

Accuracy in job analysis: toward an inference-based model

FREDERICK P. MORGESON¹* AND MICHAEL A. CAMPION²

¹*Lowry Mays College & Graduate School of Business, Texas A&M University, U.S.A.*

²*Kramert Graduate School of Management, Purdue University, U.S.A.*

Summary

Although the accuracy of job analysis information is critically important, standards for accuracy are not clear. Researchers have recently begun to address various aspects of job analysis accuracy by investigating such things as potential sources of inaccuracy in job analysis as well as attempting to reconceptualize our notions of job analysis accuracy. This article adds to the debate by first discussing how job analysis accuracy has been conceptualized. This points to difficulties in the prevalent 'true score' model upon which many of these discussions have been based. We suggest that discussions of job analysis accuracy would benefit from a consideration of the validity of job analysis *inferences*, as a complement to the more traditional focus on the validity of job analysis *data*. Toward this end, we develop a model of the inferences made in the job analysis process, outline some of the ways these inference could be tested, and discuss implications of this perspective. Copyright © 2000 John Wiley & Sons, Ltd.

Introduction

The issue of accuracy in job analysis is an important one. The notion that certain kinds of job analysis may be plagued with systematic and far-reaching inaccuracies was recently discussed by Morgeson and Campion (1997). They suggested there are a host of potential social and cognitive sources of inaccuracy in job analysis, thus calling into question the often implicit assumptions that these data are free from bias and inaccuracy. Unfortunately, comparatively little job analysis research has directly investigated this issue. It is likely that one impediment to this research has been the inherent difficulties in defining accuracy in the job analysis context.

Given these difficulties, Sanchez and Levine (1999) offer a much-needed discussion of the issue of accuracy in job analysis. The entire job analysis domain has struggled with questions about what constitutes accuracy. That is, how do we know that job analysis information is accurate or 'true'? The absence of a definitive answer to this question creates problems for both researchers and practitioners. Assessing the accuracy or quality of job analysis data is critical because it forms the foundation upon which virtually every human resource system is built (e.g., selection

* Correspondence to: Frederick Morgeson, Lowry Mays College & Graduate School of Business, Department of Management, Texas A&M University, College Station, TX 77843-4221, U.S.A.
E-mail: morgeson@tamu.edu

systems, training programmes, performance management systems; Ash, 1988). In addition, the use of job analysis is recommended by legal guidelines (Uniform Guidelines, 1978) and can play an important role in defending human resource systems from legal challenge (e.g., by increasing job-relatedness; Williamson *et al.*, 1997).

Because we agree with many of the points raised by Sanchez and Levine (1999), we have chosen to expand their discussion by first describing some of the ways accuracy has been defined in the job analysis domain. This acknowledges some of the difficulties highlighted by Sanchez and Levine (1999) and points out the inherent weaknesses in existing conceptions of job analysis accuracy. As a way to address these weaknesses and change the tenor of the debate, we draw from previous construct validity work (e.g., Schwab, 1980) and the debate over construct, content, and criterion-related validity in personnel selection research (e.g., Binning and Barrett, 1989) to outline an inference-based model of job analysis validity. This model suggests we should move away from discussing the accuracy of job analysis data and toward discussing the validity of job analysis inferences. This naturally leads into a description of the kinds of inferences that are made in the job analysis context, the types of evidence that could be used to support these inferences, and the implications of an inference-based model.

Existing Conceptualizations of Job Analysis Accuracy

Classical test theory

In conceptualizing accuracy in the job analysis context, research has generally relied on the principles of classic test theory (Campion, Morgeson and Mayfield, 1999; Harvey, 1991). Applied to the job analysis domain, classical test theory suggests that a 'true score' exists for a given job, true scores are stable over time, and any measurement variation is error that can be reduced or eliminated or aggregating across time or sources (Nunnally and Bernstein, 1994). In this spirit, researchers have commonly aggregated job analysis information (across incumbents, subject matter experts, or job analysts) to arrive at the 'true score' for a particular job. The quality of job analysis data is then typically indexed with estimates of interrater reliability.

Generalizability theory

Because classical test theory can only estimate one source of error at a time, some have suggested adopting a generalizability theory perspective (Campion *et al.*, 1999; Sanchez and Levine, 2000). Generalizability theory is concerned with the dependability of behavioral measures (Cronbach *et al.*, 1972), where dependability involves the accuracy of generalizing from an observed score to the average score obtained over a wide range of situations (Shavelson and Webb, 1991). Although generalizability theory is predicated on a stable true score (like classical test theory), it allows one to segment sources of variance into multiple sources or facets (e.g., methods of data collection, sources of data).

