Singapore Management University
Institutional Knowledge at Singapore Management University

12-2004

Easing the inferential leap in competency modeling: The effects of task-related information and subject matter expertise

Filip LIEVENS
Singapore Management University, filiplievens@smu.edu.sg

Juan I. SANCHEZ
Florida International University

Wilfred DE CORTE
Ghent University
DOI: https://doi.org/10.1111/j.1744-6570.2004.00009.x

EASING THE INFERENTIAL LEAP IN COMPETENCY MODELING: THE EFFECTS OF TASK-RELATED INFORMATION AND SUBJECT MATTER EXPERTISE

FILIP LIEVENS
Department of Personnel Management and
Work and Organizational Psychology
Ghent University, Belgium

JUAN I. SANCHEZ
Department of Management and International Business
Florida International University

WILFRIED DE CORTE
Department of Data Analysis
Ghent University, Belgium

Despite the rising popularity of the practice of competency modeling, research on competency modeling has lagged behind. This study begins to close this practice–science gap through 3 studies (1 lab study and 2 field studies), which employ generalizability analysis to shed light on (a) the quality of inferences made in competency modeling and (b) the effects of incorporating elements of traditional job analysis into competency modeling to raise the quality of competency inferences. Study 1 showed that competency modeling resulted in poor interrater reliability and poor between-job discriminant validity amongst inexperienced raters. In contrast, Study 2 suggested that the quality of competency inferences was higher among a variety of job experts in a real organization. Finally, Study 3 showed that blending competency modeling efforts and task-related information increased both interrater reliability among SMEs and their ability to discriminate among jobs. In general, this set of results highlights that the inferences made in competency modeling should not be taken for granted, and that practitioners can improve competency modeling efforts by incorporating some of the methodological rigor inherent in job analysis.

In recent years, the practice of competency modeling has made rapid inroads in organizations (Lucia & Lepsinger, 1999; Schippmann, 1999). In contrast to traditional job analysis, competency modeling ties the derivation of job specifications to the organization's strategy, which, together

with nonstrategic job requirements, are used to generate a "common language" in the form of a set of human attributes or individual competencies. This same set of competencies usually serves as a platform for various HR practices such as performance evaluation, compensation, selection, and training (Schippmann et al., 2000). The calls for strategic alignment of HR practices make competency modeling a timely mechanism to build the organization's strategy into all HR practices (Becker, Huselid, & Ulrich, 2001). An inspection of the ABI/INFORM database attests to the impetus of competency modeling, with over 500 articles on this topic published between 1995 and 2003, in sharp contrast to the 87 articles published between 1985 and 1995.

Contrary to the flourishing popularity of competency modeling among practitioners, the scientific community has regarded competency modeling with some degree of skepticism. The validity of "competencies" as measurable constructs appears to be at the core of this controversy (Barrett & Callahan, 1997; Barrett & Depinet, 1991; Lawler, 1996; Pearlman, 1997). Specifically, the process of deriving competencies requires a rather large inferential leap because competency modeling often fails to focus on detailed task statements prior to inferring competencies (Schippmann et al., 2000). In the U.S., this aspect of competency modeling appears to be problematic in light of current, quasi-legal standards such as the Uniform Guidelines on Employee Selection Procedures (1978), which require demonstrable linkages between job specifications such as knowledge, skills, abilities, and other requirements (KSAOs) on the one hand and important job behaviors on the other hand. In the absence of these linkages to important job behaviors, Sanchez and Levine (2001) observed that "*making and justifying inferential leaps on the slippery floor of behaviorally fuzzy competencies is certainly a methodological challenge*" (p. 85).

To date, there is a paucity of empirical studies actually scrutinizing the quality of the inferences required in competency modeling. Even more important, virtually no studies have compared competency modeling to more traditional job analysis approaches (for an exception, see Morgeson, Delaney-Klinger, Mayfield, Ferrara, & Campion, in press). Therefore, it is still unclear whether and how the assumed lack of task information inherent in competency modeling detracts from the quality of its inferences. The dearth of research on this issue is surprising because the quality of the inferences made in competency modeling has not only legal ramifications regarding possible violations of sound measurement procedures but also practical consequences such as HR practices that fail to truly leverage the organization's human resource capital (Becker et al., 2001).

We present three studies that begin to close this practice–science gap by examining the quality of inferences drawn in competency modeling. Specifically, we assessed the quality of inferences made in

competency modeling and whether elements of traditional job analysis (i.e., task-related information and subject matter expertise) can be fruitfully incorporated into competency modeling to enhance the quality of such inferences. We operationalized the level of "quality" in terms of both interrater reliability and discriminability among jobs and/or competencies. These criteria are important because they reflect underlying issues of reliability and discriminant validity in work analysis data (Dierdorff & Wilson, 2003; Morgeson & Campion, 1997).

Study Background

The traditional task analysis approach to job analysis provides an indirect estimation of knowledge, skills, abilities, and other characteristics (KSAOs; Gatewood & Field, 2001, pp. 367–380; Morgeson & Campion, 2000). That is, the complex inferential leap from the job to the specification of KSAOs (Cornelius & Lyness, 1980; Morgeson & Campion, 1997) is broken down into a series of more manageable steps. First, the various job tasks are identified. Next, subject matter experts (SMEs) are asked to judge the importance or criticality of these tasks. Finally, given these tasks, SMEs make inferences about which KSAOs are most important. The methodological rigor of this step-by-step approach lends credence to the SMEs' inferences. Although the widely employed task analysis approach relies on this indirect estimation method, it should be acknowledged that job-analytic approaches vary widely in the extent to which they focus on job tasks and other descriptors (e.g., Christal, 1974; Fine, 1988; Lopez, Kesselman, & Lopez, 1981; Prien, Prien, & Gamble, 2004).

Other forms of work analysis involve directly estimating KSAOs by asking SMEs to rate the importance of various KSAOs for a given job. In this direct estimation method, the intermediate step of specifying tasks is not an explicit requirement. This kind of holistic KSAO judgment calls for a larger inferential leap than asking SMEs to infer KSAOs from specific task statements. An example of direct estimation is the job element method (Primoff, 1975).

How are the focal worker attributes for a given job determined in competency modeling? Schippmann et al. (2000) tried to delineate the major characteristics of competency modeling by prompting the opinion of 37 work analysis experts and authority figures. As concluded by Schippmann et al., competency modeling is less rigorous than traditional, task-based job analysis. In most competency modeling approaches, only one type of descriptor information (i.e., competencies) is gathered, and SMEs are not provided with detailed task statements prior to making inferences about which competencies are important. Instead, similar to a direct estimation method, competencies are inferred from just a broad job description plus information about the organization's strategy.

