Open access • Journal Article • DOI:10.1037/0021-9010.89.4.674

Self-presentation processes in job analysis: a field experiment investigating inflation in abilities, tasks, and competencies. — Source link

Frederick P. Morgeson, Kelly Delaney-Klinger, Melinda S. Mayfield, Philip Ferrara ...+1 more authors

Institutions: Michigan State University, University of Missouri, New York State Unified Court System, Purdue University

Published on: 01 Aug 2004 - Journal of Applied Psychology (American Psychological Association)

Topics: Job analysis, Job attitude, Job performance, Job design and Job satisfaction

Journal of Applied Psychology
2004, Vol. 89, No. 4, 674–686

0021-9010/04/$12.00 DOI: 10.1037/0021-9010.89.4.674

Self-Presentation Processes in Job Analysis: A Field Experiment Investigating Inflation in Abilities, Tasks, and Competencies

Frederick P. Morgeson and Kelly Delaney-Klinger
Michigan State University

Melinda S. Mayfield
University of Missouri—Columbia

Philip Ferrara
New York State Unified Court System

Michael A. Campion
Purdue University

Although job analysis is a widely used organizational data collection technique, little research has investigated the extent to which job analysis information is affected by self-presentation processes. This study represents the first direct test of the propositions offered by F. P. Morgeson and M. A. Campion (1997) concerning self-presentation in job analysis measurement. Using an experimental design, the authors examined job incumbent response differences across ability, task, and competency statements. Results indicated that ability statements were more subject to inflation than were task statements across all rating scales. Greater endorsement of nonessential ability statements was responsible for the differences. This produced higher endorsement of ability items but lower mean ratings. Finally, frequency and importance ratings of global competency statements were generally higher than decomposed ability and task scales, but required-at-entry judgments demonstrated the opposite relationship.

Job analysis data is perhaps the most widely gathered type of organizational information for developing human resource (HR) management systems. It forms the foundation upon which many important HR management systems are built (Butler & Harvey, 1988), including selection systems, training programs, performance management programs, and compensation systems. The seemingly straightforward character of collecting information about jobs has led many to assume that job analysis methods result in reliable, valid, and unbiased information. It has recently been suggested, however, that job analysis information may be subject to numerous social and cognitive sources of inaccuracy (Morgeson & Campion, 1997). Such inaccuracies can negatively affect the HR systems that rely on job analysis. This study examined how self-presentation processes can serve to inflate job analysis responding and represents the first direct test of the propositions outlined by Morgeson and Campion (1997).

Understanding the nature of job analysis accuracy is particularly important given some of the new directions job analysis practice has taken. For example, the task-based Dictionary of Occupational Titles has been replaced with the more ability-based Occupational Information Network (O*NET; Peterson et al., 2001). It has been suggested that more abstract inferences are required when making ability judgments compared with task judgments (Harvey, 1991; Morgeson & Campion, 2000). There remain a number of important questions regarding the extent to which ability judgments can be accurately made. In addition, competency-oriented approaches have been forwarded as a replacement for traditional task-based job analysis methods (Shippmann et al., 2000). Again, it is not clear how accurately incumbents can make these more global judgments and how social and cognitive factors might influence competency judgments.

As a technique that can be used to collect a variety of information about job and worker requirements, job analyses can be conducted to collect information about the tasks performed, the abilities needed to perform the tasks, or the competencies needed to perform a range of tasks. For example, sometimes a job attribute can be stated as a task (e.g., "performs mental calculations"), an ability (e.g., "ability to perform mental calculations"), or a competency (e.g., "competence to perform mental calculations"), depending on the preferences of the analyst or the purpose of the project. Although it may seem that differences between these statements are trivial, there are theoretical reasons to expect differences, with some types of statements receiving higher ratings than others (Morgeson & Campion, 1997).

The inflation of certain job analysis judgments can reflect underlying self-presentation motives, in which responding is designed to shape the perceptions held by others (Schlenker, 1980). Respondents may have different self-presentational motives, including a desire to strategically influence the outcomes they receive from others or an attempt to safeguard their own self-concept (Leary, 1995). Despite differing motivations, any inflation in responses is likely to be consistent and predictable across different types of job analysis approaches to the same job. The present research examined these issues by describing how self-presentation processes influence rater judgments of different types

Frederick P. Morgeson and Kelly Delaney-Klinger, Department of Management, The Eli Broad Graduate School of Management, Michigan State University; Melinda S. Mayfield, Department of Management, University of Missouri—Columbia; Philip Ferrara, New York State Unified Court System, New York; Michael A. Campion, Department of Management, Krannert School of Management, Purdue University.

Correspondence concerning this article should be addressed to Frederick P. Morgeson, The Eli Broad Graduate School of Management, Michigan State University, N475 North Business Complex, East Lansing, MI 48824-1122. E-mail: morgeson@msu.edu

of job analysis measures by comparing task, ability, and competency statements containing comparable job content.

Understanding Inflation in Job Analysis Responding

Differences Between Ability and Task Statements

One of the most basic distinctions made in job analysis has been the distinction between job-oriented and worker-oriented information (McCormick, Jeanneret, & Mecham, 1972). Job-oriented information typically includes job tasks and work procedures, whereas worker-oriented information is concerned with generalized worker requirements such as knowledge, skill, ability, and other worker characteristics (KSAOs). Although there are obvious differences between these types of information, one key difference lies in the extent to which they are directly observable.

For example, task statements are typically very specific, concrete, and directly observable, whereas ability statements are often less discrete and less observable (Harvey, 1991). As a consequence, more abstract inferences are needed when making judgments about abilities than when making judgments about tasks (Morgeson & Campion, 1997, 2000). This suggests that ratings of ability statements will be more susceptible to inflation by job analysis respondents, because it is more difficult to verify the presence of an ability than the performance of a task. This gives respondents greater opportunity to engage in self-presentation when responding to ability statements than when responding to task statements.

Another reason ability statements could be inflated compared with task statements is that job analysis respondents may think more in terms of their own individual talents and skills and not in terms of the abilities they actually use to perform their job successfully. To the extent that individuals are underutilized in their job, the likelihood of inflation increases. The fact that ability ratings may reflect a *self-rating* compared with a *job rating* further suggests that self-presentation will be more likely with ability statements.

At its core, self-presentation is the "process by which individuals attempt to control the impressions others form of them" (Leary & Kowalski, 1990, p. 34). This process involves the presentation of information predicted to be desirable by others. Most individuals display the characteristics of which they are most proud, including traits, abilities, values, or other personal characteristics. Although there are multiple goals of self-presentation, most include an attempt to boost or maintain self-identity or self-concept.

Most of the research on self-presentation has focused on direct tactics of image management (Cialdini, 1989; Richardson & Cialdini, 1981). These tactics involve highlighting or displaying information about the self, including personality traits, abilities, and personal accomplishments. However, Richardson and Cialdini (1981) pointed out that individuals may also use indirect tactics to shape the impressions formed by others. These tactics involve the presentation of information about the people and the things to which an individual is connected, to share the positive reputation or to avoid the negative reputation of these others. One of the indirect tactics described by Cialdini (1989) is "burnishing," which involves the enhancement of favorable features of a positively linked other person or thing.

This process of burnishing is likely to apply to evaluations of jobs even when the job, not the person, is intended as the referent. Given the amount of time and effort most work situations involve, it is likely that individuals will attempt to protect their self-concepts by providing favorable information about their jobs. According to Richardson and Cialdini (1981), indirect but favorable information about one's job will more likely lead to favorable impressions others hold about the self. This suggests that incumbents, in contrast with others outside the job, are likely to demonstrate self-presentation in their ratings of a particular position. This would imply overall inflation in job-related ratings.

