THE INCREMENTAL VALIDITY OF INTERVIEW SCORES OVER AND ABOVE COGNITIVE ABILITY AND CONSCIENTIOUSNESS SCORES

JOSE M. CORTINA
George Mason University

NANCY B. GOLDSTEIN
Analysis & Technology, Inc.

STEPHANIE C. PAYNE
George Mason University

H. KRISTL DAVISON
Tulane University

STEPHEN W. GILLILAND
Department of Management and Policy
University of Arizona

Recent research has suggested that scores on measures of cognitive ability, measures of Conscientiousness, and interview scores are positively correlated with job performance. There remains, however, a question of incremental validity: To what extent do interviews predict above and beyond cognitive ability and Conscientiousness? This question was addressed in this paper by (a) conducting meta-analyses of the relationships *among* cognitive ability, Conscientiousness, and interviews, (b) combining these results with predictive validity results from previous meta-analyses to form a "meta-correlation matrix" representing the relationships among cognitive ability, Conscientiousness, interviews, and job performance, and (c) performing 9 hierarchical regressions to examine the incremental validity of 3 levels of structured interviews in best, actual, and worst case scenarios for prediction. Results suggested that interview scores contribute to the prediction of job performance over and above cognitive ability and Conscientiousness to the extent that they are structured, with scores from highly structured interviews contributing substantially to prediction. Directions for future research are discussed.

Research in the last 15 years has shown that cognitive ability, Conscientiousness, and interviews all contribute to prediction of job performance for a variety of jobs (e.g., Barrick & Mount, 1991; Huffcutt &

Arthur, 1994; Hunter & Hunter, 1984). One question that arises from these findings is: To what extent do these predictors contribute to prediction over and above one another (Landy, Shankster, & Kohler, 1994)?

Early research on interviews suggested that they lacked predictive validity (for a review, see Arvey & Campion, 1982). More recent research has indicated that interviews can have higher average validities than researchers originally believed (e.g., Huffcutt & Arthur, 1994; McDaniel, Whetzel, Schmidt, & Maurer, 1994; Wright, Lichtenfels, & Pursell, 1989). As a result, researchers have called for an assessment of the incremental validity of interviews (Arvey & Campion, 1982; Thumin & Barclay, 1980). Specifically, the usefulness of interviews is best assessed in the context of other commonly used selection devices (Tett, Jackson, Rothstein, & Reddon, 1994; Schmidt & Hunter, 1998). With this in mind, the purpose of the present study was to address the issue of incremental validity by combining the results of new and previously published meta-analyses in order to examine the contribution of interview scores to the prediction of job performance over and above cognitive ability and Conscientiousness.

Of course, it is not our intention to suggest that cognitive ability tests, Conscientiousness scales, and interview scores are the only important selection devices. Indeed, Ulrich and Trumbo (1965) and Ash (1981) found reviews of resumes and application materials to be more common than any other selection tools. Nevertheless, we have limited our focus with respect to these three predictors for four reasons. First, in spite of the widespread use of various selection materials, there is typically little research available with respect to their relationships *with other predictors*. Thus, cumulation of such results is not possible for many predictors. Second, there is far less evidence of the generalizability of validity of other selection devices across organizations and jobs. Third, if it were found that structured interviews contributed little to prediction over and above cognitive ability and Conscientiousness, then it would be unlikely that the addition of other variables would change the conclusions of the paper with respect to the lack of usefulness of structured interviews. Finally, inclusion of measures of cognitive ability and Conscientiousness allows an examination of substantive, as well as practical, considerations. This issue is discussed in more detail later.

The remainder of the introduction unfolds as follows. First, the research on the criterion-related validity of cognitive ability tests, Conscientiousness scores, and interview scores is reviewed. Second, because collinearity is important when considering incremental validity, research on the relationships *among* these predictors is summarized.

Criterion-Related Validity

Cognitive ability. Research in the last 15 years has led to the widespread belief that cognitive ability is the single best predictor of many aspects of job and training performance for a variety of jobs (Hunter & Hunter, 1984; Ree, Earles, & Teachout, 1994; Reilly & Chao, 1982; Schmidt, Ones, & Hunter, 1992). For example, Hunter and Hunter (1984) reported a mean validity between cognitive ability and job performance of .45.[1] The success of such tests in predicting performance, together with the relative ease of administration and scoring of such tests, has led many organizations to use them for selection purposes. Thus, an assessment of the usefulness of interviews must include consideration of the role played by tests of cognitive ability.