Although the provision of strategic information might impose a common frame of reference on SMEs, thereby enhancing their interrater reliability (cf. Sulsky & Day, 1994), Schippmann et al. concluded that the reduced methodological rigor (i.e., the absence of detailed task statements) of most competency modeling approaches increases the difficulty of the inferences required from SMEs.

Indeed, various empirical studies in the traditional job analysis literature found that raters are less capable to make reliable judgments for the entire job than for narrower descriptors such as task statements (Butler & Harvey, 1988; Dierdorff & Wilson, 2003; Sanchez & Levine, 1989, 1994). Hughes and Prien (1989) showed that even when SMEs were given task statements, a relatively large inferential leap was still required, as evidenced by the moderate interrater agreement reported. Finally, Morgeson et al. (in press) found that global judgments similar to those made in competency modeling were more inflated than task-level judgments.

From a theoretical point of view, the derivation of worker attributes required for job performance can be conceptualized as an inferential decision, in which job events need to be recalled and then reduced to a set of dimensions (Holland, Holyoak, Nisbett, & Thagard, 1986). Clearly, the amount of information that needs to be recalled and integrated into a set of job-level competencies exceeds that required to make similar judgments at the narrower task level. Apart from judgment theory, categorization theory (Srull & Wyer, 1980) also predicts more biases for holistic judgments than for task-based judgments. The reason is that experts might make judgments on the basis of what they think the job involves (a category) instead of on the basis of factual tasks. In short, empirical evidence and theoretical arguments suggest that the quality of inferences is higher when task-based information is available.

The SMEs in Schippmann et al.'s (2000) study hinted that the future of work analysis might consist of blending the two approaches previously discussed (i.e., task analysis and competency modeling). This blended approach represents an effort to incorporate not only the organization's strategy into the derivation of broad worker attributes or "competencies," but also the methodological rigor of task analysis, where SMEs are provided with task statements prior to inferring KSAOs.

Such a blended approach might improve the quality of the competency inferences because it capitalizes on the strengths of both the task analysis and competency modeling approaches. First, information about the organization's strategy provides SMEs with a common frame of reference regarding the strategic implications for the HR function, which should facilitate the process of identifying worker attributes or competencies aligned with such a strategy. Second, the information about important task statements should decrease the complexity of the competency

judgments required from SMEs, who would have a more concrete referent of job behaviors than that provided by a description of the organization's strategy.

The primary purpose of the studies to be presented here (see Study 1 and 3) was to systematically compare the quality of inferences in various work analysis approaches. On the basis of the research reviewed above, we expected that the quality of inferences in the blended approach would be higher than the quality of inferences in the task-based approach, which in turn would be higher than the quality of inferences in the competency modeling approach.

Besides the inclusion of task information, traditional job analysis further outperforms competency modeling in terms of its methodological rigor when composing SME panels. As noted by Schippmann et al. (2000), raters barely familiar with the job are sometimes employed in competency modeling. In addition, in some competency modeling projects, only a few people select the competencies deemed to be important. To date, we do not know how insufficient subject matter expertise might impact the quality of inferences made in competency modeling. In a similar vein, it is unknown how many SMEs or which type of SMEs are needed to obtain reliable competency ratings.

Traditional job-analytic research might shed light on these questions. In fact, job analysis approaches have typically preferred job incumbents because the quality of their ratings is superior to that of ratings made by naïve raters (usually college students) (Cornelius, DeNisi, Blencoe, 1984; Friedman & Harvey, 1986; Voskuijl & Van Sliedregt, 2002). However, familiarity with the job such as that possessed by job incumbents might be a necessary albeit insufficient requirement for accurately determining job specifications. Specifically, it has been argued that other sources such as supervisors, HR specialists, and internal customers should probably supplement the information provided by job incumbents (Brannick & Levine, 2002; Sanchez, 2000). For instance, Hubbard, McCloy, Campbell, Nottingham, Lewis, and Rivkin (1999) argued that job incumbents might have difficulty judging the relevance to their job of abstract attributes such as competencies because many workers might have never distinguished between their personal attributes and those required by their job. The literature comparing incumbent to nonincumbent ratings has also shown that, depending on the level of job satisfaction and occupational complexity, incumbent ratings do not necessarily agree with ratings from other sources (Gerhart, 1988; Sanchez, 2000; Sanchez, Zamora, & Viswesvaran, 1997; Spector & Jex, 1991). Another reason for supplementing incumbent ratings with those from other sources is that incumbents may lack sufficient foresight and knowledge of technological innovations to define strategically aligned work requirements such as those demanded by competency modeling.

In short, a second purpose of this paper was to compare the quality of inferences between a group of naïve student raters (see Study 1) and a group of experienced SMEs (incumbents, supervisors, HR specialists, and internal customers) in a real organization (see Studies 2 and 3). On the basis of the research mentioned above, we expected that the quality of inferences would be higher among a variety of job experts than among naïve raters.

Overview of Studies

In the following sections, we present three studies that, together, enabled us to test our predictions about the effects of subject matter expertise and task-related information. Study 1 was a laboratory study using naïve student raters wherein we compared the quality of inferences of all three work analysis approaches discussed above (i.e., task-based job analysis, competency modeling, and blended approach). Study 2 focused on competency modeling. In particular, the quality of inferences as made in competency modeling was investigated in an actual organization among different types of SMEs (i.e., incumbents, supervisors, HR specialists, and internal customers). Finally, Study 3 was a quasi-experiment wherein the competency modeling and blended approaches were compared using a group of SMEs from the same organization.

Note that the jobs rated differed across these three studies. Therefore, job-specific effects might have been confounded with our manipulations. However, we felt that keeping familiarity with the job constant across studies was more important than keeping the job content per se constant (see Hahn & Dipboye, 1988). Students are typically not familiar with jobs such as those employed in Study 2 and 3, and therefore their ability to rate such jobs is questionable. Therefore, jobs that were familiar to student raters were chosen in Study 1. In a similar vein, familiarity with the job served as prime criterion for including individuals as SMEs in Studies 2 and 3.

Study 1

Method

Study 1 compared the quality of inferences in three work analysis approaches (i.e., task-based job analysis, competency modeling, and blended approach) among student raters. Participants were 39 graduate students (31 men, 8 women; mean age = 22.1 yrs.) in industrial and organizational psychology. First, they received a 2-hour lecture about work analysis and a 1-hour workshop about a specific competency modeling

technique (see below) wherein they practiced this technique by determining the competencies of one job (assistant to the human resources manager). Then, they received feedback about the competencies that were chosen by a panel of SMEs for this job. Next, their task was to determine the competencies of three jobs (accountant, executive secretary, and sales manager). In a pilot study, a similar group of 15 graduate students (11 women, 4 men; mean age = 21.4 yrs.) had rated their familiarity with these jobs on a 5-point scale, ranging from $1 = I$ *am not at all familiar with the content of this job* to $5 = I$ *am very familiar with the content of this job*. Results showed that students were relatively familiar with these jobs: accountant ($M = 3.40$; $SD = .91$), executive secretary ($M = 3.47$; $SD = .74$), and sales manager ($M = 3.20$; $SD = .86$).