The ability to estimate multiple sources of measurement error simultaneously is particularly useful because different facets of job analysis are subject to different sources of inaccuracy (see Morgeson and Campion, 1997, Table 2). This makes it possible to explicitly examine how

different methods, sources of data, locations, or other potentially important factors (e.g., incumbent ability, personality, or experience) impact the accuracy of job analysis data. Although not all systematic variability reflects inaccuracy (it may reflect real differences), identifying the sources of variance may provide researchers with a deeper understanding of the ways in which job analysis information can be affected.

Components of overall accuracy

Another perspective involves utilizing Cronbach's (1955) method of partitioning overall accuracy scores into four related components: elevation; differential elevation (and subcomponent elevation correlation); stereotype accuracy (and subcomponent stereotype correlation); and differential accuracy (and subcomponent differential correlation). Elevation reflects the way raters use response scales and would be a function of the differences between the average of an individual's ratings and the average expert score (frequently operationalized as the average of all job incumbents or all job analysts). Differential elevation reflects how closely an average job rating (across all dimensions) would be to an average expert rating whereas elevation correlation reflects an analyst's ability to correctly rank order jobs in terms of their overall mean rating.

Stereotype accuracy reflects the ability to predict the profile of dimension means (across jobs) whereas stereotype correlation reflects the ability to correctly rank order rating dimension averages. Finally, differential accuracy reflects the ability to predict differences between jobs on individual dimensions whereas differential correlation reflects the ability to correctly rank order jobs on each dimension (averaged across dimensions). This approach has been used by Harvey and Lozada-Larsen (1988) to examine average incumbent ratings. In an experimental study, they found that the ability to match the average expert score (elevation) and the ability to predict differences between jobs (differential accuracy) were the largest components of overall accuracy.

Multidimensional conception

Finally, Morgeson and Campion (1997) suggested that inaccuracy can be indexed in at least six different ways: interrater reliability; interrater agreement; discriminability between jobs; dimensionality of factor structures; mean ratings; and completeness of job information. Interrater reliability reflects consistency across different raters and indexes rater covariation (Shrout and Fleiss, 1979). As noted earlier, this form of reliability has been the most commonly used index of job analysis accuracy. Interrater agreement reflects the absolute level of agreement across different raters. As such, it indexes the extent to which raters make similar ratings (Kozlowski and Hattrup, 1992). Discriminability between jobs reflects the ability to distinguish between different jobs. Dimensionality of factor structures reflects the extent to which factor structures are complex or multidimensional (Stone and Gueutal, 1985). Mean ratings refer to elevated or depressed ratings. Completeness reflects the relative comprehensiveness of the job analysis data.

There are several aspects of this conception that deserve mention. First, these six different indices reflect underlying issues of reliability (e.g., reliability and agreement) and validity (e.g., discriminability, dimensionality, mean ratings, and completeness). As such, it represents a multidimensional conception of accuracy. Second, this perspective acknowledges that accuracy can be indexed in many different ways and that any single estimate may fail to adequately assess the accuracy of job analysis data. Third, it is possible that job analysis data might be affected along only one or two of the dimensions, depending on the underlying psychological processes at

work. Fourth, some of these indices require researchers and practitioners to make judgments about what is reasonable in their job analysis data. That is, in some instances (e.g., assessing the completeness of a job analysis data collection) they will have to decide what is appropriate given their situation (e.g., deciding if the data are complete enough). Finally, higher *or* lower levels of these indices could indicate inaccuracy. For example, some processes can serve to inflate these measures (e.g., conformity could inflate reliability and agreement), thereby giving researchers an inappropriate sense of security about the accuracy of their job analysis data.

Summary

Interestingly, all these different conceptions of accuracy are based on principles of classical test theory. In particular, they rely on the notion that a 'true score' exists. But as Sanchez and Levine (2000) suggest, such an assumption may be problematic in the job analysis context. For example, some have suggested that work settings and environments are becoming more dynamic in nature (Carson and Stewart, 1996). Such a high rate of change in the organizational context would seem to preclude conceptualizing jobs as static entities. Still others have suggested that over time the nature of tasks performed by incumbents can change quite dramatically (Borman, Dorsey and Ackerman, 1992). If jobs change over time, then the true score model would not appear to apply. Finally, as Sanchez and Levine (2000) note, jobs are partly a social construction, reflecting the opinions and predispositions of those who 'construct' the job from more fundamental elements. If jobs are socially constructed, then a 'true score' (as defined in classical test theory) cannot exist.