Conscientiousness. Three recent meta-analyses (Barrick & Mount, 1991; Mount & Barrick, 1995; Tett, Jackson, & Rothstein, 1991) of the relationships between the Big Five personality factors and job performance have found Conscientiousness to have a weak to moderate relationship with job performance across a variety of jobs and settings. Barrick and Mount (1991) found the overall true score correlation (i.e., corrected for range restriction and unreliability in the predictor and criterion) between Conscientiousness (as derived from facets of Conscientiousness, cf. Mount & Barrick, 1995) and performance to be approximately .22, although Tett et al. (1991) found the relationship to be .179. Mount and Barrick (1995) found an uncorrected correlation .18 which, when corrected for criterion unreliability, range restriction, and facet intercorrelation, comes to .267.

There are, of course, personality factors other than Conscientiousness. Nevertheless, in the present paper, we focus on Conscientiousness for two reasons. First, in multiple studies, Conscientiousness has demonstrated a significant relationship with performance in most jobs (Barrick & Mount, 1991; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990), although other dimensions of personality, such as Extraversion, have proven to be significant predictors of performance for some jobs but not others (Barrick & Mount, 1991). Second, results from different meta-analyses are consistent with respect to the validity of Conscientiousness (e.g., Barrick & Mount, 1991; Mount & Barrick, 1995). Due in large part to these findings, much of the recent research in the area of personality has focused on Conscientiousness, with considerably less interest in the other personality dimensions (Barrick, Mount, & Strauss,

---

[1] The more commonly reported .53 value actually represents a combination of cognitive ability and psychomotor ability.

1993; Schmidt et al., 1992; Stewart, Carson, & Cardy, 1996). Thus, we restrict our focus to Conscientiousness.

Interviews. Interviews are one of the most common selection tools used in organizations. For example, Ulrich and Trumbo (1965) reported that 99% of the organizations they surveyed used some form of interview in their selection procedures. Consistent with this finding, Arvey (1979) cited the interview as the most commonly used selection technique. Despite the fact that early research suggested that many forms of interview offered little in the way of performance prediction (Mayfield, 1964; Schmitt, 1976; Wagner, 1949), employers continued to use them as selection tools. Fortunately, recent evidence suggests that modern techniques offer improvements in the reliability and validity of interviews (Schmidt & Hunter, 1998).

There have been many narrative reviews of the validity of selection interviews (e.g., Arvey & Campion, 1982; Mayfield, 1964; Wagner, 1949), and in the last 10 years there have been five meta-analyses of the criterion-related validity of selection interviews (Huffcutt & Arthur, 1994; Marchese & Muchinsky, 1993; McDaniel et al., 1994; Wiesner & Cronshaw, 1988; Wright et al., 1989). These studies have examined the validity of selection interviews as a function of format (individual vs. board), structure, content, and criterion type. In general, these recent studies have found that interview scores are predictive of job performance, and that structured interview scores are more strongly related to job performance than are unstructured interview scores. Specifically, these studies have shown the overall relationship between interview scores and job performance to be between .37 and .47, with the validity for structured interviews lying between .44 and .63 and the validity for unstructured interviews lying between .20 and .35.

The results presented in these five meta-analyses are not consistent with respect to the effect of format on interview validity. For example, McDaniel et al. (1994) found interviews conducted by individuals to have higher validities than did those conducted by panels (.43 vs. .32), whereas Wiesner and Cronshaw (1988) found no difference in validities (.44 in each case). It should also be noted that, although McDaniel et al. found considerable support for the criterion-related validity of interviews, these validity values were lower in those cases where the interviewer had prior knowledge of the cognitive ability test scores of the interviewee. Nevertheless, although the precise nature of the relationship between scores from all forms of interviews and performance is not entirely clear, there is consensus on the general levels of the validities involved.

Given that structure appeared to be an important moderator of the predictive validity of interviews, Huffcutt and Arthur (1994) divided the

structure of an interview into three categories: no standardization of questions and no standardization of scoring (Level 1); slight to moderate constraints on questions and slight to moderate standardization of scoring (Level 2); and moderate to full constraints on both questions and scoring (Level 3–4).[2] Their results were consistent with those of McDaniel et al. (1994) and Wiesner and Cronshaw (1988). McDaniel et al. (1994) broke interviews down by content and found that situational interviews were superior to job related and psychological interviews with respect to the prediction of performance (.50 vs. .39 and .29 respectively), though it should be noted that structure and content may have been confounded in this analysis.

Relationships Among the Predictors

Cognitive ability and Conscientiousness. Theoretically, some of the Big Five factors should be related to cognitive ability (e.g., Openness to Experience, Neuroticism). However, these relationships may in fact be fairly weak (Costa & McCrae, 1992). The relationship between Conscientiousness and cognitive ability, however, is less clear. On the one hand, Conscientiousness and cognitive ability are related to many of the same outcome variables (e.g., job performance, training performance, delinquency). On the other hand, empirical examinations have suggested very little relationship exists between these variables (e.g., McHenry, Hough, Toquam, Hanson, & Ashworth, 1990). In an attempt to resolve some of this conflict, we performed a meta-analysis of the relationship between cognitive ability and Conscientiousness scores.