Participants were randomly assigned to one of the following conditions. In the *first condition ("competency modeling approach")*, they received only a description of the business and HR strategy of the company (e.g., core values of the company). This description originated from an actual HR report of an organization. In the *second condition ("task-based approach"),* participants received detailed information about the tasks related to each of the three jobs. No information about the HR strategy was provided. The *third condition ("blended approach")* was a combination of Conditions 1 and 2. Participants received detailed information about both the tasks associated with the jobs and the business and HR strategy. Afterwards, all participants were instructed to determine independently the competencies of the three jobs using the card sorting method. The sorting and rating of the jobs lasted for approximately $2^{1}/_{2}$ hours.

The Portfolio Sort Cards of the LEADERSHIP ARCHITECT® (Lominger Limited) consist of 67 cards each describing a competency according to behaviorally anchored definitions (Lombardo & Eichinger, 2003). The Portfolio Sort Cards are a Q-sort method in which SMEs sort 67 cards (competencies) in 5 rating categories: $1 =$ *essential for success*, $2 =$ *very important or necessary*, $3 =$ *nice to have*, $4 =$ *less important*, and $5 =$ *not important*. To reduce rating inflation, the Portfolio Sort Cards limits the number of cards that can be sorted in each category. SMEs are required to provide 6 times a rating of 1, 16 times a rating of 2, 23 times a rating of 3, 16 times a rating of 4, and 6 times a rating of 5. Given this forced distribution, the competency ratings assigned by a rater to a job represent a set of ipsative scores. However, the dependency between the competency ratings was very small, as illustrated by the average correlation between ratings, which can be estimated as $1/(67 - 1) = .015$ (Clemans, 1966; Greer & Dunlap, 1997). Nevertheless, we employed a data-analytic approach that took ipsativity into account (VanLeeuwen & Mandabach, 2002).

We chose the Portfolio Sort Cards because it is a commercially available method for competency modeling employed by organizations (see Tett, Guterman, Bleier, & Murphy, 2000). In addition, the Portfolio Sort Cards converge closely with the features typically associated with competency determination approaches as described by Schippmann et al. (2000). For example, the Portfolio Sort Cards focus on just one type of descriptor (i.e., competencies) and operationalize it with broad labels consisting of narrative definitions (cf. behavioral anchors). Another similarity with typical competency modeling approaches outlined by Schippmann et al. is that data are collected from a number of content experts. Finally, Schippmann et al. characterized the typical protocol for determining competencies as a semistructured one. The Q-sort method employed by the Portfolio Sort Cards exemplifies such a semistructured protocol.

Analyses

We employed generalizability analysis (Brennan, 1992; Cronbach, Gleser, Nanda, & Rajaratnam, 1972) to understand the sources of variance in competency ratings. A key advantage of generalizability theory, in contrast to classical test theory, is that measurement error is regarded as *multi*faceted. Whereas classical reliability theory distinguishes only between true and error variance, generalizability theory permits the simultaneous estimation of various sources of variance. In our studies, generalizability analysis was used to estimate simultaneously the following sources of variance: competencies, jobs, raters, and their interactions. As a second advantage of generalizability analysis, the variance components can be used to estimate a generalizability coefficient, which is an intraclass correlation defined as the ratio of the universe score variance to the expected observed score variance (Brennan, 1992). This coefficient is similar to the classical reliability coefficient, although it is more accurate because multiple sources of error are taken into account. Finally, generalizability analysis allows projecting reliability estimates under different measurement conditions (Greguras & Robie, 1998). For instance, one might examine whether the reliability of competency ratings would increase when more raters are used, enabling the making of prescriptions regarding ideal measurement conditions.

Prior to a generalizability analysis, the researcher typically specifies the object of measurement (i.e., universe score) and the factors (so-called facets) affecting the measurement process (Brennan, 1992). When considered in the context of this study, the variance due to raters is seen as undesirable variance (see also Dierdorff & Wilson, 2003). A large variance component due to raters suggests substantial variation in competency ratings across raters and, therefore, is indicative of low interrater

reliability. Conversely, variance due to competencies and variance due to jobs are desirable sources of variance because they indicate discriminant validity across competencies and jobs. As it is not possible to compute a generalizability coefficient with two objects of measurement in the analysis, we followed the same strategy used in prior generalizability studies in other domains (Greguras & Robie, 1998; Greguras, Robie, Schleicher, & Goff, 2003) and conducted within-competency generalizability analyses as well as within-job generalizability analyses. In the within-competency generalizability analyses, jobs served as the object of measurement and raters as the facet. In the within-job generalizability analyses, competencies served as the object of measurement and raters as the facet.

In the within-competency generalizability analyses, the ipsative nature of the competency ratings was of no consequence because each analysis involved only one competency. However, the within-job generalizability analyses involved the full set of competency ratings that are ipsative by design and, therefore, not fully independent. To account for this dependency, we replaced the usual computational approach employed for generalizability analyses by a procedure proposed by VanLeeuwen and Mandabach (2002). For each within-job analysis, this procedure resulted in a corrected estimate of the universe score variance, which corresponded to the corrected variance component associated with the object of measurement (i.e., the competencies) and in a corrected estimate of the relative error variance, also referred to as the error variance associated with a relative decision. The generalizability coefficient was then determined in the usual way as the ratio between the estimate of the universe score variance and the sum of this estimate and the estimated relative error variance.