A variety of research supports the view that self-presentation will be more pronounced when describing abilities rather than tasks. For example, Fiske and Taylor (1991) have suggested that individuals will tend to overstate their abilities unless they believe their actual abilities will be verified. DeNisi and Shaw (1977) demonstrated this in an organizational context by finding that self-reported abilities evidenced little convergence with test scores of abilities. This suggests that self-presentation is particularly likely when describing abilities. The framework of job analysis inaccuracy developed by Morgeson and Campion (1997) suggested that similar inflation would occur when incumbents are asked to report the abilities needed to perform a job compared with the tasks performed on the job.

Smith and Hakel (1979) also reported a general tendency for supervisors and incumbents to inflate their responses compared with job analysts on socially desirable items in a job analysis questionnaire. This suggests that job analysis items that sound more socially desirable will be judged as occurring more frequently and as being more important than items lower in social desirability. Task statements are likely to evidence less inflation than ability statements because ability statements sound more socially desirable and personally evaluative and are less verifiable.

For all these reasons, we expected ability statements to be more influenced by self-presentation processes than comparable task statements across all types of response scales (do you perform, frequency, importance, required at entry). This was likely to manifest itself in several different ways. First, incumbents are likely to indicate that a greater number of ability statements are part of their job.

> *Hypothesis 1:* More ability statements will be endorsed as being part of the job compared with task statements.

Second, the fact that a greater number of ability statements are endorsed as being part of the job will lead to a larger number of ability statements being rated by incumbents (e.g., in terms of frequency, importance, and required at entry). Differences in responding due to self-presentation processes between ability and task statements, however, will be most evident on statements that are less essential to the job. This is because ability and task statements that are clearly needed or performed will be correctly identified and rated. In effect, these ability and task statements are easily recognized and not subject to self-presentation. Self-presentation processes will affect less essential ability statements, however, because these statements require greater subjective judgment to determine whether they are part of the job. The endorsement and rating of less essential ability statements will produce higher ability statement ratings.

Hypothesis 2: Summed ability statement ratings will be higher than comparable task statement ratings.

Hypothesis 1 and Hypothesis 2 represent two distinct ways to index inflation in ability statement endorsement. They both involve examining how self-presentation affects the aggregation of ability and task statements. Yet the most common way ability and task statements are operationalized is by calculating the mean across rated ability or task items. If both the sum of items (Hypothesis 2) and the number of items endorsed (Hypothesis 1) increase (which are the two components of the mean), what will be the effect when the mean is calculated? The endorsement and rating of less essential ability statements serve to deflate mean ability ratings compared with mean task ratings because of the inclusion of more lower rated nonessential ability statements.

An example highlights why this is the case. Let us assume that one component of a job (e.g., data entry) has seven different ability and task statements. When rating tasks, incumbents may indicate that only three of these tasks are part of their job and rate them highly (because they are essential to the job). When rating abilities, however, incumbents may indicate that five of the abilities are part of their job, of which three would be rated highly (the same three that are identified as essential when tasks are rated) and two would be rated lower (which are the non-essential tasks). The mean of the task ratings would be higher than the mean of the ability ratings because the addition of the two nonessential abilities would be proportionally lower compared with high ratings of the fewer task statements. In effect, this serves to pull down or otherwise deflate the ability mean compared with the task mean. The ultimate effect of calculating only means would be to disguise the impact of self-presentation processes.

Hypothesis 3: Mean ability statement ratings will be lower than comparable task statement ratings.

Understanding Nonessential Job Analysis Statements

The general social desirability of ability statements was hypothesized to result in more nonessential ability statements being endorsed than task statements, leading to higher summed ability statement ratings and lower mean ability statement ratings. Because this comparison involves actual task and ability statements (i.e., those that some incumbents actually do perform), it is not clear when a particular statement becomes nonessential. To more directly test the issue of how the essentiality of an item affects ability and task statement endorsement, one must examine truly nonessential items.

One way to do this is to use bogus job analysis items (i.e., items that no respondent should endorse because the item describes a fictitious ability or task for the job in question). These kinds of items have been used in previous research and have been operationalized as a carelessness index (Green & Stutzman, 1986) or an inflation scale (Anderson, Warner, & Spencer, 1984). Both Green and Stutzman (1986) and Anderson et al. (1984) found that an alarming number of respondents endorsed bogus items, and consequently the researchers developed techniques to identify and separate these respondents from those who did not endorse bogus items.

For our purposes, the use of bogus items was one way to test the nonessential ability hypothesis because bogus items are truly nonessential, in that no respondents should endorse any bogus items as part of their job. We expected that more respondents would say that bogus ability statements are part of their job than bogus task statements because abilities appear more socially desirable. The tendency to indicate that bogus ability statements are part of their job will occur regardless of how these scales are operationalized (i.e., sum or mean) because all statements are nonessential.

Hypothesis 4: Summed or mean bogus ability ratings will be higher than will comparable bogus task ratings.

Differences Between Competency, Ability, and Task Statements

Competency modeling has emerged as a major force in HR practice in the last 10 years. Shippmann et al. (2000) noted that between 75% and 80% of surveyed companies have some form of competency-related applications in place. One of the reasons cited by Shippmann et al. for the growth of competency-based approaches is a concern that standard job analysis procedures are not well suited for organizations in which the nature of work departs from traditional conceptualizations of fixed jobs. Although there are innumerable ambiguities associated with the practice of competency modeling, perhaps one of the most vexing issues involves actually defining a competency.

Definitions of competencies have included demonstrated knowledge, skills, or abilities (Ulrich, Brockbank, Yeung, & Lake, 1995); a mixture of knowledge, skills, abilities, motivations, beliefs, values, and interests (Fleishman, Wetrogan, Uhlman, & Marshall-Mies, 1995); and a motive, trait, skill, aspect of one's self-image or social role, or a body of knowledge (Boyatzis, 1982). Notwithstanding the problems associated with different conceptualizations of competencies (Barrett & Depinet, 1991; Shippmann et al., 2000), these different definitions share the view that competencies are at a more global or comprehensive level of job description than job tasks and abilities. This global character of competencies is viewed as an advantage because it provides for a more flexible taxonomy of work given the demands of dynamic and changing organizational environments.

Yet these advantages need to be balanced against a large body of research in the decision-making literature that suggests that these kinds of global or holistic judgments are less accurate and of lower quality than are more decomposed judgments (Armstrong, Denniston, & Gordon, 1975; Dawes & Corrigan, 1974; Einhorn, 1972; Goldberg, 1971; Kleinmuntz, Fennema, & Peecher, 1996; Meehl, 1954, 1986; Miller, 1956; Morera & Budescu, 1998, 2001; Ravinder, 1992). Holistic methods directly assign overall values to a given stimulus, whereas decomposed methods divide the judgment task into a simpler set of subtasks (Fischer, 1977). In the job analysis context, holistic strategies involve incumbents using their knowledge about a job to make overall judgments about the job. This is similar to making global competency judgments. Decomposed strategies involve incumbents making judgments about the individual elements of a job. These individual judgments are then combined to derive an overall judgment about the job (Cornelius & Lyness, 1980). This involves breaking a global competency down to its component parts (e.g., specific abilities and tasks).

There are a variety of reasons why decomposed judgments are likely to be superior to holistic judgments. First, human informa-

tion processing is limited (Kahneman, Slovic, & Tversky, 1982; Miller, 1956). Decomposing the judgment task reduces the information-processing burden and may help reduce the ineffective simplifying strategies individuals often use when faced with complex judgment tasks (Einhorn, 1972). Second, decomposed judgments are likely to be more finely tuned than holistic judgments. Because rating scales contain a finite number of discrete values, multiple decomposed judgments (compared with single holistic judgments) allow for finer distinctions in the objects being judged (Ganzach, Kluger, & Klayman, 2000). This may increase judgment quality because it decreases the coarseness of the judgment being made. Third, decomposed judgments allow for the consideration of a larger number of object attributes (Armstrong et al., 1975; Fischer, 1977; Shepard, 1964). In decomposed judgments, it is less likely that important attributes will be ignored.