Cognitive ability and interview scores. There is ample reason to expect that interview scores are related to cognitive ability test scores. For example, Campion and his colleagues (Campion, Campion, & Hudson, 1994; Campion, Pursell, & Brown, 1988) found consistent relationships between scores on structured interviews and scores on cognitive ability tests. In addition, it has been suggested by some researchers that structured interviews are nothing more than poor measures of cognitive ability (Campion et al., 1988; Wright et al., 1989), and their predictive validity derives from their overlap with direct tests of cognitive ability. This, in turn, would suggest that interview scores might not explain performance variance over and above that explained by cognitive ability scores.

At least three studies have not found support for the incremental validity of interviews over and above cognitive ability tests. Campion et al. (1988) investigated the incremental validity of structured interviews

---

[2]Levels 3 and 4 were combined because they show no differences with respect to validity (Huffcutt & Arthur, 1994).

over and above a battery of cognitive aptitude tests when predicting su-
pervisory evaluations. Although the tests explained additional variance
in the performance appraisal beyond that explained by the interview, the
interview did not explain variance in the appraisal beyond that explained
by the tests. Shahani, Dipboye, and Gehrlein (1991) examined the in-
cremental contribution of an interview over and above verbal and math
Scholastic Aptitude Test scores when predicting university admission de-
cisions and freshman grade point average. They found that the interview
was valid, but not incrementally so. Similarly, Walters, Miller, and Ree
(1993) examined the incremental validity of a structured interview over
and above two aptitude measures previously validated as predictors of
pilot training for a pilot-training program. They also found that the in-
terview was valid but not incrementally.

At the same time, some researchers have found support for the in-
cremental validity of interviews over and above cognitive ability tests.
For example, Campion et al. (1994) found that structured interviews
had incremental validity above and beyond a battery of cognitive ability
tests when predicting supervisory ratings of job performance. Similarly,
Latham and Sue-Chan (1999) found a situational interview to predict
grade point average for nursing students above and beyond the Wonder-
lic Personnel Test and a tacit knowledge test. The incremental validity
of interviews may be a function of other variables like interviewer char-
acteristics. For example, Dipboye (1989) found individual differences
between interviewers in the incremental validity of interview scores.

Although there is empirical and theoretical support for the notion
that interview scores are related to cognitive ability test scores, there is
little support for the notion that interview scores are composed only of
cognitive ability and error (Harris, 1989; Huffcutt, Roth, & McDaniel,
1996; Landy, 1976). This literature suggests a moderate, positive rela-
tionship between interview scores and cognitive ability test scores that
does not vary across levels of interview structure. Most recently, Huf-
fcutt et al. (1996) conducted a meta-analysis of this relationship across
two levels of interview structure and found an overall range-restriction
corrected validity of .33 with almost no differences in validities associ-
ated with differences in structure. In an attempt to build upon the Huf-
fcutt et al. meta-analysis, we conducted a literature search for the rela-
tionship between cognitive ability and interview scores that resulted in
the identification of 142 effect size values. This allowed separate analysis
of the cognitive ability-interview relationship for three levels of interview
structure and for studies that specifically measured cognitive ability (as
opposed to "lower-order" abilities such as perceptual speed).

Conscientiousness and interview scores. There are also theoretical
reasons to expect a moderate relationship between Conscientiousness

and interview scores. It has been suggested that interviews are useful
for predicting many of the same components of performance as are per-
sonality tests (Schneider & Schmitt, 1986). Indeed, Schmitt (1976) sug-
gested that interpersonal skill and motivation may be best assessed with
interviews. Further, *all* of the interview studies that we examined that
described the dimensions tapped by their interview questions included
some aspect of Conscientiousness (e.g., initiative, work ethic, accepting
responsibility, Campion et al., 1994; thoroughness, Campion et al., 1988;
organizing and planning, Tziner & Dolan, 1982a). Research on biodata
measures intended to tap constructs similar to Conscientiousness has
yielded positive correlations between interview scores and these biodata
measures (Dalessio & Silverhart, 1994). In addition, more general work
in the personality domain has shown some evidence that interviewers are
able to assess applicant personality accurately (Paunonen, Jackson, &
Oberman, 1987). Other evidence suggests that supervisors, coworkers,
and even customers are able to generate moderately valid assessments of
Conscientiousness (Mount, Barrick, & Strauss, 1994). Finally, it should
be noted that the selection setting is an example of a "strong" situation
in which spontaneous behavior is replaced by more planned behavior.[3]
Thus, factors such as social desirability are likely to affect interview be-
havior and Conscientiousness item responses similarly, causing the cor-
relation between these two sets of variables to increase.