Results and Discussion

We first conducted the within-competency generalizability analyses. A total of 67 generalizability analyses were conducted, one for each competency. Table 1 presents the results collapsed across competencies, but broken down by condition. The values in the body of the table are the variance components and the mean percentages of variance explained by jobs, raters, and their interaction across analyses. As shown in Table 1, the blended approach produced on average the least variability among raters (18.94%), followed by the task-based approach (21.64%), and the competency modeling approach (22.20%). The opposite trend was apparent for the variance due to jobs. As explained before, rater variance was considered undesirable because it represents unreliability, whereas job variance was considered desirable because it represents discriminant validity. There

TABLE 1

Summary of Within-Competency Generalizability Analyses of Study 1

Effect	Competency modeling		Task-based approach		Blended approach		F	p
	VC	%	VC	%	VC	%		
Raters	.26	22.20%	.26	21.64%	.21	18.94%	.54	.59
Jobs	.12	12.85%	.13	13.95%	.61	16.79%	.93	.40
Jobs × Raters	.58	64.95%	.61	64.41%	.57	64.27%	.68	.51
G-coefficient (13 raters)	.72		.74		.75		1.20	.30
G-coefficient (12 raters)	.70		.72		.74		–	–
G-coefficient (11 raters)	.69		.71		.72		–	–
G-coefficient (10 raters)	.67		.69		.71		–	–
G-coefficient (9 raters)	.65		.67		.69		–	–
G-coefficient (8 raters)	.63		.65		.67		–	–
G-coefficient (7 raters)	.60		.62		.64		–	–
G-coefficient (6 raters)	.57		.59		.61		–	–
G-coefficient (5 raters)	.54		.55		.57		–	–
G-coefficient (4 raters)	.49		.51		.52		–	–

Notes. $N = 67$ as there are 67 competencies. All values in body of table are averages across generalizability analyses conducted within competency. ANOVAs were conducted on the variance components. Dashes indicate that no ANOVAs were conducted because these generalizability coefficients were projected estimates. VC = Variance component.

were virtually no differences among conditions in the variance explained by the interaction between jobs and raters.

To examine whether these differences were statistically significant, we conducted a MANOVA using the variance components due to raters, jobs, and their interaction as dependent variables and condition as the independent variable. No significant multivariate main effect emerged, $F(6, 392) = .61$, *ns*, Wilks lambda = .98. Follow-up ANOVAs per dependent variable also failed to yield significant differences among conditions.

An inspection of the generalizability coefficients revealed the highest value in the blended approach (.75), followed by the task-based (.74), and competency modeling (.72) approaches. At first sight, these values appear acceptable. However, these results reflect generalizability over 13 raters (recall there were 13 raters per condition). Given that so many raters may not be available in every application, we projected the generalizability coefficient under different sets of measurement conditions (i.e., different number of raters, see lower part of Table 1). For instance, in our experience, practitioners often have access to no more than four SMEs (one supervisor, one job analyst, and two job incumbents) per job. Table 1 shows that when four raters were used, the projected generalizability coefficients barely reached .50.

Whereas all prior analyses were conducted within competencies, we also conducted within-job generalizability analyses. The results were identical to those in Table 1.

Two important conclusions follow from Study 1. First, the provision of task information to student raters did not produce beneficial effects in terms of increasing interrater reliability and discriminant validity among jobs. Although the blended approach performed slightly better than the task-based approach, which in turn fared better than the competency modeling approach, no statistically significant differences among the three conditions emerged. A second important conclusion is that Study 1 showed that regardless of the work analysis approach, the inferential leap from jobs to competencies might have been too large for the student raters. Indeed, generalizability coefficients would never surpass the .50s if only four student raters were used. Therefore, a practical implication of Study 1 is that practitioners interested in competency modeling should be cautious about using naïve raters in their SME panels. As noted by Schippmann et al. (2000), raters barely familiar with the job are sometimes used in competency modeling.

The generalizability of our results might be weakened by the use of student raters. As these students lacked the organizational context, they probably had limited interest in making ratings that had no real impact on HR applications. In addition, given that the jobs targeted (e.g., sales manager, accountant) were relatively common, it is possible that the students were influenced by shared job stereotypes that might have overshadowed the information provided in the various conditions, thereby reducing differences across conditions.

Given the limitations inherent in the use of student samples and hypothetical jobs, the next two studies examined competency modeling in an actual organization as carried out by a diverse group of SMEs (i.e., incumbents, supervisors, HR specialists, and internal customers). In addition, these SMEs' competency inferences had real impact because they affected training and development interventions in the organization.

Study 2

Method

In Study 2, we examined the quality of inferences in competency modeling among a diverse group of SMEs (incumbents, supervisors, HR specialists, and internal customers) in a multinational company producing specialty materials. Three jobs were selected because the organization had expressed the need to determine the competencies of these jobs as input

TABLE 2

Summary of Within-Competency Generalizability Analyses of Study 2

Effect	VC	% of variance explained
Raters	.15	15.84%
Jobs	.20	19.30%
Jobs × Raters	.59	64.85%
G-coefficient (4 raters)	.62	

Notes. $N = 67$ as there are 67 competencies. All values in body of table are averages across generalizability analyses conducted within competency. VC = Variance component.

for future training and development plans. These jobs were design and manufacturing engineer (translates orders into production process specifications such as blueprints, materials, machines needed, and standards), technical production operator (handles machines, materials, and tools by studying blueprints, selecting relevant actions, and verifying the operations to determine whether standards were met), and management accountant (conducts financial analyses and provides decisionmakers with this information).

The Portfolio Sort Cards were used to determine job competencies. A SME panel was assembled for each of the three jobs involved. Each SME panel consisted of one representative of the following four information sources: job incumbents (two men, one woman; mean age = 28.8 years; mean tenure in the organization = 3.5 years), supervisors (all men; mean age = 35.7 years; mean tenure = 5.8 years), HR specialists (two men, one woman; mean age = 33.4 years; mean tenure = 6.7 years), and internal customers (colleagues; all men; mean age = 42.4 years; mean tenure = 17.4 years). Familiarity with the focal job was the primary selection criterion for panel membership. All SMEs were also knowledgeable about the business and the HR strategies of the organization and had completed a $\frac{1}{2}$ day training session that familiarized them with the Portfolio Sort Cards. This training session explained the 67 competencies, their behaviorally anchored definitions, and the Q-sort method used. At the end of the training session, all SMEs received a manual and a set of competency sort cards.

Results and Discussion

Table 2 presents the results of the within-competency generalizability analyses. Raters explained 15.84% of the variance. This figure is lower than the corresponding one in Study 1 (i.e., 22.20%), suggesting there is less rater variability and therefore higher interrater reliability among experienced SMEs. In a similar vein, jobs accounted for 19.30% of the

TABLE 3

Summary of Within-Job Generalizability Analyses of Study 2

Effect	VC
Job 1: Design and manufacturing engineer	
Competencies	.47[a]
Error	.18[b]
G-coefficient	.72
Job 2: Technical production operator	
Competencies	.34[a]
Error	.22[b]
G-coefficient	.61
Job 3: Management accountant	
Competencies	.70[a]
Error	.12[b]
G-coefficient	.85

[a]Variance component.
[b]Relative error variance (see VanLeeuwen & Mandabach, 2002).

variance, which is higher than the corresponding 12.85% in Study 1. This finding suggests that experienced SMEs are better able to discriminate between different competencies across jobs than student raters are. The generalizability coefficient for the four rater types was .62, which is also higher than the corresponding generalizability coefficient of four student raters in Study 1.