Although the use of holistic and decomposed judgments has been investigated in the job analysis domain, the superiority of decomposed strategies has not been fully supported. For example, Cornelius and Lyness (1980) found few differences between holistic and decomposed job analysis ratings when making overall evaluations of worker requirements and motivational characteristics of work. Similarly, Sackett, Cornelius, and Carron (1981) found no differences in job classification results when using a comprehensive task analysis or a simplified paired-comparison procedure. However, Butler and Harvey (1988) found virtually no convergence between holistic ratings of Position Analysis Questionnaire dimensions and decomposed ratings of individual Position Analysis Questionnaire items.

These inconsistent results raise questions about whether global competency judgments will be different from more decomposed judgments. Further compounding this problem is the fact that none of the research testing holistic versus decomposed judgments in job analysis has used directly comparable holistic and decomposed rating stimuli. For example, Sackett et al. (1981) used a series of 28 paired comparisons among eight jobs for their holistic judgments and used task ratings for their decomposed judgments. Because the content of the rating stimuli was so different, it is not clear whether any differences (or failures to find a difference) were due to the judgment process or the rating format and stimuli.

To address this problem and directly test whether holistic competency ratings are inflated compared with decomposed ratings, the present research compared competency statements and decomposed ratings with comparable content. This was accomplished by creating global competencies that were combinations of underlying abilities and tasks, allowing a direct and unambiguous test of the extent to which competency items are inflated relative to decomposed items.

One of the problems with the global nature of competencies is that it offers greater opportunities for inflated responding. This is the case for two reasons. First, because competency statements represent combinations of tasks and KSAOs, the respondent is not given the opportunity to separately rate the individual parts. Thus, the inclusion of potentially inapplicable content along with applicable content will serve to inflate the overall rating. In essence, respondents are forced to include job elements they may skip if rated separately.

Second, because competency statements generally are larger and more complex than the individual items, incumbents will view competencies as more important. Furthermore, this complexity is likely to have different effects depending on the particular rating scale. In terms of frequency and importance ratings, competencies are likely to be higher than decomposed item judgments because of the complexity inherent in a more global measure. In terms of whether a competency is required at entry to the job, however, competencies are likely to be rated lower (i.e., less required-at-entry) because of their apparent complexity and difficulty for new employees to perform. In addition, because needed-at-entry ratings reflect expectations individuals must meet when starting on the job, they must be able to perform all aspects of the competency.

Hypothesis 5: Frequency and importance ratings will be higher for competencies than for ability and task statements.

Hypothesis 6: Needed-at-entry ratings will be lower for competencies than for ability and task statements.

Method

Participants

Job analysis surveys were completed by 494 office clerical employees of a large statewide public organization in the United States. These clerical jobs had four different titles (followed by the percentage of respondents in the sample): senior office assistant (23%), principal office assistant (36%), senior office typist (23%), and principal office typist (18%). These jobs are similar enough to justify a common task survey. To empirically test whether there were any differences among the job titles, we ran one-way analyses of variance with the frequency, importance, or required-at-entry ratings for each of the job components as the dependent variable and the job title code number as the grouping variable. These analyses indicated no significant difference among job titles. As such, we chose to analyze the jobs as a single group. Most participants had been employed at the organization for at least 3 years (77.6%), were full-time employees (93.8%), and had some college education (70.7%). Approximately one quarter of the respondents were minority group members. Survey response rates varied across different job titles but averaged 40%. There were similar response rates for each of the survey forms (see below). Only full-time employees who had worked at the organization for longer than 1 year were included in the final sample, yielding a final sample size of 431.

Measures

Task, ability, and competency statements. The organization's existing office clerical series job analysis survey was modified to update items on the basis of subject matter expert input, to add bogus items, and to add the competency statements. Each survey contained 12 components that comprised the major duties of all jobs (see Table 1). For each of the components, there was a list of the specific tasks or abilities associated with that component, ranging from 4 to 13 items per component. Task and ability statements were matched for content. Specifically, task statements were translated into ability statements by adding the phrase *ability to* at the beginning of the statements. For example, the task "maintain appointment calendars, make travel and meeting arrangements, etc.," was translated into "ability to maintain appointment calendars, make travel and meeting arrangements, etc." This provided identical item content across statements, thus ensuring that differences across statements were not due to differences in content.

After examining the existing literature on competencies identified earlier, we wrote competency statements by combining task and ability statements into global competency statements and including the phrase *demonstrate competence in* the particular elements of each component. This is consistent with the notion that competencies are combinations of

Table 1
Major Job Components and Number of Survey Items for Each Job Component

Name of job component	No. of survey items
1. Preparation of written materials	7
2. Filing	4
3. Record keeping	7
4. Office management	5
5. Work supervision	4
6. Statistical and financial	7
7. Typing and data entry	11
8. Computer applications	4
9. Managing and maintaining computer systems	4
10. Providing information to others	4
11. Contributing to service quality	13
12. Teamwork	10

several KSAOs or abilities or capabilities to perform a group of related tasks (Fleishman et al., 1995; Ulrich et al., 1995). As such, competencies were direct summaries of the entire list of task and ability statements for that job component.

These global competency judgments were compared with all of the relevant task and ability statements for that competency. As such, the rating stimuli are comparable in scope, differing only in the way in which the ratings are made (holistic vs. decomposed). This method of comparing holistic and decomposed stimuli is similar to methods used in the decision-making literature. Table 2 provides an example of the preparation of written materials component and the corresponding competency, ability, and task statements. This example shows that all of the task and ability statement content is reflected in the global competency.

Bogus items. Subject matter experts were used to write 14 bogus items. These items consisted of fictitious task and ability statements that no respondent should have endorsed. Example bogus items (along with an explanation as to why they are bogus) included "cross-check written materials against MLA standards" (there were no such standards at this organization), "matrix files and court records for archiving purposes" (matrix represents a fictitious activity and these employees were not involved in archiving), and "use a computer terminal or personal computer for fundamental analysis" (there was no such thing as fundamental analysis in this organization). As before, the bogus tasks were translated into an ability statement by adding the phrase, *ability to* at the beginning of each statement. Bogus items were distributed throughout the survey.

Rating scales. All competency, task, ability, and bogus statements were rated on three scales. The first was a 4-point frequency scale (in which 4 = *daily performance* and 1 = *yearly or less frequent performance*). Average internal consistency reliabilities across the 12 components (with 4–13 items each) were .68 for the task items and .69 for the ability items. The second scale was a 3-point importance scale (in which 3 = *very important* and 1 = *not very important*). Average internal consistency reliabilities across the 12 components were .77 for the task items and .80 for the ability items. The third scale was a 3-point required-at-entry scale ranging from 3 = *should be able to perform immediately,* 2 = *not expected to perform immediately but can be quickly learned on the job,* 1 = *not expected to perform immediately but can after formal training is provided.* Average internal consistency reliabilities across the 12 components were .76 for the task items and .88 for the ability items. It is important to recognize that internal consistency reliability is generally lower in job analysis applications because some individuals do not perform given parts of the job, thereby reducing reliability estimates. We report it here because it gives some sense of the homogeneity of the ratings within a job component.