Given that interviews are often designed to assess some aspect of
Conscientiousness, that interviewers are generally able to assess the
Conscientiousness of interviewees, and that there are aspects of the se-
lection situation that should affect Conscientiousness scores and inter-
view scores similarly, there should be at least a moderate relationship
between these variables. Because of the limited number of studies that
have examined the relationship between Conscientiousness and inter-
view scores, it was not possible to analyze the correlations separately
by amount of interview structure. Therefore, we performed one meta-
analysis of these values and used functions of the resulting value in the
hierarchical regression analyses.

Before moving on, it should be noted that Schmidt and Hunter (1998)
recently published a paper examining the incremental validity issue with
regard to a number of predictors. Specifically, these authors examined
the contribution to prediction made by a variety of individual predictors
over and above cognitive ability. With respect to Conscientiousness and
interviews, these authors suggested the following: (a) Conscientiousness
contributes to prediction over and above cognitive ability, (b) structured
interview scores contribute to the prediction over and above cognitive

---

[3]Thanks to an anonymous reviewer for pointing out this aspect of the selection situation.

ability, and (c) unstructured interview scores contribute only slightly to prediction over and above cognitive ability.

The present study builds on these findings in several ways. First, whereas Schmidt and Hunter (1998) relied on assumed relationships between cognitive ability and Conscientiousness, this relationship is estimated empirically in the present study. Second, in the present study, the contribution of interview scores to prediction is assessed over and above cognitive ability and Conscientiousness *in conjunction*. This has both practical and theoretical benefits. The practical benefit is that the contribution of interviews, which are more expensive to develop and administer (Hogan & Hogan, 1992), can be examined relative to *two* commonly cited but less expensive predictors, cognitive ability and Conscientiousness. The theoretical benefit is that consideration of both cognitive ability and Conscientiousness allows us to better understand the substantive determinants of interview scores. Third, the present study incorporates the rigorous definition of interview structure developed by Huffcutt and Arthur (1994), thus allowing more precise consideration of the issue.

Method

In order to examine the incremental validity of structured interviews on job performance over and above cognitive ability and Conscientiousness, six correlations are required: three validities and the three correlations among the three predictors. As depicted in Table 1, values *a, b*, and *c*, which represent the relationships between cognitive ability and Conscientiousness, cognitive ability and interview scores, and Conscientiousness and interview scores were generated from meta-analyses conducted as part of the present study. Values for *d, e*, and *f* representing the relationships between cognitive ability and job performance, Conscientiousness and job performance, and interview scores and job performance were taken from three previous meta-analyses. The remainder of the method section is organized as follows. First, the reviews from which criterion-related validity values were taken are described, and the values themselves are offered. Second, the general meta-analytic methods used to generate correlations among predictors are described. Third, the analyses used to establish incremental validity are described.

Values from Previous Studies

We used .45 from Hunter and Hunter (1984) to represent the relationship between cognitive ability tests and job performance. Given the consistency of criterion-related validities across different measures of cognitive ability (Hunter & Hunter, 1984), we felt justified in using this

TABLE 1

*Proposed "Meta-Correlation Matrix" for Cognitive Ability Scores,
Conscientiousness Scores, Interview Scores, and Job Performance*

	1	2	3	4
1. Cognitive ability	–			
2. Conscientiousness	a	–		
3. Interview score	b	c	–	
4. Job performance	d	e	f	–

Note: a, b, and c are generated in the present study and d, e, and f are taken from previous meta-analytic work.

value to represent the criterion-related validity of cognitive ability tests in general.

We used .267 from Mount and Barrick (1995) to represent the relationship between Conscientiousness and job performance. This value is simply the .31 value reported by Mount and Barrick uncorrected for predictor unreliability.

Finally, we adopted Huffcutt and Arthur's (1994) approach to structure of interviews and therefore thought it most appropriate to use their validity values (by structure level) in our regression analyses. Specifically, the validities corrected for criterion unreliability and range restriction were .20, .35, and .56 for structure Levels 1, 2, and 3–4, respectively.

All three of the meta-analyses from which we borrowed validities are seminal pieces in their respective areas. They are all comprehensive in that they conducted extensive searches to locate studies. They used similar inclusion criteria. For example, all three meta-analyses included unpublished studies. A summary of their results is provided in Appendix A. Given that we intended to combine our results with these values, we conducted our meta-analyses similarly.

General Meta-Analytic Methods

We conducted extensive searches for studies that reported relationships between any of the three predictors of interest. Manual and computer searches involving a variety of journals and keywords were combined with a review of the studies included in the meta-analyses of Barrick and Mount (1991), Huffcutt and Arthur (1994), McDaniel et al. (1994), Tett et al. (1991), and Wiesner and Cronshaw (1988). This search yielded 178 coefficients from 24 published papers, 2 unpublished papers, and 8 technical reports. Citations for these studies and the values taken from them are contained in Appendix B.