This would suggest that it is difficult (if not impossible) to establish a single 'gold standard' or unquestionably correct description of a job (Sanchez and Levine, 2000). If accuracy is viewed as convergence to a known standard, then speaking in terms of the 'accuracy' of job analysis data is inappropriate, in part because there are rarely unambiguous standards against which to judge these data. Instead, there are multiple ways to index convergence, and this convergence may or may not reflect accuracy. As noted, the conceptions highlighted earlier employ different means to describe and assess this convergence.

This points to a seemingly intractable problem: if one focuses on the accuracy of job analysis data, it is relatively easy to demonstrate how these data might not be stable or objective in an absolute sense. Because of this, we may never be certain in our knowledge about the data's accuracy. In an attempt to avoid these difficulties, Sanchez and Levine (2000) focused on consequential validity as a standard for job analysis accuracy. Such a perspective provides a needed shift in thinking about the issue. And as Morgeson and Hofmann (1999) note, focusing on outcomes has substantial precedent in other areas of organizational science.

But as a standard for job analysis accuracy, consequential validity has two difficulties. First, it reflects usefulness more than accuracy. This does not diminish usefulness as an important consideration, but it is qualitatively different than notions of accuracy. Second, and more importantly, the same problems associated with making judgments about job analysis data remain. In other words, Sanchez and Levine's (2000) conception does not address the difficulties associated with tying job analysis data to some evaluative standard (be it accuracy or usefulness). Again, these criticisms do not obviate the importance of Sanchez and Levine's (or any other) perspective on job analysis accuracy. They all do provide unique perspectives on the quality and uses of job analysis data. But what they all fail to do is solve the basic dilemma of how to actually validate job analysis data.

Because of this, we suggest that it might be useful to shift from an exclusive focus on the accuracy (or consequential validity) of job analysis data to a consideration of the validity of job

analysis *inferences*. This shift in focusing to validating job analysis inferences has ample precedent in the personnel selection literature (Binning and Barrett, 1989; Lawshe, 1985), avoids the problems that have plagued discussions of job analysis accuracy, and implicitly includes Sanchez and Levine's (2000) notion of usefulness or utility.

Toward an Inference-based Model of Job Analysis Validity

To understand the nature of job analysis inferences and how they might be validated, it is necessary to describe the two primary reasons job analyses are conducted. This highlights some of the assumptions that are made when conducting job analyses, thereby revealing the range of inferences commonly made. It is then possible to discuss the nature of these inferences, some ways in which the inferences could be tested, and some implications of this conceptualization.

Reasons job analysis information is collected

Job analyses have long been conducted to determine the tasks and duties performed by job incumbents. Toward this end, job analysts often write job descriptions, which describe the work activities that are performed in the job (McCormick, 1979). Thus, job descriptions identify *what* is done in the job. This is often helpful for such things as identifying major job responsibilities for inclusion in a performance management system or defining the content of training programmes.