Rating instructions varied depending on whether a task or ability was being rated for each of the response scales. For example, for task frequency, incumbents were instructed to answer "How often do you typically perform this aspect of your job?" For ability frequency, incumbents were instructed to answer "How often is this ability typically required in the job?" Similarly parallel instructions were provided for the importance and the required-at-entry

Table 2
Example Competency, Ability, and Task Statements for the Preparation of Written Materials Job Component

Competency	
Demonstrate competence in recording routine information, answering and composing original correspondence, preparing minutes of meetings, proofreading and correcting written materials, and verifying legal citations and references.	
Ability statements	Task statements
Ability to record phone messages and other routine information.	Record phone messages and other routine information.
Ability to answer correspondence (e.g., letters, memos, etc.) using form letters or standard wording to answer inquiries or provide basic information.	Answer correspondence (e.g., letters, memos, etc.) using form letters or standard wording to answer inquiries or provide basic information.
Ability to compose original correspondence (e.g., letters, memos, etc.), without using form letters or standard wording to answer inquiries or provide information.	Compose original correspondence (e.g., letters, memos, etc.), without using form letters or standard wording to answer inquiries or provide information.
Ability to prepare minutes of meetings, conferences, and similar events.	Prepare minutes of meetings, conferences, and similar events.
Ability to proofread and correct written materials (e.g., letters, memos, reports, etc.) for errors in punctuation, spelling, grammar, etc.	Proofread and correct written materials (e.g., letters, memos, reports etc.) for errors in punctuation, spelling, grammar, etc.
Ability to edit and correct written materials (e.g., letters, memos, reports, etc.) to correct meaning and/or style of narrative content.	Edit and correct written materials (e.g., letters, memos, reports, etc.) to correct meaning and/or style of narrative content.
Ability to verify legal citations and references by checking information in various legal reference books.	Verify legal citations and references by checking information in various legal reference books.

rating scales. This created parallel (and comparable) rating scales across task and ability statements and focused incumbents on the job itself.

In addition, respondents indicated whether the ability, task, and bogus items were part of their job. Because our main concern for the bogus items was the tendency to indicate that these items were part of the job, the items were combined to create an overall bogus endorsement scale. Internal consistency reliabilities were .75 for the bogus task scale and .82 for the bogus ability items.

Procedures

Surveys contained either competency and task statements or competency and ability statements. Given the isomorphism between the task and ability statements, we decided to use a between-subjects design with random assignment. Thus, every participant rated the competencies, but half the sample rated the task statements and the other half rated the ability statements. The order of the competency statements was also varied, with roughly half of the surveys presenting the competency statements (within each component) prior to the component's task or ability statements and the other half presenting the competency statements after the component's task or ability statements. This was done to assess whether the presentation of the competency statement influenced subsequent task or ability statement judgments and vice versa. A statistical analysis of competency means across the different presentation orders indicated that there were no order effects.

Participants were instructed to first read through all components and indicate whether each component was part of their job. Participants were then instructed to return to the first component. If they had indicated that the component was part of their job, they were instructed to read the competency statement associated with the component and to make their ratings. Next, participants were instructed to read each of the tasks or abilities associated with the component and to indicate whether each was part of their job. Finally, participants were instructed to make frequency, importance, and required-at-entry ratings for each task or ability statement that was part of their job.

The four different survey forms were randomly distributed to participants by the organization's HR department. Random assignment of participants to condition has the effect of equating the groups on any unmeasured factors. Participants completed the surveys and returned them in sealed envelopes to the HR department.

Results

Primary Findings

Tables 3, 4, and 5 contain the means and standard deviations for the scales across the 12 job components. Regression with dummy coding was used to examine the hypotheses (Cohen & Cohen, 1983). Thus, depending on the hypothesis being tested, we used one or two dummy codes to examine the mean differences. For example, when comparing the differences between tasks and abilities, we coded task ratings with a 1 and ability ratings with a 0. If the regression parameter is significant, then the mean rating would be significantly different. Statistical power was more than 90% to detect a small effect ($d = .20$) and more than 99% to detect a medium effect ($d = .50$; $p < .05$, one-tailed; Cohen, 1988). To understand the magnitude of any observed differences, we reported average effect sizes (d) . Small ($d = .20$), medium ($d = .50$), and large ($d = .80$) effects are so noted.

Hypothesis 1 predicted that the number of ability statements endorsed as being part of the job would be higher than the number of task statements so endorsed. For 11 of 12 job components, the number of ability statements endorsed was significantly greater than the total number of task statements endorsed ($p < .05$; see Table 3). The mean number of ability statements endorsed was between 0.11 and 2.74 higher than the mean number of task statements endorsed across the 12 job components. On average, ability statements were endorsed 0.80 more than task statements, representing a medium average effect ($d = .52$). These results provide strong support for Hypothesis 1.

Hypothesis 2 predicted that summed ability ratings would be higher than summed task ratings. For 9 of 12 job components, the summed ability frequency ratings were significantly higher than the summed task frequency ratings ($p < .05$; see Table 4). The mean of the summed frequency ratings of ability statements was between 0.54 and 8.70 higher than the mean of the summed frequency ratings of task statements across the 12 job components. On average, ability frequency ratings were 2.67 higher than task frequency ratings, representing a medium average effect ($d = .42$). For 10 of 12 job components, the summed ability importance ratings were significantly higher than the summed task importance ratings ($p < .05$). The mean of the summed importance ratings of ability statements was between 0.59 and 6.58 higher than the mean of the summed importance ratings of task statements across the 12 job components. On average, ability importance ratings were 2.25 higher than task importance ratings, representing a medium average effect ($d = .49$). For 11 of 12 job components, the summed ability required-at-entry ratings were significantly higher than the summed task required-at-entry ratings ($p < .05$). The mean of the summed required-at-entry ratings of ability statements was between 0.70 and 5.60 higher than was the mean of the summed required-at-entry ratings of task statements across the 12 job

Table 3
Numbers of Ability and Task Statements Endorsed as Part of the Job, for Each Component

	Job component											
Statement type	1	2	3	4	5	6	7	8	9	10	11	12
Ability												
M	5.04_a	3.33_a	5.63_a	3.18_a	2.63_a	4.21_a	7.75_a	3.06_a	2.86_a	3.63_a	10.28_a	7.98_a
SD	1.75	0.89	1.63	1.28	0.96	1.74	2.40	1.11	0.81	0.69	3.23	2.66
Task												
M	3.88_b	2.65_b	4.97_b	2.41_b	2.05_b	3.60_b	6.67_b	2.74_b	2.75_a	3.40_b	9.59_b	5.24_b
SD	1.79	1.02	1.62	1.14	0.72	1.59	2.46	1.04	0.88	0.79	3.23	2.73

Note. Within each component, values with corresponding subscripts are not significantly different at $p < .05$ (one-tailed). Higher values indicate that more statements were endorsed as part of the job (0 = *not part of job* and 1 = *part of job*).

Table 4
Summed Ability and Task Ratings for Each Job Component

Statement type	Job component 1	2	3	4	5	6	7	8	9	10	11	12
Frequency scale												
Ability												
M	16.48_a	10.53_a	19.61_a	10.08_a	7.16_a	12.86_a	27.71_a	9.93_a	10.09_a	13.86_a	38.23_a	27.49_a
SD	5.93	3.61	6.06	3.86	3.88	6.18	8.79	4.20	2.98	3.14	11.12	10.54
Task												
M	13.75_b	8.71_b	18.06_b	8.02_b	5.63_b	11.83_a	24.73_b	9.56_a	10.17_a	13.32_b	36.08_b	18.79_b
SD	6.52	3.43	6.13	3.44	3.00	5.80	9.38	4.17	3.31	3.14	12.63	9.40
Importance scale												
Ability												
M	12.98_a	8.61_a	15.36_a	8.35_a	7.16_a	11.63_a	20.62_a	7.55_a	7.29_a	9.84_a	27.78_a	20.81_a
SD	4.95	2.75	4.59	3.45	2.98	5.18	7.38	3.14	2.39	2.54	8.64	7.91
Task												
M	10.53_b	6.98_b	13.95_b	6.43_b	5.29_b	10.19_b	18.01_b	7.08_a	7.50_a	9.25_b	25.83_b	14.23_b
SD	5.15	2.84	4.75	3.09	2.03	4.89	7.03	3.02	2.63	2.60	9.44	7.46
Required-at-entry scale												
Ability												
M	11.33_a	7.39_a	11.06_a	6.29_a	4.24_a	8.52_a	18.35_a	6.12_a	5.30_a	8.08_a	21.56_a	18.57_a
SD	4.66	2.70	4.52	3.01	2.45	4.04	7.50	2.92	2.17	2.61	7.84	7.72
Task												
M	8.49_b	5.95_b	8.96_b	4.54_b	3.08_b	5.91_b	14.53_b	4.96_b	5.18_a	7.38_b	19.21_b	12.97_b
SD	4.14	2.61	3.81	2.48	1.70	3.53	6.85	2.79	2.28	2.58	7.56	7.07

Note. Within each component, values with corresponding subscripts are not significantly different at $p < .05$ (one-tailed). Higher values indicate that the statements were performed more frequently, were considered more important, and were more required at entry.

components. On average, ability required-at-entry ratings were 2.32 higher than task importance ratings, representing a medium to large average effect ($d = .65$). These results provide strong support for Hypothesis 2.