Studies were included in a meta-analysis if they reported correlations between any two of the three predictors under consideration or the in-

formation necessary to compute them. Three of the authors coded studies for possible inclusion in one or more of the meta-analyses. Previous descriptions of cognitive ability, the coding scheme used in Barrick & Mount (1991), and the coding scheme used by Huffcutt and Arthur (1994) were used by three of the authors to code studies with respect to measures of cognitive ability, Conscientiousness, and interview structure respectively. Perfect agreement was achieved for all three types of measure.

The meta-analyses involved the calculation of sample-size weighted mean correlations, sample-size weighted variance estimates from the observed study results, and values based on residual variability using procedures described in Hunter and Schmidt (1990). Because these analyses involved predictors and because organizations must use selection instruments with error, the correlations among these predictors were not corrected for unreliability and facet intercorrelations (Ones, Viswesvaran, & Reiss, 1996). Thus, the ρs are meant to represent true validities as opposed to true score correlations. In addition, because so little information was available with respect to range restriction, the values reported in previous meta-analyses were used to correct correlations for range restriction. Specifically, the mean ratio of incumbent standard deviation to applicant standard deviation reported in Huffcutt and Arthur (.74, 1994) was used to correct correlations involving interviews and the value of .67 reported in Hunter and Hunter (1984) was used to correct correlations involving tests of cognitive ability. The final steps involved the calculation of the proportion of variance in observed correlations due to sampling error and the calculation of confidence intervals and credibility values associated with ρs. It should be noted that although conceptual replications do not bias estimates of mean effect size, they can affect variance estimates (Hunter & Schmidt, 1990, p. 453). Thus, whereas all individual values were used separately to generate mean correlation values, conceptual replication-based values were collapsed within studies, and single mean values from each study were used to generate variance values.

Incremental Validity Analyses

Hierarchical regression was used to examine incremental validity of interview scores. For each of the hierarchical regression analyses presented, performance was regressed onto cognitive ability and Conscientiousness (Step 1) followed by the interview (Step 2) for each of the three levels of interview structure. Entering the interview last, however, somewhat stacks the deck against it. In addition, there is reason to believe that some of the correlations among the variables considered here

may vary somewhat across situations. Therefore, we examined the incremental validity of interview scores for varying relationships between cognitive ability and performance and Conscientiousness and performance. This was done by using the standard deviation of ρ values associated with these relationships.

According to Hunter and Hunter (1984) and Schmitt, Gooding, Noe, and Kirsch (1984), the SD_ρ of the cognitive ability-job performance relationship is approximately .14. According to Mount and Barrick (1995) the SD_ρ of the Conscientiousness-job performance relationship is approximately .09.[4] If we add and subtract these values from their respective true validities, we can create matrices that represent "best" and "worst" case scenarios for prediction using the interview. In other words, the interview has the greatest opportunity to account for significant variance when the true validities of the other two predictors are low (i.e., .45 –.14 for cognitive ability and .267 –.09 for Conscientiousness). Likewise, the interview has less of a chance to account for significant variance in job performance when the true validities of the other two predictors are high (i.e., .45 + .14 for cognitive ability and .267 + .09 for Conscientiousness). In an effort to examine the various possibilities, we conducted nine regressions: best, actual, and worst case scenario for prediction with the interview for all three levels of structure.

Results

The results section of the paper unfolds as follows. First, the results of the meta-analysis of the relationship between cognitive ability test scores and Conscientiousness scores are presented and discussed. Second, the results of meta-analyses of the relationships between cognitive ability and interview scores with three levels of structure are presented and discussed. Third, the results of the meta-analysis of the relationship between Conscientiousness scores and interview scores are presented and discussed. Finally, the results of regression-based incremental validity analyses are presented and discussed.

---

[4] Because the amount of variability attributable to between study differences in predictor unreliability was not reported in Mount & Barrick (1995), we were unable to "uncorrect" the population SD value for predictor unreliability. Nevertheless, given that little variance in effect sizes is typically attributable to artifacts other than sample size (Hunter & Schmidt, 1990), and given that predictor unreliability was only one of several such artifacts, it is unlikely that consideration of predictor unreliability would have changed the estimate of the population SD to two decimal places.