It would be worthwhile to know which one of the four types of raters (incumbents, supervisors, HR specialists, and internal customers) provided the most reliable and differentiated ratings. However, such within-rater type analyses were not possible because we had only one rater per source. However, our data permitted an examination of differences among rater types. To this end, we repeated the previous generalizability analyses four times, excluding one of the rater types each time. For instance, we ran a generalizability analysis including incumbents, supervisors, and HR specialists, but excluding internal customers. Interestingly, the variance component associated with type of rater dropped only when internal customers were left out. This finding suggested that the ratings of internal customers were most different from the ratings of the other sources. Therefore, when the perspective of internal customers is not considered important, their exclusion may facilitate a higher interrater reliability.

The results of the within-job generalizability analyses are shown in Table 3. Because we conducted these analyses according to the procedure proposed by VanLeeuwen and Mandabach (2002), the presentation of the results differs somewhat from the presentation of our other analyses.

In particular, Table 3 indicates the estimate of the variance component of the measurement object (i.e., the competencies), the estimate of the relative error variance, and the resulting generalizability coefficient, broken down by job. There were substantial differences in the generalizability coefficients among jobs. For instance, when the four types of raters rated 67 competencies for the job of management accountant, the generalizability coefficient was .85. However, the generalizability coefficient was only .61 for the technical production operator job. The job of design and manufacturing engineer produced a generalizability coefficient of .72.

Three noteworthy conclusions follow from a comparison between Study 1 and Study 2. First, the competency modeling approach yielded more acceptable levels of interrater reliability when the SME panel consisted of job incumbents, supervisors, internal customers, and HR specialists than when it consisted of student raters like those employed in Study 1. Second, experienced SMEs seemed better able to discriminate among the relative importance of each competency for each job than student raters in Study 1 were. Finally, competency modeling seemed to work better for some jobs than for others because within-job generalizability coefficients varied considerably, with the lowest value found for the job of technical production operator and the highest value obtained for the job of management accountant. Perhaps the Portfolio Sort Cards are better suited for describing certain jobs (i.e., managerial jobs) than others (i.e., entry-level jobs). However, the Portfolio Sort Cards consist of a large number (67) of competencies in order to increase their applicability to a wide variety of jobs. Still another explanation for the lower generalizability coefficient of the production operator is that this job might have a slightly different content across departments and employees (see also Borman, Dorsey, & Ackerman, 1992). In fact, many employees across different departments were working as production operators for this organization. Conversely, the job of management accountant was a newly defined job and there was only one management accountant in the organization.

In general, the results of Study 2 are somewhat more encouraging for competency modeling than those of Study 1 because the use of a diverse panel of SMEs increased the quality of inferences. Study 1 and Study 2 employed different mechanisms (i.e., task information in Study 1 and use of different types of SMEs in Study 2) for increasing the quality of inferences. However, given the limitations of a lab study (Study 1), as well as the fact that practitioners may choose not just one but the two mechanisms, it was deemed appropriate to examine their combined effects. For this reason, Study 3 compared the quality of inferences made in a blended versus a competency modeling approach using a diverse panel of SMEs in a real organizational setting.

Method

The data were gathered at the same multinational company employed in Study 2. To identify competencies, a SME panel was assembled consisting of one representative from the following four information sources: job incumbents (3 men, 1 woman; mean age = 39.2 years; mean tenure in the organization = 14.5 years), supervisors (all men; mean age = 39.1 years; mean tenure = 3.7 years), HR specialists (3 men, 1 woman; mean age = 30.7 years; mean tenure = 4.11 years), and internal customers (colleagues; 3 men, 1 woman; mean age = 36.9 years; mean tenure = 4.8 years). Familiarity with the focal job served as the primary selection criterion for these panels. Similar to Study 1, all SMEs had completed a $^1/_2$ day training session in which they became familiar with the Portfolio Sort Cards. Note that the SMEs of Study 2 were different from the SMEs in Study 3.

Similar to our previous studies, the SMEs used the Portfolio Sort Cards to independently determine the competencies. In the first condition (i.e., "competency modeling approach"), SMEs received a description of the business and the HR strategies of the company. In the second condition ("blended approach"), SMEs did not only receive a description of the business and the HR strategies but also detailed information regarding the tasks performed on each job. These tasks were previously defined by a job analyst who interviewed SMEs not included in this study.

Recall that Study 2 had shown that the focal job might affect the quality of inferences drawn. Therefore, the ideal design of Study 3 would have asked for ratings of the exact same set of jobs using the competency modeling and the blended approach. Although this design might be feasible in a laboratory experiment, it was not possible in this field setting, where highly priced SME time was not to be spent on redundant panels. Therefore, we tried to ensure that the (dis)similarity of the two jobs chosen per condition was equivalent. To this end, HR specialists in the organization were asked to judge the similarity of various jobs for which the competencies still had to be determined. Another criterion for inclusion was that only jobs that were relatively stable were considered. This process led to two sets of jobs. A first set consisted of the jobs of technical sales coordinator (coordinates the sales process starting from the initial order until possible after-sales complaints) and technical services manager (plans, coordinates, and monitors new investments projects by directing project engineers). Although both jobs had a technical component, the latter also included a managerial component. Therefore, these jobs were seen as relatively dissimilar. The second set of jobs consisted of the jobs of maintenance technician (conducts reparatory and preventive maintenance activities on machines and

TABLE 4

Summary of Within-Competency Generalizability Analyses of Study 3

Effect	Competency modeling		Blended approach			
	VC	%	VC	%	*t*	*p*
Raters	.23	23.51%	.10	10.15%	3.38	.00
Jobs	.14	12.78%	.42	26.19%	−3.02	.00
Jobs × Raters	.60	63.71%	.62	63.66%	.09	.99
G-coefficient (four raters)	.68		.74		−2.05	.04

Notes. $N = 67$ as there are 67 competencies. All values in body of table are averages across generalizability analyses conducted within competency. *T* tests were conducted on the variance components. VC = Variance component.

production facilities) and technical production supervisor (plans, coordinates, and monitors production activities by directing technical operators). Again, both of these jobs were technical in nature. However, the latter also included an important managerial component. Therefore, the two jobs in this second set were also seen as relatively dissimilar. We also examined whether the (dis)similarity of the two jobs chosen per condition was equivalent. To this end, we counted how many of the 10 competencies in the final[1] competency profiles overlapped. We found that the competency profiles of the technical sales coordinator and technical services manager shared 4 competencies. Exactly the same overlap (4 competencies) was found for the jobs of technical production supervisor and maintenance technician. Together, these comparisons attested to the fact that the two sets of jobs chosen in each condition were relatively equivalent. The technical sales coordinator and technical services manager were randomly assigned to the competency modeling condition, leaving the technical production supervisor and maintenance technician in the blended approach condition.