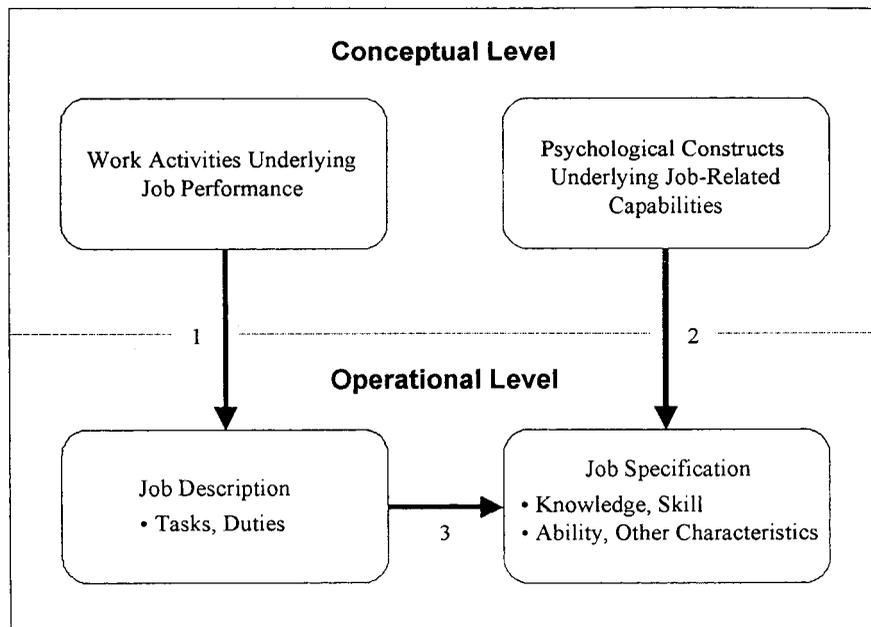


Figure 1. A model of the inferences made in the job analysis process

Note: 1 = Job descriptive inference. 2 = Job specification inference. 3 = Operational inference

Job analyses have also been conducted to determine the knowledge, skill, ability, and other characteristics (KSAOs) needed to perform a given job. Toward this end, job analysts often write job specifications, which set forth the job requirements for prospective job candidates or incumbents (McCormick, 1979). Thus, job specifications identify *how* a job gets done. This is often helpful for such things as establishing job requirements for personnel selection purposes or developing compensation programmes that reward employees based on their skill levels.

Critical job analysis inferences

The process of creating job descriptions or job specifications involves at least three critical inferences (Figure 1). The first inference involves the extent to which a job description and lists of tasks and duties adequately represents the work activities that underlie job performance. The second inference involves the extent to which a job specification and the KSAOs identified adequately represent the psychological constructs underlying job-related capabilities. The third inference involves the extent to which the KSAOs are needed to perform job tasks and duties.

Inference 1 can be thought of as *job descriptive inference* because it concerns the relationship between the work activities that underlie job performance and the manner in which a job is defined or established via an identification of tasks and duties. This is the inference that the description of tasks or duties actually represents the physical and mental activities that are performed in the job. There are, however, some difficulties associated with validating this inference. First, some of the mental activities performed on the job are not directly observable. Second, jobs are really collections of demands with imprecise boundaries, making it difficult to definitively identify where one job stops and another starts. Third, a lack of adequate job characteristics taxonomies (Fleishman and Quaintance, 1984) further hinders knowledge of whether all relevant aspects of work have been described.

Notwithstanding these difficulties one way to support this inference is by first identifying major work activity domains (e.g., by using existing taxonomies or classification systems) and then deriving the tasks associated with the domains. In this sense a content validity strategy may be appropriate, where evidence is presented that demonstrates how the task analysis adequately samples from the work activities domain. Another way to establish the linkage is to observe and carefully describe what the worker does, and how, to whom or what, and why a worker does it (Goldstein, Zedeck and Schneider, 1993). This can provide a conceptual and empirical justification for a task or set of tasks and duties.

Inference 2 can be thought of as *job specification inference* because it concerns the relationship between the psychological constructs underlying job-related capabilities and the KSAOs identified as important for successful job performance. This is the inference that a specification of KSAOs actually represents the psychological constructs needed to perform the job. As with inference 1, there are some difficulties associated with validating this inference. First, because they reflect psychological constructs, KSAOs are not directly observable. Second, some have suggested that making judgments about KSAOs (and AOs in particular) require a much greater inferential leap than occurs when simply describing tasks or duties (Harvey, 1991). This suggests that inferring KSAOs are much more difficult than describing tasks. In addition, there is reason to believe that inferring abilities and other characteristics is much more difficult than inferring knowledge and skills (Harvey, 1991). This is primarily due to the fact that knowledge and skills are often directly linked to the performance of tasks and defined in terms of learned, observable behaviors, whereas ability and other characteristics are much more hypothetical and not generally tied to specific behaviors identified in a job description.

Notwithstanding these difficulties, one way to support this inference is by first defining the constructs that comprise the range of job-related capabilities and then outlining more specific KSAOs that tap into these constructs. Again, a content validity strategy may be appropriate, where evidence is presented that demonstrates how specific KSAOs adequately sample the range of factors that underlie successful job performance. In this respect it is critical to make certain that the KSAOs are observable and operationally defined, as opposed to abstract traits.

Inference 3 can be thought of as an *operational linkage* because it concerns the relationship between the job description and the job specification and occurs at the operational level. This is the inference that the KSAOs are needed to perform the identified tasks and duties. Establishing this inference is one of the more perplexing problems in job analysis (Arvey, Salas and Gialluca, 1992). Because of this, a number of researchers have discussed ways to measure this linkage. For example, it may be possible to identify specific KSAOs that are needed to effectively perform tasks, identifying how they function, and the level at which they are used (Goldstein *et al.*, 1993). Such a process may be facilitated by devising a task-by-ability matrix where each KSAO is matched to specific tasks (Tenopyr, 1977). Finally, it may be possible to have subject matter experts rate the relevance of KSAOs to the tasks in order to empirically examine the linkages (Goldstein *et al.*, 1993).

