon.
- Murphy, K., & Cleveland, J. N. (1995). *Understanding performance appraisal: Social, organizational and goal-based perspectives*. Thousand Oaks, CA: Sage.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
- 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.
- Prahalad, C. K., & Hamel, G. (1990). The core competence of the corporation. *Harvard Business Review*, *68*, 79–91.
- Pulakos, E. D. (1986). The development of training programs to increase accuracy with different rating tasks. *Organizational Behavior & Human Decision Processes*, *38*, 76–91.
- Riggio, R. E., & Lee, J. (2007). Emotional and interpersonal competencies and leader development. *Human Resource Management Review*, *17*, 418–426.
- Roberts, J. K. (1999). Canonical redundancy (Rd) coefficients: They should (almost never) be computed and interpreted. In B. Thompson (Ed.), *Advances in social science methodology* (pp. 333–341). Stamford, CT: JAI Press.
- Ruth, D. (2006). Frameworks of managerial competence: Limits, problems and suggestions. *Journal of Managerial Psychology*, *30*, 206–226.
- Schippmann, J. S., Ash, R. A., Battista, M., Carr, L., Eyde, L. D., Hesketh, B., et al. (2000). The practice of competency modeling. *Personnel Psychology*, *53*, 703–740.
- Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, *1*, 199–223.
- Sherry, A., & Henson, R. K. (2005). Conducting and interpreting canonical correlation analysis in personality research: A user-friendly primer. *Journal of Personality Assessment*, *84*, 35–46.
- Spencer, L. M., & Spencer, S. M. (1993). *Competence at work*. New York: Wiley.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Pearson.
- Tett, R. P., Guterman, H. A., Bleier, A., & Murphy, P. J. (2000). Development and content validation of a “hyperdimensional” taxonomy of managerial competence. *Human Performance*, *13*, 205–251.
- Thompson, B. (1984). *Canonical correlation analysis: Uses and interpretation*. Newbury Park, CA: Sage.
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, *81*, 557–574.
- Wernimont, P. F., & Campbell, J. P. (1968). Signs, samples, and criteria. *Journal of Applied Psychology*, *52*, 372–376.
- Williams, K. M., & Crafts, J. L. (1997). Inductive job analysis: The job/task inventory method. In D. L. Whetzel & G. R. Wheaton (Eds.), *Applied measurement methods in industrial psychology* (pp. 51–88). Palo Alto, CA: Davies-Black Publishing.
- Wilson, M. A., Levine, J. D., Morgeson, F. P., Thomsen, D. J., & Harvey, R. J. (2007). *Unanswered questions in competency modeling and job analysis*. Paper presented at the 22nd Annual SIOF Conference, New York.
- Woodruffe, C. (1991). Competent by any other name. *Personnel Management*, *23*, 30–33.

- Woodruffe, C. (1993). *Assessment centres: Identifying and developing competence* (2nd ed.). London: Institute of Personnel Management.
- Xiao, J. (2006). Survey ranking of job competencies by perceived employee importance: Comparing China's three regions. *Human Resource Development Quarterly*, 17, 371–402.

*Duncan J. R. Jackson is with The University of Seoul.*

*Helena D. Cooper-Thomas is with the Department of Psychology, The University of Auckland.*

*Marco van Gelderen is with the Faculty of Economics and Business Administration, Department of Management & Organization, VU Centre for Entrepreneurship.*

*Jane Davis is with the Warehouse Group Limited.*