Results and Discussion

Table 4 shows the results of the within-competency generalizability analyses. The blended approach outperformed the competency modeling approach in terms of reducing rater variability (10.15% vs. 23.51%)

---

[1]The final competency profiles were the ones used by the organization to derive training plans. These final competency profiles were determined by averaging the competency ratings across SMEs. When the ratings differed considerably, discrepancies were discussed in a meeting among SMEs. For practicality, only the 10 most important competencies per job were kept in the final competency profile.

TABLE 5

Summary of Within-Job Generalizability Analyses of Study 3

| | Competency modeling | |
	Manager technical services	Technical sales coordinator
Competencies	.46[a]	.43[a]
Error	.19[b]	.19[b]
G-coefficient	.71	.69

| | Blended approach | |
	Technical production supervisor	Maintenance technician
Competencies	.56[a]	.63[a]
Error	.16[b]	.14[b]
G-coefficient	.78	.82

[a]Variance component.
[b]Relative error variance (see VanLeeuwen & Mandabach, 2002).

and increasing discriminant validity among jobs (26.19% vs. 12.78%). To examine whether these differences were statistically significant, we conducted a MANOVA using the variance components due to raters, jobs, and their interaction as dependent variables, and condition as the independent variable. A significant multivariate main effect was found, $F(3, 130) = 6.39, p < .001$, Wilks lambda $= .87$, and eta$^2 = .13$. Followup t-tests for each dependent variable also showed significant differences across conditions for the variance component related to jobs as well as for the variance component related to raters. In addition, there was a statistically significant difference in generalizability coefficients, with the blended approach producing the highest coefficient across the four types of raters (.73).

Table 5 presents the results of the within-job generalizability analyses obtained through the method advocated by VanLeeuwen and Mandabach (2002). These results supported the findings previously reported in Table 4 because the blended approach led to the largest increases in interrater reliability and discriminant validity. For the two jobs, the generalizability coefficients in the competency modeling approach were lower than the coefficients in the blended approach (.69 and .71 vs. .78 and .82, respectively).

The results depicted in Table 5 are important in light of the possible confound between jobs and conditions. Although we undertook substantial efforts to ensure that the (dis)similarity between the two jobs chosen was equal per condition (see above), some might argue that the two jobs in the blended approach were more dissimilar from each other than the two jobs in the competency modeling approach. However, Table 5 argues

against this alternate explanation because, for all jobs, the variance due to competencies in the blended approach was always higher than the variance due to competencies in the competency modeling approach.

General Discussion

A spirited debate about competency modeling has erupted in recent times. A central aspect of this debate has been the questionable quality and job-relatedness of the broad inferences required by competency modeling. This study is among the first to empirically examine the quality of inferences made in competency modeling, along with potential improvements of the methodology. We used generalizability analysis as a systematic tool for examining the quality of competency inferences. A broad conclusion that can be drawn from our data is that the quality of inferences in competency modeling should not be taken for granted. A second general conclusion is that at least two procedural factors (provision of task information and use of a variety of job experts) can enhance the quality of the inferences drawn in competency modeling. We discuss each of these conclusions together with their practical implications in further detail in the next sections.

Task-related Information

Study 3 revealed that blending a task analysis approach with a competency-based approach might enhance the quality of the inferences that job experts draw about competency requirements. This result is not only important for competency modeling but also for other approaches that directly estimate KSAOs from a job description (e.g., Primoff, 1975). When organizations combine two types of descriptor information (i.e., tasks from the task analysis approach with strategy descriptions from the competency modeling approach), they seemingly capitalize on the strengths inherent in either approach. Task information appears to reduce the inferential leaps required from SMEs and provide a common frame of behavioral referents, thereby enhancing the reliability of SMEs' inferences and their ability to discriminate among jobs. Besides, SMEs are also reminded of the HR strategy prior to inferring competencies. The superiority of the blended approach evidenced here suggests that behaviorally based work analysis and strategic thinking about job specifications are not mutually exclusive (Schippmann et al., 2000).

The gains afforded by the blended approach over the competency modeling approach are illustrated by comparing the number of raters needed to ensure acceptable levels of interrater reliability in each. Specifically, in Study 3, four raters using the blended approach are sufficient to obtain a

generalizability coefficient of .74, whereas at least six raters are needed to achieve a similar coefficient in the competency modeling approach. Thus, the additional time, material, and labor resources demanded by the combination of task-based and strategic information in the blended approach appear to provide a significant return on investment, at least in the form of increased measurement quality of the inferences.

Generally, our results concur with Schippmann et al. (2000), who argued for increasing the methodological rigor in competency modeling while at the same time preserving its hallmarks (i.e., creating a nexus between an organization's strategy and the job specifications of strategically aligned HR practices). It appears that such methodological improvements may borrow from traditional job analyses, which supply helpful "rating aids" such as the opportunity to study job tasks when judging the relevance of competencies. Future studies are warranted to identify other design recommendations that might increase the quality of competency modeling endeavors. For example, it would be interesting to examine whether the quality of inferences increases when raters are given not only a list of task statements, but are also required to make task-competency linkages beforehand. In fact, some authors have proposed the use of a two-way matrix (job task × worker attributes) to facilitate determinations of underlying worker requirements (Drauden & Peterson, 1974; Guion, 1980). Although introducing this rigor in competency modeling makes this technique less flexible and moves it closer to traditional job analysis, rigor might pay off by increasing the quality and legal defensibility of the final output.

Use of a Variety of Job Experts

Our findings suggested that organizations interested in obtaining acceptable levels of reliability in competency modeling should carefully select their raters among a pool of experienced SMEs representing different views or "takes" on the jobs (e.g., incumbents, supervisors, internal customers). For instance, a comparison between the results from Study 1 and those from Studies 2 and 3 hints that rating quality suffers when naïve and inexperienced raters are not sufficiently familiar with the job. Even though the student raters in Study 1 were carefully trained on the competency modeling technique and were also familiar with the target jobs, there was substantial disagreement on competency inferences from rater to rater. Conversely, there was less variation in competency ratings across experienced SMEs in Studies 2 and 3. These results are consistent with those from the traditional job analysis literature, which had revealed considerably lower reliabilities for naïve raters than for experienced raters

(Friedman & Harvey, 1986; Harvey & Lozada-Larsen, 1988; Voskuijl & Van Sliedregt, 2002).