Implications of inference-based model

This model of job analysis inferences has a number of implications for job analysis research and theory. First, it shifts the focus from the validation of job analysis data to the validation of job analysis inferences. This shift is needed because it is difficult to validate job analysis data in a definitive, 'true score' sense.

Second, the validation of some inferences may not be needed for certain kinds of job analyses. For example, if a job analysis is conducted to develop a new performance management system, it is likely that a task or duty analysis will be conducted. In such situations, validating the job specification inference may be unneeded. This suggests that an important consideration when validating any inference concerns the reason or purpose for which the job analysis data is being collected. That is, the validation process should be driven by the eventual use of the job analysis information (Sanchez and Levine, 2000). The use identifies which inferences are important, the kinds of evidence or standards to employ during the validation process, and the confidence one can have in the job analysis output.

Third, the validity of the operational inference (inference no. 3) is dependent on the confidence in both the job description inference (inference no. 1) and the job specification inference (inference no. 2). If these two latter inferences are not well established, it will be very difficult to have confidence in the relationship between tasks and duties and KSAOs. This is because inference numbers 1 and 2 represent both content and construct validity concerns. If the task analysis does not adequately represent the work activities and if the KSAO analysis does not adequately represent the psychological constructs, it is difficult to have confidence in the inference between tasks and KSAOs (inference no. 3).

Fourth, the conceptual level of the model is more abstract, construct-based, and unobservable whereas the operational level is more concrete, operational, and observable. In a related way, the inferences highlighted in Figure 1 can be arranged hierarchically in terms of the magnitude of the inferential leap that must be made. That is, some of the inferences in the model are relatively straightforward and concrete, whereas others are more abstract. For example, the job description inference (no. 1) is less subject to multiple interpretations because of the fact that tasks are commonly defined in terms of specific behaviors that are readily observable. Because of this, we

generally have high confidence in making this inference. The job specification inference (no. 2), on the other hand, requires considerably more judgment and involves more hypothetical constructs, particularly when ability and other characteristics are inferred (Harvey, 1991).

Finally, as the model demonstrates, there are at least two different ways to estimate the KSAOs that are needed to adequately perform the job. Traditionally, KSAOs have been estimated via a three-step process that involves three separate inferences. First, jobs are analyzed and described in terms of a set of tasks and duties. These tasks are assumed to represent the work activities (inference no. 1). Second, the KSAOs are identified and are assumed to represent the psychological constructs (inference no. 2). Third, the tasks are linked to the KSAOs (inference no. 3). Because of the range of different inferences that are made and the fact that the process begins with an analysis of tasks, this can be thought of as an *indirect estimation model*. One of the drawbacks to such an approach is the fact that a large number of judgments need to be made when linking tasks to KSAOs. The linkage process can become quite cumbersome as the number of tasks increase, thereby increasing the potential for inaccuracy in the judgments.

More recently, practitioners and researchers have directly estimated KSAOs by asking job incumbents or job analysts to rate the job in terms of the level or importance of various KSAOs. For example, the new Occupational Information Network (O*Net) developed by the Department of Labor to replace the Dictionary of Occupational Titles has defined a number of different content domains (e.g., skills, abilities) that directly assess these job requirements (Peterson *et al.*, 1999). Because these estimates involve only one inference (no. 2), this can be thought of as a *direct estimation model*. Although this direct estimation model involves fewer inferences, it relies on an assumption that possessing these capabilities will enable job holders to adequately perform relevant work activities. If this assumption is not warranted, the validity of direct estimates may be called into question. This suggests that at least in some instances, directly estimating KSAOs may prove to be problematic.

Conclusion

As we have highlighted, there are a number of different inferences that must be made when conducting a job analysis. In addition, there is no single standard or test for validating job analysis inferences. As such, the present model focuses on the inferences between the various job analysis components and the kinds of evidence used to validate these inferences. It is hoped that this brief discussion adds to the debate on accuracy in job analysis and helps clarify some of the issues involved when conducting job analyses in organizational settings.