Meta-Analysis of the Relationship Between Conscientiousness
Scores and Cognitive Ability

The literature search resulted in 27 coefficients representing the relationship between Conscientiousness and cognitive ability, and a total sample size of 53,144. As can be seen in Table 2, there is a weak relationship between scores on cognitive ability tests and scores on measures of Conscientiousness. Even after correction for range restriction, this relationship remains near zero (ρ = .075) and the 95% credibility value falls below zero, indicating a magnitude of relationship that is not clearly distinguished from zero. A relatively small portion of the variance in observed correlations associated with this relationship was attributable to sampling error and range restriction. Given that artifacts other than sample size typically account for relatively little variance, this suggests that moderator variables may have some impact on observed correlations between Conscientiousness scores and cognitive ability test scores. Nevertheless, potential moderators notwithstanding, these results suggest a relationship between these two variables that is very weak in most situations.

Meta-Analysis of the Relationship Between Cognitive Ability
and Interview Scores

The literature search resulted in 7 coefficients for Level 1 interviews, 76 coefficients for Level 2 interviews, and 59 coefficients for Level 3–4 interviews. The total sample sizes for these three analyses were 1,366; 8,042; and 11,885, respectively. The uncorrected correlations between cognitive ability and interviews at each of the three levels of structure were .041, .202, and .181, respectively. For some reason, many of the studies included in the Level 3–4 analysis used measures of cognitive ability that were oriented toward memory capacity, perceptual speed, and other capacities that are not traditionally considered to be pure measures of cognitive ability. (This was not the case for the other levels of structure.) Thus, the correlation involving Level 3–4 interviews was reestimated without these studies. This analysis of the remaining 43 coefficients yielded a value of .219 (N = 8,049). The results for all four analyses are presented in Table 2.

As can be seen in Table 2, there is a weak relationship between scores on unstructured interviews and cognitive ability. Even after correction for range restriction, this relationship remains near zero (ρ = .055). The relationship between scores on interviews with Level 2 structure and cognitive ability was considerably higher. After correction for range restriction, this value (ρ = .253) is slightly smaller than those reported in Huff-

TABLE 2

Results of Meta-Analyses for the Cognitive Ability–Conscientiousness, Cognitive Ability–Interview Score, and Conscientiousness–Interview Score Relationships

Variable	N	Observed mean r^a	S_r^2	S_e^2	ρ	S_ρ^2	% of variance attributed to sampling error	95% credibility value
Relationships with cognitive ability								
1. Level 1 interview	1,366	.041	.0040	.0043	.055	−.0003	107	.055
2. Level 2 interview	8,042	.202	.0290	.0038	.253	.039	13	−.134
3. Level 3–4 interview^b	8,049	.219	.0139	.0035	.270	.016	25	.022
4. Conscientiousness	53,144	.052	.0050	.0011	.075	.0081	22	−.101
Relationship between Conscientiousness and interviews								
	966	.206	.0038	.0072	.258	−.0033	187	.258

^{a} Although the mean r and ρ values are based on the full datasets, variance computations were based on sets of values that included correlations collapsed within studies to account for conceptual replications as suggested by Hunter & Schmidt (1990).

^{b} The values in this row are based on the group of coefficients in which perceptual speed measures were excluded.

cutt et al. (1996). If we focus on the estimate of the relationship between Level 3–4 interviews and the purer measures of cognitive ability, we see that this relationship ($\rho = .27$) is stronger than those involving interviews with less structure, but is still slightly below the values reported in Huffcutt et al. (1996).

Though all of the variance in correlations associated with interviews with Level 1 structure was attributable to sampling error and range restriction, relatively small portions of the variance in observed correlations associated with Level 2 and Level 3–4 structure were so attributable. Thus, there is evidence of the existence of moderators that have an impact on observed correlations between interview scores and cognitive ability test scores. It is worth noting that removal of a single outlier from the Level 2 analysis reduced the residual variance value to below zero.

Meta-analysis of the Relationship between Conscientiousness and Interview Scores

The literature search resulted in nine coefficients and a total sample size of 966. All nine coefficients were based on Level 3–4 interviews. As a result, we were only able to calculate a meta-analytic value between Conscientiousness and Level 3–4 interviews. As can be seen in Table 2, there is a moderate relationship between interview scores and scores on measures of Conscientiousness. After correction for range restriction, correlation between these variables is .258. All of the variance in observed correlations associated with this relationship was attributable to sampling error and range restriction.