A related conclusion is that organizations interested in competency modeling should have a sufficient number of raters in their panel. Having only one or two raters in competency modeling is unlikely to yield inferences that can sustain legal challenges, although this caveat seems to be sometimes ignored (Schippmann et al., 2000). Across the studies, having at least four raters appeared to be the minimum to achieve acceptable levels of interrater reliability. This figure is supported by prior research on the related domain of job evaluation, where Doverspike, Carlisi, Barrett, and Alexander (1983) found that substantial gains in reliability could be expected with an increase from one to four raters (see also Fraser, Cronshaw, & Alexander, 1984).

Finally, this study presents preliminary evidence regarding the effects of mixing different types of raters in competency modeling. Study 2 revealed that internal customers differed the most from the other three rater types (incumbents, supervisors, and HR specialists), suggesting that internal customers hold a somewhat different take on the job than more traditional types of raters (Brannick & Levine, 2002).

Future research should scrutinize this finding more closely. In particular, if multiple raters per rater type are available, within-rater source generalizability analyses should be conducted to determine which type of raters provides the most reliable and differentiated ratings.

Methodological Implications

The most recent version of the *Principles for the Validation and Use of Personnel Selection Procedures* (Society of Industrial and Organizational Psychology, 2003) makes an explicit reference to competency modeling, stating that any work analysis method should have reasonable psychometric characteristics. From a methodological point of view, Studies 2 and 3 underscore the need to document that competency inferences have adequate measurement properties because our data suggest large variability in SMEs' competency ratings as a function of the type and circumstances surrounding the focal jobs.

Another methodological contribution lies in illustrating that generalizability analysis is a useful tool for examining the reliability of competency ratings. Even in job analysis and job evaluation, generalizability analysis has been scarcely used (Doverspike et al., 1983; Fraser et al., 1984). This absence is unfortunate because generalizability analysis enables the estimation of multiple sources of variance and the projection of reliability under different measurement conditions.

Limitations

The generalizability of our results is bounded by our focus on one organization and one competency modeling technique. Although our competency modeling condition mirrored the general procedure advocated by the proponents of this approach, future studies are needed on different competency modeling techniques. For example, our studies do not examine "generative" competency modeling procedures such as those in which the organization starts with a blank slate and has yet to define a set of competencies. When considering the reliability of such unstructured procedures to competency modeling, the reliability and discriminant validity figures obtained in this study might be best seen as upper-bound estimates.

A caveat is in order regarding the two dependent variables employed (i.e., interrater reliability and discriminability among jobs and/or competencies). As argued by Morgeson and Campion (1997), these criteria reflect important issues of reliability and discriminant validity in work analysis data. However, they capture only part of the picture. This criticism, however, is not unique to this study but applies to work analysis research in general (Sanchez & Levine, 2000). On a broader level, organizations should always take the purpose of work analysis into account when deciding upon the dependent variables to be used (Sackett & Laczo, 2003).

A final limitation is the relatively small sample sizes in our studies. However, it is important to remember that samples like the ones we employed involve SMEs whose time is precious to the organization. In fact, it can be argued that our sample sizes are representative of the typical SME panel size in the field.

Conclusion

Although competency modeling has made rapid inroads in many organizations, research has lagged behind. Our three studies represent a first step in closing this practice–research gap. Our investigation was predicated on the assumption that competency modeling and job analysis need not be mutually exclusive. On the contrary, we maintain that there is plenty of room for a fruitful cross fertilization between the two domains. First, competency modeling confers strategic alignment to work analysis by incorporating the organization's strategy into the derivation of the attributes. Second, we found that the quality of the inferences made in competency modeling can be enhanced by borrowing from traditional job-analytic methodology. Specifically, our studies identified design considerations (i.e., provision of task information, composition of SME panels, number of raters) that might make a key difference in terms of the quality of the

inferences drawn in competency modeling. We hope that our studies begin to draw the roadmap towards a rigorous methodology capable of turning the "art" of competency modeling into a truly scientific endeavor (Lucia & Lepsinger, 1999).

REFERENCES

Barrett GV, Callahan CM. (1997, April). Competencies: The Madison Avenue approach to professional practice. In Page RC (Chair), *Competency models: What are they and do they work?* Practitioner Forum presented at the 12th Annual Conference of the Society for Industrial and Organizational Psychology, St. Louis, Missouri.

Barrett GV, Depinet R. (1991). A reconsideration of testing for competence rather than intelligence. *American Psychologist, 46*, 1012–1024.

Becker BE, Huselid MA, Ulrich D. (2001). *The HR scorecard. Linking people, strategy, and performance*. Boston: Harvard Business School Press.

Borman WC, Dorsey D, Ackerman L. (1992). Time-spent responses as time allocation strategies: Relations with sales performance in a stockbroker sample. PERSONNEL PSYCHOLOGY*, 45*, 763–777.

Brannick MT, Levine EL. (2002). *Job analysis*. Thousand Oaks, CA: Sage.

Brennan RL. (1992). *Elements of generalizability theory*. Iowa City, IA: American College Testing Program.

Butler SK, Harvey RJ. (1988). A comparison of holistic versus decomposed ratings of Position Analysis Questionnaire work dimensions. PERSONNEL PSYCHOLOGY*, 37*, 453–464.

Christal TE. (1974). The U.S. Air Force occupational research projects. *JSAS Catalog of Selected Documents in Psychology, 4*, 61.

Clemans WV. (1966). An analytical and empirical examination of some properties of ipsative measures. *Psychometric Monographs, 14*.

Cornelius ET, DeNisi AS, Blencoe AG. (1984). Expert and naïve raters using the PAQ: Does it matter? PERSONNEL PSYCHOLOGY*, 41*, 761–771.

Cornelius ET, Lyness KS. (1980). A comparison of holistic and decomposed judgment strategies in job analysis by job incumbents. *Journal of Applied Psychology, 65*, 155–163.

Cronbach LJ, Gleser GC, Nanda H, Rajaratnam N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.

Dierdorff EC, Wilson MA. (2003). A meta-analysis of job analysis reliability. *Journal of Applied Psychology, 88*, 635–646.

Doverspike D, Carlisi AM, Barrett GV, Alexander RA. (1983). Generalizability analysis of a point-method job evaluation instrument. *Journal of Applied Psychology, 68*, 476–483.

Drauden GM, Peterson NG. (1974). *A domain approach to job analysis*. St. Paul, MN: Minnesota Department of Personnel Test Research and Development Section.

Fine SA. (1988). Functional job analysis. In Gael S, Cornelius III ET, Levine EL, Salvendry G (Eds.), *The job analysis handbook for business, industry, and government*, *Vol. II* (pp. 1019–1035). New York: Wiley.

Fraser SL, Cronshaw SF, Alexander RA. (1984). Generalizability analysis of a point-method job evaluation instrument: A field study. *Journal of Applied Psychology, 69*, 643–647.