References

- Arvey RD, Salas E, Gialluca KA. 1992. Using task inventories to forecast skills and abilities. *Human Performance*, **5**: 171–190.
- Ash RA. 1988. Job analysis in the world of work. In *The Job Analysis Handbook for Business, Industry, and Government*, Vol. 1, Gael S (ed.) Wiley: New York; pp. 3–13.
- Binning JF, Barrett GV. 1989. Validity of personnel decisions: a conceptual analysis of the inferential and evidential bases. *Journal of Applied Psychology*, **74**: 478–494.

- Borman WC, Dorsey D, Ackerman LD. 1992. Time-spent responses as time allocation strategies: relations with sales performance in a stockbroker sample. *Personnel Psychology*, **45**: 763–777.
- Campion MA, Morgeson FP, Mayfield M. 1999. O*Net's theoretical contributions to job analysis research. In *An Occupational Information System for the 21st Century: The Development of O*NET*, Peterson NG, Borman WC, Mumford MD, Jeanneret PR, Fleishman EA (eds) APA: Washington, DC; pp. 297–304.
- Carson KP, Stewart GL. 1996. Job analysis and the sociotechnical approach to quality: a critical examination. *Journal of Quality Management*, **1**: 49–65.
- Cronbach LJ. 1955. Processes affecting scores on 'understanding of others' and 'assumed similarity'. *Psychological Bulletin*, **52**: 177–193.
- Cronbach LJ, Gleser GC, Nanda H, Rajaratnam N. 1972. *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles*, Wiley: New York.
- Fleishman EA, Quaintance MK. 1984. *Taxonomies of Human Performance: The Description of Human Tasks*, Academic Press: Orlando, FL.
- Goldstein IL, Zedeck S, Schneider B. 1993. An exploration of the job analysis–content validity process. In *Personnel Selection in Organizations*, Schmitt N, Borman WC and Associates (eds) Jossey-Bass: San Francisco, CA; pp. 3–34.
- Harvey RJ. 1991. Job analysis. In *Handbook of Industrial and Organizational Psychology*, Vol. 2, 2nd edn. Dunnette MD, Hough LM (eds) Consulting Psychologists Press: Palo Alto, CA; pp. 71–163.
- 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.
- Kozlowski S. W. J., Hattrup K. 1992. A disagreement about within-group agreement: disentangling issues of consistency versus consensus. *Journal of Applied Psychology*, **77**: 161–167.
- Lawshe CL. 1985. Inferences from personnel tests and their validities. *Journal of Applied Psychology*, **67**: 237–238.
- McCormick EJ. 1979. *Job Analysis: Methods and Applications*, ANACOM: New York.
- Morgeson FP, Campion MA. 1997. Social and cognitive sources of potential inaccuracy in job analysis. *Journal of Applied Psychology*, **82**: 627–655.
- Morgeson FP, Hofmann DA. 1999. The structure and function of collective constructs: implications for multilevel research and theory development. *Academy of Management Review*, **24**: 249–265.
- Nunnally JC, Bernstein IH. 1994. *Psychometric Theory*, 3rd ed. McGraw-Hill: New York.
- Peterson NG, Mumford MD, Borman WC, Jeanneret PR, Fleishman EA. 1999. *An Occupational Information System for the 21st Century: The Development of O*Net*, APA: Washington, DC.
- Sanchez JI, Levine EL. 2000. Accuracy of consequential validity: which is the better standard for job analysis data?. *Journal of Organizational Behavior*, **21**: 809–818.
- Schwab DP. 1980. Construct validity in organizational behavior. In *Research in Organizational Behavior*, Vol. 2, Staw BM, Cummings LL (eds) JAI Press: Greenwich, CT; pp. 3–43.
- Shavelson RJ, Webb NW. 1991. *Generalizability Theory: A Primer*, Sage: Newbury Park, CA.
- Shrout PE, Fleiss JL. 1979. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, **86**: 420–428.
- Stone EF, Gueutal HG. 1985. An empirical derivation of the dimensions along which characteristics of jobs are perceived. *Academy of Management Journal*, **28**: 376–396.
- Tenopyr ML. 1977. Content–construct confusion. *Personnel Psychology*, **30**: 47–54.
- Uniform guidelines on employee selection procedures. 1978. *Federal Register*, **43**, 38290–38315.
- Williamson LG, Campion JE, Malos SB, Roehling MC, Campion MA. 1997. Employment interview on trial: linking interview structure with litigation outcomes. *Journal of Applied Psychology*, **82**: 900–912.