Hypothesis 3 predicted that mean ability ratings would be lower than mean task ratings. For 10 of 12 job components, the mean ability frequency ratings were significantly lower than the mean task frequency ratings ($p < .05$; see Table 5). The mean ability frequency ratings were between 0.08 and 0.24 lower than the mean task frequency ratings across the 12 job components. On average, ability frequency ratings were 0.17 lower than task frequency ratings, representing a small to medium average effect size ($d = .31$). For 4 of 12 job components, the mean ability importance ratings were significantly lower than the mean task importance ratings ($p < .05$). The mean ability importance ratings were between 0.11 and 0.15 lower than the mean task importance ratings across the 12 job components. On average, ability importance ratings were 0.13 lower than task frequency ratings, representing a small average effect size ($d = .28$). For 1 of 12 job components, the mean ability required-at-entry ratings were significantly lower than the mean task required-at-entry ratings ($p < .05$). These results provide mixed support for Hypothesis 3.

Hypothesis 4 predicted that summed or mean bogus ability ratings would be higher than comparable bogus task ratings. The summed bogus ability ratings ($M = 2.04$, $SD = 2.26$) were significantly higher than the summed bogus task ratings ($M = 1.00$, $SD = 1.52$; $p < .05$), representing a medium effect size ($d = .55$). The mean bogus ability ratings ($M = 0.20$, $SD = 0.22$) were significantly larger than the mean bogus task ratings ($M = 0.09$, $SD = 0.12$; $p < .05$), representing a medium to large effect size ($d = .65$). These results provide strong support for Hypothesis 4.

Hypothesis 5 predicted that frequency and importance ratings would be higher for competencies than for ability and task statements. Given that only a single rating was made for each competency on the response scales, the competencies must be compared with the mean ability and task ratings. The frequency ratings for 10 of 12 job components were significantly higher for the competencies than for the mean ability frequency ratings ($p < .05$; see Table 5). The competency frequency ratings were between 0.11 and 0.52 higher than the mean ability frequency ratings across the 12 job components. On average, competency frequency ratings were .30 higher than mean ability frequency ratings, representing a medium average effect ($d = .53$). In addition, the frequency ratings for 8 of 12 job components were significantly higher for the competencies than for the mean task frequency ratings ($p < .05$). The competency frequency ratings were between 0.08 and 0.48 higher than the mean task frequency ratings across the 12 job components. On average, competency frequency ratings were 0.21 higher than the mean task frequency ratings, representing a small to medium effect size ($d = .37$).

The importance ratings for 6 of 12 job components were significantly higher for the competencies than for the mean ability importance ratings ($p < .05$). The competency importance ratings were between 0.12 and 0.20 higher than the mean ability impor-

Table 5
Competency, Mean Ability, and Task Ratings for Each Job Component

Statement type	Job component 1	2	3	4	5	6	7	8	9	10	11	12
Frequency scale												
Competency												
M	3.74_a	3.77_a	3.81_a	3.28_a	3.01_a	3.39_a	3.72_a	3.16_a	3.60_a	3.95_a	3.88_a	3.70_a
SD	0.62	0.61	0.53	0.99	1.15	0.86	0.66	1.06	0.82	0.30	0.40	0.65
Ability												
M	3.32_b	3.25_b	3.49_b	3.17_b	2.69_b	3.02_b	3.43_b	3.14_a	3.56_a	3.79_b	3.63_b	3.48_b
SD	0.59	0.72	0.47	0.66	1.05	0.82	0.54	0.86	0.52	0.39	0.46	0.58
Task												
M	3.54_c	3.29_b	3.63_c	3.37_a	2.67_b	3.26_a	3.66_a	3.35_b	3.71_b	3.87_c	3.71_c	3.59_b
SD	0.54	0.69	0.49	0.62	1.04	0.70	0.43	0.75	0.41	0.35	0.38	0.55
Importance scale												
Competency												
M	2.73_a	2.81_a	2.83_a	2.62_a	2.64_a	2.80_a	2.72_a	2.42_a	2.62_a	2.81_a	2.81_a	2.69_a
SD	0.48	0.43	0.40	0.58	0.52	0.41	0.52	0.67	0.59	0.41	0.42	0.52
Ability												
M	2.56_b	2.61_b	2.70_b	2.64_a	2.68_a	2.76_a	2.55_b	2.40_a	2.57_a	2.69_b	2.61_b	2.63_a
SD	0.40	0.41	0.37	0.37	0.45	0.38	0.41	0.56	0.46	0.40	0.40	0.40
Task												
M	2.67_a	2.61_b	2.82_a	2.66_a	2.58_a	2.81_a	2.68_c	2.51_a	2.72_b	2.69_b	2.67_b	2.66_a
SD	0.39	0.46	0.28	0.40	0.55	0.34	0.37	0.51	0.39	0.42	0.39	0.44
Required-at-entry scale												
Competency												
M	2.02_a	2.27_a	1.83_a	1.75_a	1.52_a	1.64_a	2.03_a	1.71_a	1.71_a	2.10_a	2.11_a	2.26_a
SD	0.72	0.68	0.70	0.67	0.67	0.66	0.75	0.71	0.62	0.65	0.66	0.68
Ability												
M	2.47_b	2.24_a	1.95_b	1.97_b	1.57_a	2.05_b	2.25_b	1.95_b	1.86_b	2.21_b	2.04_a	2.34_a
SD	0.52	0.54	0.56	0.59	0.64	0.61	0.52	0.60	0.56	0.54	0.50	0.58
Task												
M	2.19_b	2.23_a	1.84_a	1.87_b	1.49_a	1.62_a	2.12_a	1.74_a	1.87_b	2.14_b	1.99_a	2.48_b
SD	0.53	0.56	0.58	0.59	0.63	0.59	0.53	0.67	0.54	0.54	0.47	0.56

Note. Within each component, values with corresponding subscripts were not significantly different at $p < .05$ (one-tailed). Higher values indicate that the statements were performed more frequently, were considered more important, and were more required at entry.

tance ratings across the 12 job components. On average, competency importance ratings were 0.17 higher than the mean ability ratings, representing a small to medium average effect ($d = .39$). In addition, the importance ratings for 4 of 12 job components were significantly higher for the competencies than for the mean task importance ratings ($p < .05$). The competency importance ratings were between 0.04 and 0.20 higher than the mean task importance ratings across the 12 job components. On average, competency importance ratings were 0.13 higher than the mean task importance ratings, representing a small effect ($d = .23$). In total, these results provide moderate support for Hypothesis 5.

Hypothesis 6 predicted that required-at-entry ratings would be lower for competencies than for ability and task statements. As indicated in Table 5, the required-at-entry ratings for 8 of 12 job components were significantly lower for the competencies than for the mean ability required-at-entry ratings ($p < .05$). The competency required-at-entry ratings were between 0.11 and 0.45 lower than the mean ability required-at-entry ratings across the 12 job components. On average, competency required-at-entry ratings were 0.24 lower than the mean ability required-at-entry ratings, representing a small to medium effect ($d = .38$). In addition, the required-at-entry ratings for 5 of 12 job components were significantly lower for the competencies than for the mean task required-at-entry ratings ($p < .05$). The competency required-at-entry ratings were between 0.04 and 0.22 lower than the mean task ratings across the 12 job components. On average, the competency required-at-entry ratings were 0.14 lower than the mean task required-at-entry ratings, representing a small effect ($d = .23$). In total, these results provide moderate support for Hypothesis 6.