Friedman L, Harvey RJ. (1986). Can raters with reduced job descriptive information provide accurate position analysis questionnaire (PAQ) ratings? PERSONNEL PSYCHOLOGY*, 39*, 779–789.

Gatewood RD, Field HS. (2001). *Human Resource Selection* (5th ed.). Orlando, FL: Harcourt.

Gerhart B. (1988). Sources of variance in incumbent perceptions of job complexity. *Journal of Applied Psychology, 73*, 154–162.

Greer T, Dunlap WP. (1997). Analysis of variance with ipsative measures. *Psychological Methods, 2*, 200–207.

Greguras GJ, Robie C. (1998). A new look at within-source interrater reliability of 360-degree feedback ratings. *Journal of Applied Psychology, 83*, 960–968.

Greguras GJ, Robie C, Schleicher DJ, Goff MG III. (2003). A field study of the effects of rating purpose on the quality of multisource ratings. PERSONNEL PSYCHOLOGY, *56*, 1–21.

Guion RM. (1980, May). *Job analysis and validation procedures: Applications to construct identification*. Third Annual Conference on EEO Compliance and Human Resource Utilization, Chicago, IL.

Hahn DC, Dipboye RL. (1988). Effects of training and information on the accuracy and reliability of job evaluations. *Journal of Applied Psychology, 73*, 146–153.

Harvey RJ, Lozada-Larsen SR. (1988). Influence of amount of job descriptive information on job analysis rating accuracy. *Journal of Applied Psychology, 73*, 457–461.

Holland JH, Holyoak KJ, Nisbett RE, Thagard PR. (1986). *Induction: Processes of inference, learning, and discovery*. Cambridge, MA: The MIT Press.

Hubbard M, McCloy R, Campbell J, Nottingham J, Lewis P, Rivkin D. (1999). *Revision of O*NET data collection instruments*. Raleigh, NC: National Center for O*NET Development.

Hughes GL, Prien EP. (1989). Evaluation of task and job skill linkage judgments used to develop test specifications. PERSONNEL PSYCHOLOGY, *42*, 283–292.

Lawler E. (1996). Competencies: A poor foundation for the new pay. *Compensation and Benefits Review, 28*(6), 20–22.

Lombardo M, Eichinger R. (2003). *The LEADERSHIP ARCHITECT® norms and validity report*. Minneapolis: Lominger Limited, Inc.

Lopez FM, Kesselman GA, Lopez EE. (1981). An empirical test of a trait-oriented job analysis technique. PERSONNEL PSYCHOLOGY, *34*, 479–502.

Lucia AD, Lepsinger R. (1999). *The art and science of competency models*. San Francisco: Jossey-Bass.

Morgeson FP, Campion MA. (1997). Social and cognitive sources of potential inaccuracy in job analysis. *Journal of Applied Psychology, 79*, 475–480.

Morgeson FP, Campion MA. (2000). Accuracy in job analysis: Toward an inference-based model. *Journal of Organizational Behavior, 21*, 819–827.

Morgeson FP, Delaney-Klinger KA, Mayfield MS, Ferrara P, Campion MA. (in press). Self-presentation processes in job analysis: A field experiment investigating inflation in abilities, tasks, and competencies. *Journal of Applied Psychology*.

Pearlman K. (1997, April). Competencies: Issues in their application. In Page RC (Chair), *Competency Models: What are they and do they work*? Practitioner Forum presented at the 12th Annual Conference of the Society for Industrial and Organizational Psychology, St. Louis, Missouri.

Prien EP, Prien KO, Gamble LG. (2004). Perspectives on nonconventional job analysis methodologies. *Journal of Business and Psychology, 18*, 337–348.

Primoff ES. (1975). *How to prepare and conduct job-element examinations*. Washington, D.C.: U.S. Government Printing Office.

Sackett PR, Laczo RM. (2003). Job and work analysis. In Borman WC, Ilgen DR, Klimoski RJ (Eds.), *Handbook of psychology* (Vol. 12, pp. 21–37). Hoboken, NJ: Wiley.

Sanchez JI. (2000). Adapting work analysis to a fast-paced and electronic business world. *International Journal of Selection and Assessment, 8*, 207–215.

Sanchez JI, Levine EL. (1989). Determining important tasks within jobs: A policy capturing approach. *Journal of Applied Psychology, 74*, 336–342.

Sanchez JI, Levine EL. (1994). The impact of raters' cognition on judgment accuracy: An extension to the job analysis domain. *Journal of Business and Psychology, 9*, 47–57.

Sanchez JI, Levine EL. (2000). Accuracy or consequential validity: Which is the better standard for job analysis data? *Journal of Organizational Behavior, 21*, 809–818.

Sanchez JI, Levine EL. (2001). The analysis of work in the 20th and 21st centuries. In Anderson N, Ones D, Sinangil HK, Viswesvaran C (Eds.), *Handbook of industrial, work, and organizational psychology: Personnel psychology* (Vol. 1, pp. 71–89). London/New York: Sage.

Sanchez JI, Zamora A, Viswesvaran C. (1997). Moderators of agreement between incumbent and non-incumbent ratings of job characteristics. *Journal of Occupational and Organizational Psychology, 70*, 209–218.

Schippmann JS. (1999). *Strategic job modeling*. Mahwah, NJ: Erlbaum.

Schippmann JS, Ash RA, Battista M, Carr L, Eyde LD, Hesketh B, et al. (2000). The practice of competency modeling. *Personnel Psychology, 53*, 703–740.

Society of Industrial and Organizational Psychology, Inc. (2003). *Principles for the validation and use of personnel selection procedures* (3rd ed.). College Park, MD: Author.

Spector PE, Jex SM. (1991). Relations of job characteristics from multiple data sources with employee affect, absence, turnover intentions, and health. *Journal of Applied Psychology, 76*, 46–53.

Srull TK, Wyer RS. (1980). Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments. *Journal of Personality and Social Psychology, 38*, 841–856.

Sulsky LM, Day DV. (1994). Effects of frame-of-reference training on rater accuracy under alternative time delays. *Journal of Applied Psychology, 79*, 535–543.

Tett RP, Guterman HA, Bleier A, Murphy PJ. (2000). Development and content validation of a "hyperdimensional" taxonomy of managerial competence. *Human Performance, 13*, 205–251.

Uniform Guidelines on Employee Selection Procedures. (1978). Federal Register, 1978, 43, No. 166, 38290–38309.

VanLeeuwen DM, Mandabach KM. (2002). A note on the reliability of ranked items. *Sociological Methods and Research, 31*, 87–105.

Voskuijl OF, Van Sliedregt T. (2002). Determinants of interrater reliability of job analysis: A meta-analysis. *European Journal of Psychological Assessment, 18*, 52–62.