Incremental Validity

Based on the results of the analyses presented above, we can now fill in values for a, b, and c in Table 1. Specifically, value a can be filled in with the results from the present meta-analysis of the relationship between Conscientiousness scores and cognitive ability (.075). In addition, we can include as values for b the range corrected Level 1 correlation from the present study (.055), the range corrected Level 2 value from the present study (.253), and the range corrected Level 3–4 value from the present study (.270). Finally, value c can be filled in for Level 3–4 interviews with the results from the present meta-analysis of the relationship between Conscientiousness scores and interview scores (.258). Of course, this value is based only on the Level 3–4 interviews. Thus, the values used for the Conscientiousness/interview relationship for Level

TABLE 3

Meta-Correlation Matrices to be Used in Regression Analyses

Variable	1	2	3	4
1. Cognitive ability	–			
2. Conscientiousness	.075	–		
3. Level 1 interview	.055	.092	–	
4. Job performance	$.450^a$	$.267^b$	$.200^c$	–
1. Cognitive ability	–			
2. Conscientiousness	.075	–		
3. Level 2 interview	.253	.161	–	
4. Job performance	$.450^a$	$.267^b$	$.350^c$	–
1. Cognitive ability	–			
2. Conscientiousness	.075	–		
3. Level 3–4 interview	.270	.258	–	
4. Job performance	$.450^a$	$.267^b$	$.560^c$	–

[a] .310 (i.e., .45−.14) and .590 (i.e., .45 + .14) were substituted for this validity for best and worst case scenario, respectively.

[b] .177 (i.e., .267−.09) and .357 (i.e., .267 + .09) were substituted for this validity for best and worst case scenario, respectively.

[c] From Huffcutt and Arthur (1994).

1 and 2 interviews reflected the ratio of relevant validities.[5] For example, because the ratio of validities between Level 1 and Level 3–4 interviews was $\frac{.2}{.56}$ = .357, the value used to represent the Level 1 interview-Conscientiousness relationship was .258 × .357 = .092. This issue is discussed later. At this point, we will simply foreshadow our results by mentioning that conclusions with respect to incremental validity would have been very much the same even if different values has been used for the relationship between Conscientiousness and interviews with less structure. In any case, the collinearity values can then be added to the values from previous meta-analyses to form the matrices presented in Table 3.

In the "actual" case scenario, cognitive ability and Conscientiousness scores account for a total of 26% of the variance in performance. In the best and worst cases, they account for 12% and 45% of the variance in performance.

As can be seen in Table 4, the incremental validity for Level 1 interview scores ranges from .9 to 2.2 % of the variance in job performance. When the actual validities for cognitive ability and Conscientiousness are entered, Level 1 interviews only account for an additional 1.5% of the variance. Thus, unstructured interviews would very rarely contribute meaningfully to prediction.

---

[5]Thanks to an anonymous reviewer for this suggestion.

TABLE 4

Hierarchical Regression Results

Variable	β	ΔR^2	R^2
Best case			
1. Cognitive ability +	.298		
Conscientiousness	.155	.120	.120
2. Level 1 interview	.154	.022	.142
2. Level 2 interview	.267	.062	.182
2. Level 3–4 interview	.505	.222	.342
Actual value			
1. Cognitive ability +	.432		
Conscientiousness	.235	.257	.257
2. Level 1 Interview	.124	.015	.272
2. Level 2 Interview	.205	.037	.294
2. Level 3–4 Interview	.440	.169	.426
Worst case			
1. Cognitive ability +	.566		
Conscientiousness	.315	.446	.446
2. Level 1 interview	.094	.009	.455
2. Level 2 interview	.143	.018	.464
2. Level 3–4 interview	.375	.123	.569

As one might expect, results for Level 2 interview scores were a little more optimistic. Level 2 interview scores explain an additional 1.8 to 6.2% of the variance in job performance. When the actual validities for cognitive ability and Conscientiousness are entered, Level 2 interviews account for an additional 3.7% of the variance.

Level 3–4 interviews offered the most incremental validity. These interviews explained an additional 12.3%–22.2% of the variance in job performance over and above cognitive ability and Conscientiousness. When the actual validities for cognitive ability and Conscientiousness are entered, Level 3–4 interviews account for an additional 16.9% of the variance above and beyond the first two predictors.

Overall, these analyses suggest one very important conclusion. The contribution to prediction made by interview scores depends almost entirely on the amount of structure in the interview such that unstructured interviews contribute very little, even under ideal circumstances, and interviews high in structure contribute as much, if not more, to prediction as do cognitive ability scores. Even in "worst" case scenarios, interviews high in structure contribute substantially to prediction.

General Discussion

The purpose of this paper was to investigate the incremental validity of structured interview scores over and above cognitive ability and Con-

scientiousness scores in the prediction of job performance. This was ac-
complished by conducting meta-analyses of the relationships among cog-
nitive ability test scores, Conscientiousness scores, and interviews scores
and combining the results of these analyses with the results of previous
meta-analyses to produce a matrix that could be submitted to multiple
regression analysis. It was concluded that cognitive ability and Consci-
entiousness scores as a pair contribute consistently to the prediction of
performance, and interviews contribute to prediction to the extent that
they are structured, with scores from highly structured interviews con-
tributing substantially to prediction.