Supplemental Analyses

The results of Hypotheses 1–4 provide support for the notion that ability ratings are more subject to self-presentation processes. Yet there are potentially other explanations for these differences. For example, the differences observed in incumbents might be due to actual differences in task and ability requirements. To begin to rule out this and other alternative explanations, we collected additional task and ability ratings from clerical supervisors and trained job analysts. Because these individuals did not actually

perform the jobs and were rating the jobs of others, they did not have the same motivation to bias and were not likely to be subject to self-presentation processes. Thus, we expected no difference between ability and task ratings for supervisors and job analysts.

Surveys were distributed to 55 supervisors of clerical workers across a representative sample of organizational locations. Completed surveys were returned by 36 supervisors for a response rate of 65%. The surveys were completed and returned directly to the HR department. The sample was intentionally selected to be representative of the entire organization. Participating supervisors had from 5 to 10 years' experience in a supervisory capacity and thus were in an excellent position to make judgments about the job. In addition, 12 trained job analysts (master's and PhD-level psychologists) completed surveys.

Because of concerns over adequate levels of statistical power to detect significant effects, we combined the supervisor and job analyst data ($N = 48$) and interpreted one-tailed significance tests. Statistical power was 53% to detect a medium effect ($d = .50$) and 86% to detect a large effect ($d = .80$; $p < .05$, one-tailed; Cohen, 1988). The range of effect sizes found in the incumbent sample suggests that we had adequate statistical power to detect significant differences if such differences existed. Although we expected to find no differences (in essence accepting the null hypothesis), Cortina and Folger (1998) have suggested that there are circumstances when this is warranted. The supplemental study used the same measures, administered in the same fashion, with a similar type of sample, which did produce differences among incumbents. This suggests that the supplemental analyses were sensitive enough to produce and detect differences if they were present among supervisors and job analysts.

The same general methodology used with the job incumbents was used in the supplemental data collection. Specifically, participants received either a task or an ability survey. Of the 48 completed surveys, 26 were task surveys and 22 were ability surveys. Because of concerns about the amount of time it would take to complete the surveys, only frequency and importance ratings were collected. In addition, bogus items were not included because of concerns about negative participant reactions. Thus, we were able to conduct parallel analyses for Hypotheses 1–3. To interpret the results, we examined the statistical significance of the differences, differences in the pattern of means for nonsignificant comparisons, and average effect sizes.

When testing Hypothesis 1, we examined the extent to which supervisors and job analysts indicated that more ability statements were part of the job than task statements. As expected, there were no significant mean differences across all 12 job components. In examining the pattern of means, for 2 components the ratings were essentially identical, for 5 components more ability statements were viewed as part of the job, and for 5 components more task statements were viewed as part of the job. Finally, the average effect size was quite small ($d < .01$).

When testing Hypothesis 2, we examined the extent to which supervisors and job analysts had higher summed ability ratings than task ratings. For the frequency ratings, 2 of 12 job components were significantly different. Specifically, for Job Component 2, ability ratings were higher than task ratings. For Job Component 3, however, task ratings were higher than ability ratings. In examining the pattern of mean differences for the nonsignificant relationships, for 1 component the ratings were essentially identical, for 5 components ability ratings were higher than task ratings, and for 6 components task ratings were higher than ability ratings. The average effect size was small ($d = .09$). For the importance ratings, none of the 12 job components were significantly different. In examining the pattern of mean differences, for 5 components ability ratings were higher than task ratings, and for 7 components task ratings were higher than were ability ratings. The average effect size was small ($d = .04$).

Finally, when testing Hypothesis 3, we examined the extent to which supervisors and job analysts had lower mean ability ratings than task ratings. For the frequency ratings, 2 of 12 job components were significantly different. Specifically, for Job Components 2 and 4, ability ratings were higher than task ratings. In examining the pattern of mean differences for the nonsignificant relationships, for 7 components the ratings were essentially identical, and for 3 components ability ratings were higher than task ratings. Finally, the average effect size was small ($d = .17$). For the importance ratings, only 1 of the 12 job components was significantly different. Specifically, for Job Component 11, task ratings were higher than ability ratings. In examining the pattern of mean differences for the nonsignificant relationships, for 6 components the ratings were essentially identical, for 4 components ability ratings were higher than task ratings, and for 1 component task ratings were higher than ability ratings. Finally, the average effect size was small ($d = .01$).

To summarize these analyses, there were few significant differences between supervisor and job analyst ratings of ability and task statements. In fact, out of the 60 possible mean comparisons, only 5 were significantly different. This is about what would be expected by chance alone. In addition, there was no consistent pattern of inflation of ability statements. Finally, any differences that were found were quite small in magnitude (average $d = .06$). These results increase our confidence in the findings for the incumbent sample because we would expect job incumbents to self-present and have inflated ability ratings (and they did) but did not expect supervisors and job analysts to self-present and have inflated ability ratings (and they did not).

Discussion

The purpose of this study was to examine the extent to which job analysis information is affected by self-presentation processes. First, we found that incumbents endorsed more ability than task statements as being part of their job. Second, summed frequency, importance, and required-at-entry ratings were larger for ability than for task statements. Third, mean frequency ratings were smaller for ability statements than for task statements, but mean importance and required-at-entry scales showed fewer differences. Fourth, bogus ability statements were endorsed more often than bogus task statements. Finally, global competency statements were generally higher than abilities and tasks when frequency and importance judgments were made but were lower than abilities and tasks when required-at-entry judgments were made. The lack of similar levels of inflation in the supervisor and job analyst data lends further confidence to the conclusion that the differences we did detect in the incumbent sample were a result of self-presentational factors.

This study makes several contributions to the job analysis literature and has a number of implications for both research and

practice. The experimental design allowed us to directly examine how identical task and ability scales compare with one another. This was the first study to perform such a direct comparison, and it enabled a direct test of the degree to which self-presentation processes may be affecting job analysis responding.

Our results suggest that ability statements are more susceptible to self-presentation tactics. As such, this is the first direct support for the propositions outlined in Morgeson and Campion (1997). This may be because ability judgments require more abstract inferences than do task judgments (Morgeson & Campion, 1997, 2000). It may also be because ability statements are more difficult to observe and less verifiable than task statements, as well as sounding more socially desirable. Respondents may take this opportunity (consciously or unconsciously) to engage in self-presentation, thereby endorsing more nonessential abilities as being more frequently performed, more important, and more required at entry.

Yet it is important to recognize that simply adding the phrase *ability to* to a task statement does not actually produce an ability statement (because abilities are human attributes rather than activities). There are two reasons why ability was operationalized in this manner. First, to make equivalent comparisons between task and ability statements, we felt it was essential to make sure the different statements had identical item content. The statements had to possess the same content (e.g., recording phone messages) because if the content were different, any observed differences could be due to the different statement type (i.e., task or ability) or they could be due to differences in underlying item content. By maintaining the same item content (at the potential expense of creating less than ideal ability statements), this alternative explanation is eliminated. Second, this operationalization reflects what sometimes occurs in applied settings. We have witnessed many job analysis projects in which supervisors and incumbents generated ability statements by simply adding the phrase *ability to* to task statements. This is inappropriate because it fails to make the distinction between domains of work behaviors (tasks) and human attributes needed to perform such tasks (abilities). Although we adopted this operationalization to maintain similar content, we do not advocate creating ability statements in this manner outside of a research setting such as this.

What is remarkable about the present study is that there were no substantive differences between the ability and the task statements (i.e., simply the inclusion of the phrase *ability to*). As such, this can be viewed as a very weak manipulation because there were only superficial differences between task and ability statements. Prentice and Miller (1992, p. 160) have suggested that "the statistical size of effect is heavily dependent on the operationalization of independent variables." If minimal manipulation of the independent variable still accounts for some variance in the dependent variable, the effects should be regarded as very important (Fichman, 1999; Prentice & Miller, 1992). It is likely that there would be much more self-presentation if the ability statements were more abstract.

Including bogus items in the questionnaire allowed us to examine how self-presentation processes have an impact. Consistent with the idea that respondents would endorse more nonessential ability statements than task statements, we found that bogus ability statements are more likely to be endorsed than bogus task statements. Thus, on statements no incumbent should endorse, abilities were more frequently identified as part of the job. This suggests that the motivation to present oneself favorably may be stronger than the ability to differentiate the actual abilities associated with a particular position.

As expected, the detection of inflation was affected by the manner in which the data were examined. When incumbents were asked whether a statement was part of their job or when response scales were summed, the expected inflation was observed. When a mean was computed, however, either the effect was reversed or no differences were found. Given the typical job analysis practice of calculating means on frequency and importance ratings, this suggests that the effect of self-presentation processes may have been disguised in prior job analysis practice. This can lead to a false sense of security about the accuracy of resultant job analysis data. Future research should use the more sensitive measures used in the present research to determine whether responses have been inflated. This is important because HR systems that use level ratings (such as performance appraisals or compensation systems) may be adversely affected by inflation. These systems may not recognize high levels of performance (because the standards are set too high) or they may overreward employees (because jobs are rated too highly).

Of interest, past job analytic research has tended to focus on covariance-based measures (Dierdorff & Wilson, 2003), such as interrater reliability (DeNisi, Cornelius, & Blencoe, 1987; Dierdorff & Wilson, 2003), test–retest reliability (Wilson, Harvey, & Macy, 1990), correlations between different response scales (Butler & Harvey, 1988; Sanchez & Fraser, 1992), and correlations between different data sources (Smith & Hakel, 1979). The present study's use of the mean levels is valuable for three reasons. First, convergence in level ratings yields information distinctly different from that of covariance-based measures. In fact, the effects of self-presentation will be detected only by covariance-based measures if people self-present in different amounts (i.e., rank ordering changes). Level ratings thus provide a more sensitive measure to detect self-presentation. Second, as noted, many HR systems are developed through the use of frequency or importance ratings (i.e., level ratings are what are interpreted). If certain types of job analysis items are systematically under- or overestimated, this can affect the HR systems that result. Third, level ratings can be compared with each other instead of with some true score. This circumvents the problems associated with developing an absolute standard of accuracy.

These findings also offer insight into the importance of respondent motivation to self-present in a job analysis setting. Job analyses are conducted for a variety of reasons. The purpose of the job analysis is likely to have a pronounced effect on incumbent motivation to self-present (Morgeson & Campion, 1997). In the present study, respondents were told that the questionnaire data were being used to simply update existing task, knowledge, skill, and ability information for job-related examinations. Because the results of a job analysis done for this purpose minimally impact the respondent (i.e., any HR system that may result does not directly affect the respondent), the motivation to self-present is likely to be low. Yet respondents did self-present by inflating certain ratings. In situations in which respondents might be more motivated to self-present, there will likely be even greater inflation. For example, if the job analysis is conducted to determine compensation, job classification, or training needs, the uses of the job analysis data

will have a pronounced effect on respondents. This is likely to further increase self-presentation and distortion in job analysis responding. This is an important area for future research.

These findings also have implications for job analysis systems that include worker-oriented descriptor domains (e.g., ability, skills, personality) like those used in the O*NET. Although ability statements used in this study are much more specific and narrow than those used in the O*NET, they still evidenced differences from task statements. These results are likely to become even more problematic, as future O*NET data collection is largely planned to occur through incumbent self-reports, which the present study shows are subject to inflation. It may be that all such worker-oriented domains that are not directly tied to specific tasks will be vulnerable to the sort of inflation found here. As Morgeson and Campion (2000) have noted, ability statements can be directly or indirectly judged. Traditionally, ability requirements have been derived indirectly by inferring them from tasks. Future research is needed to determine the extent to which direct judgments of ability requirements (and other worker-oriented descriptors) can be accurately made. This becomes more critical as job analysis systems such as the O*NET increase in prominence.

Another key implication concerns the use of ability statement data for selection and other HR system development. For example, the identification of important job-related abilities is essential when developing and choosing selection measures. If ability statements are systematically inflated, this may result in a more complex selection system than is actually needed. This suggests that steps must to be taken to guard against inflation. As the supplemental data analysis suggests, nonincumbent judgments (those of supervisors and job analysts) are less likely to be systematically inflated. These can serve as an important check against incumbent judgments of ability statements. Given the extensive use of job incumbent self-reports in job analysis practice, however, a key question for future research concerns whether there are ways to structure job-incumbent ability statement data collection to avoid problems of inflation. One possibility might be to give incumbents explicit task–ability linkages so that they can have the appropriate frame of reference and can anchor their ability judgments in the tasks performed. Research on strategies designed to avoid job incumbent inflation is sorely needed.

This study also has implications for the use of global competency judgments within the job analysis domain. Our results suggest that competency modeling and other techniques that require such global judgments can be subject to inflation in responding. This study eliminated a key problem with previous comparisons of holistic and decomposed judgments by creating competencies that were direct combinations of the individual tasks and abilities. Even though this resulted in competencies that were fairly concrete, we still found evidence that incumbents inflate their ratings.

As organizations and researchers turn to the competency modeling approach as an alternative to traditional job analysis, this research sounds a cautionary note. Clearly, additional research should be conducted to examine the conditions under which respondents can make accurate global judgments. It is likely that as competency statements become more abstract and less verifiable, the possibility of distortion increases. In addition, the purpose of the analysis (e.g., strategic clarification vs. organizational restructuring) is likely to exert a strong influence on incumbent responding.

There are several potential limitations to this research that need to be kept in mind. One of these issues is that participants in this study represented a single occupational job family, office clerical workers. The extent to which these findings generalize to other job families, particularly those with greater complexity and mental demands, needs further investigation. The fact that most aspects of clerical jobs such as these are directly observable, however, serves to limit self-presentation because incumbents would know that others would be able to easily verify the information. In contrast, more complex jobs with greater mental demands would be less observable by an outsider and thus make it more likely that incumbents would engage in self-presentation. This suggests that this study is a conservative test of the effects of self-presentational processes in job analysis.

Another potential concern about the task and ability statement comparisons concerns the scope of the ability and task statements. For example, although one may rarely need to "compose original correspondence," one might need to have the ability to do so in case it does come up. Similarly, there may be cases in which important abilities simply do not exist as tasks. Both instances may result in higher ratings of ability statements. This is not likely to have occurred in the present study for three reasons. First, incumbents were asked to focus on the job itself. This is important because it emphasizes that the rating target is the job and not the individuals in the job. Second, there is a parallelism between the task and the ability ratings in that they focus on task performance or ability requirements. This is important because the ratings focus solely on what is typically done or required and not what might be done or required. Third, incumbents in this study made the same judgments (i.e., frequency and importance) on the exact same statements represented as either tasks or abilities. As a consequence, the absence of a corresponding task or ability does not pose a problem in this study. They are all represented. This would be a greater problem when one was generating task and ability statements as opposed to judging them.

It should also be acknowledged that we have viewed the higher ability and competency ratings as reflecting inflation. Absent some true score, however, it is difficult to definitively establish whether these ratings are truly inflated. To circumvent this problem, we focused on comparisons between task, ability, and competency statements and interpreted differences as inflation for two reasons. First, task judgments are the most common and traditional kind of job information collected. In addition, tasks are the most concrete and observable kind of statement (when compared with ability and competency statements) that require the fewest inferences (Harvey, 1991). This suggests that these kinds of statements would least likely be subject to inflation. Second, the theoretical arguments we forward and the literature on which they are based predict that both ability and competency statements would be rated higher than task statements. The fact that the data conform to these a priori expectations further suggests that we have observed some level of inflation. Notwithstanding these reasons, further research should compare task, ability, and competency ratings with an objective true score.

A final concern with this study is related to how competencies were operationalized. A competency was created for each job component by combining the task and ability statements into a single unit. Some researchers have included more nonobservable characteristics, such as motivation, beliefs, values, and aspects of

self-image in their definition of competencies (Boyatzis, 1982). We operationalized competencies in this way to enable direct comparisons between holistic and decomposed judgments without the usual confounding of item content. Once again, the level of observability of our statements likely provides a conservative test of self-presentation processes. We would expect that broader competency statements with less observable job requirements would be more likely to lead to greater self-presentation by incumbents.

References

Anderson, C. D., Warner, J. L., & Spencer, C. C. (1984). Inflation bias in self-assessment examinations: Implications for valid employee selection. *Journal of Applied Psychology, 69,* 574–580.

Armstrong, J. S., Denniston, W. B., & Gordon, M. M. (1975). The use of the decomposition principle in making judgments. *Organizational Behavior and Human Performance, 14,* 257–263.

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

Boyatzis, R. E. (1982). *The competent manager: A model for effective performance.* New York: Wiley.

Butler, S. K., & Harvey, R. J. (1988). A comparison of holistic versus decomposed rating of Position Analysis Questionnaire work dimensions. *Personnel Psychology, 41,* 761–771.

Cialdini, R. B. (1989). Indirect tactics of image management: Beyond basking. In R. A. Giacalone & P. Rosenfeld (Eds.), *Impression management in the organization* (pp. 45–56). Hillsdale, NJ: Erlbaum.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

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

Cortina, J. M., & Folger, R. G. (1998). When is it acceptable to accept a null hypothesis: No way, Jose? *Organizational Research Methods, 1,* 334–350.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81,* 95–106.

DeNisi, A. S., Cornelius, E. T., & Blencoe, A. G. (1987). Further investigation of common knowledge effects on job analysis ratings. *Journal of Applied Psychology, 72,* 262–268.

DeNisi, A. S., & Shaw, J. B. (1977). Investigation of the uses of self-reports of abilities. *Journal of Applied Psychology, 62,* 641–644.

Dierdorff, E. C., & Wilson, M. A. (2003). A meta-analysis of job analysis reliability. *Journal of Applied Psychology, 88,* 635–646.

Einhorn, H. J. (1972). Expert measurement and mechanical combination. *Organizational Behavior and Human Performance, 7,* 86–106.

Fichman, M. (1999). Variance explained: Why size does not (always) matter. In R. I. Sutton & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 21, pp. 295–331). Stamford, CT: JAI Press.

Fischer, G. W. (1977). Convergent validation of decomposed multiattribute utility assessment procedures for risky and riskless decisions. *Organizational Behavior and Human Performance, 18,* 295–315.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Fleishman, E. A., Wetrogan, L. I., Uhlman, C. E., Marshall-Mies, J. C. (1995). Abilities. In N. G. Peterson, M. D. Mumford, W. C. Borman, P. R. Jeanneret, & E. A. Fleishman (Eds.), *Development of prototype occupational information network content model* (Vol. 1, pp. 1–10). Salt Lake City: Utah Department of Employment Security.

Ganzach, Y., Kluger, A. N., & Klayman, N. (2000). Making decisions from an interview: Expert measurement and mechanical combination. *Personnel Psychology, 53,* 1–20.

Goldberg, L. R. (1971). Five models of clinical judgment: An empirical comparison between linear and nonlinear representations of the human inference process. *Organizational Behavior and Human Performance, 6,* 458–479.

Green, S. B., & Stutzman, T. (1986). An evaluation of methods to select respondents to structured job-analysis questionnaires. *Personnel Psychology, 39,* 543–564.

Harvey, R. J. (1991). Job analysis. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, 2nd ed., pp. 71–163). Palo Alto, CA: Consulting Psychologists Press.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* London: Cambridge University Press.

Kleinmuntz, D. N., Fennema, M. G., & Peecher, M. E. (1996). Conditioned assessment of subjective probabilities: Identifying the benefits of decomposition. *Organizational Behavior and Human Decision Processes, 66,* 1–15.

Leary, M. R. (1995). *Self-presentation: Impression management and interpersonal behavior.* Boulder, CO: Westview Press.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin, 107,* 34–47.

McCormick, E. J., Jeanneret, P. R., & Mecham, R. C. (1972). A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire (PAQ) [Monograph]. *Journal of Applied Psychology, 56,* 347–368.

Meehl, P. E. (1954). *Clinical versus statistical prediction.* Minneapolis: University of Minnesota Press.

Meehl, P. E. (1986). Causes and effects of my disturbing little book. *Journal of Personality Assessment, 50,* 370–375.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Morera, O. F., & Budescu, D. V. (1998). A psychometric analysis of the "divide and conquer" principle in multicriteria decision making. *Organizational Behavior and Human Decision Processes, 75,* 187–206.

Morera, O. F., & Budescu, D. V. (2001). Random error reduction in analytic hierarchies: A comparison of holistic and decompositional decision strategies. *Journal of Behavioral Decision Making, 14,* 223–242.

Morgeson, F. P., & Campion, M. A. (1997). Social and cognitive sources of potential inaccuracy in job analysis. *Journal of Applied Psychology, 82,* 627–655.

Morgeson, F. P., & Campion, M. A. (2000). Accuracy in job analysis: Toward an inference-based model. *Journal of Organizational Behavior, 21,* 819–827.

Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., Fleishman, E. A., Campion, M. A., et al. (2001). Understanding work using the Occupational Information Network (O*NET): Implications for practice and research. *Personnel Psychology, 54,* 1–51.

Prentice, D. A., & Miller, D. T. (1992). When small effects are impressive. *Psychological Bulletin, 112,* 160–164.

Ravinder, H. V. (1992). Random error in holistic evaluations and additive decompositions of multiattribute utility—An empirical comparison. *Journal of Behavioral Decision Making, 5,* 155–167.

Richardson, K. D., & Cialdini, R. B. (1981). Basking and blasting: Tactics of indirect self-presentation. In J. T. Tedeschi (Ed.), *Impression management theory and social psychological research* (pp. 41–53). New York: Academic Press.

Sackett, P. R., Cornelius, E. T., & Carron, T. J. (1981). A comparison of global judgment vs. task oriented approaches to job classification. *Personnel Psychology, 34,* 791–804.

Sanchez, J. I., & Fraser, S. L. (1992). On the choice of scales for task analysis. *Journal of Applied Psychology, 77,* 545–553.

Schlenker, B. R. (1980). *Impression management: The self-concept.* Melbourne, FL: Krieger.

Shepard, R. N. (1964). On subjectively optimum selection among multiattribute alternatives. In M. W. Shelly & G. L. Bryan (Eds.), *Human judgments and optimality* (pp. 257–281). New York: Wiley.

Shippmann, J. S., Ash, R. A., Battista, M., Carr, L., Eyde, L. D., Hesketh, B., et al. (2000). The practice of competency modeling. *Personnel Psychology, 53,* 703–740.

Smith, J. E., & Hakel, M. D. (1979). Convergence among data sources, response bias, and reliability and validity of a structured job analysis questionnaire. *Personnel Psychology, 32,* 677–692.

Ulrich, D., Brockbank, W., Yeung, A. K., & Lake, D. G. (1995). Human resource competencies: An empirical assessment. *Human Resource Management, 34,* 473–495.

Wilson, M. A., Harvey, R. J., & Macy, B. A. (1990). Repeating items to estimate the test–retest reliability of task inventory ratings. *Journal of Applied Psychology, 75,* 158–163.

Received August 29, 2002
Revision received September 9, 2003
Accepted September 16, 2003 ■