Previous research on the incremental validity of interviews has been
mixed. Similar to Campion et al. (1994) and Latham and Sue-Chan
(1999), our study provides support for the incremental validity of struc-
tured interviews over and above cognitive ability. More specifically, as
the interview increased in structure, it accounted for more variance in
job performance. This finding is at odds with some of the previous work
in the area. One possible explanation for this discrepancy is that the
studies reporting findings that interviews did not predict above and be-
yond cognitive ability measures (e.g., Campion et al., 1988; Shahani et
al., 1991; Walters et al., 1993) involved interviews that, either by design
or in spite of it, had lower levels of structure. In fact, all of these studies
were published before Huffcutt and Arthur (1994) published their sem-
inal piece distinguishing among the three levels of interview structure.

Our study also found support for incremental validity of structured
interviews over and above measures of Conscientiousness. This finding
is a unique contribution to the literature as this has not been previously
explored. Although it has been suggested that each of these predictors
capture similar constructs (Schneider & Schmitt, 1986), our results sug-
gest that structured interviews tap more than Conscientiousness and cog-
nitive ability.

Although the present study addresses some important questions,
other questions remain. For example, one limitation of the present
study was the small number of values available for assessment of the
Conscientiousness–interview score relationship. Future research might
endeavor to examine this relationship further using different levels of
interview structure, although the lack of incremental validity associated
with less structured interviews suggests that the practical implications of
such research would be limited.

Another unanswered question is why do structured interviews ac-
count for more variance than unstructured interviews? Some researchers
have explained the effectiveness of structured interviews in terms of
both method and content including standardization, multiple interview-
ers, content validity, and training of interviewers (Campion et al., 1988).

Structured interviews appear to be less vulnerable to interviewer biases including attributional, recall, and rating biases (Dipboye & Gaugler, 1993). Consistent with this finding, structured interviews have higher levels of interrater reliability than unstructured interviews (Conway, Jako, & Goodman, 1995; McDaniel et al., 1994). Other attributes not withstanding, the greater legal defensibility of structured interviews (Williamson, Campion, Malos, Roehling, & Campion, 1997) makes them an attractive alternative to unstructured interviews.

Given that structured interviews account for additional variance above and beyond cognitive ability and Conscientiousness, the next question is what are they measuring? Although there has been some effort to assess the dimensions tapped by interviews (e.g., Landy, 1976), the nature of the information provided by interviews is still unclear. It may be that some individual characteristics are more amenable to assessment by interview than others, though relatively little work has been done to identify such characteristics.

Another interesting possibility is that structured interviews tap job knowledge. Job knowledge has been shown to contribute to the prediction of performance over and above cognitive ability (Ree, Carretta, & Doub, 1998/1999; Ree, Carretta, & Teachout, 1995). Perhaps part of the criterion space covered by structured interviews is related to job knowledge.[6] Along these same lines, interviews may be most useful in tapping fine-grained aspects of experience that are more difficult to uncover with traditional measures of experience such application blanks. For example, Quiñones, Ford, & Teachout (1995) suggested that measures of experience can be classified by specificity (organizational, job, and task). It may be that the application blank is useful for measuring experience at the organizational and job level, and interviews are useful for measuring experience at the task level (e.g., amount of time spent performing certain kinds of tasks, number of times the tasks were performed, complexity of the tasks). This task-level experience might then translate into job knowledge and a variety of other attributes that directly influence performance.

Future research might also attempt to quantify the various individual, organizational, and societal costs/benefits associated with these predictors (e.g., diversity, test development/administration/scoring, litigation, etc.) and compare them to the level of expected job performance associated with their usage. This research suggests that the additional cost required to develop and implement a structured interview may be worthwhile.

---

[6]Thanks to an anonymous reviewer for suggesting this possibility.

Finally, future research might endeavor to discover what if any situations lead to the contribution to performance prediction of Conscientiousness and/or other personality variables over and above cognitive ability and structured interviews. Existing Conscientiousness measures tend to require little in the way of administration and scoring costs. Nevertheless, even these minimal costs are unjustified unless a contribution to prediction can be documented.

The issues raised here are important if we are to increase our understanding of the ways that our predictors combine to predict performance. The present study is a first step that suggests that, other issues notwithstanding, cognitive ability and Conscientiousness scores as a set are generally useful in the prediction of performance, and interview scores contribute over and above these predictors to the extent that they are structured.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analyses.

Arvey RD. (1979). Unfair discrimination in the employment interview: Legal and psychological aspects. *Psychological Bulletin, 86*, 736–765.

Arvey RE, Campion JE. (1982). The employment interview: A summary and review of recent research. PERSONNEL PSYCHOLOGY, *35*, 281–322.

Ash RA. (1981). Comparison of four approaches to the evaluation of job applicant training and work experience. *Dissertation Abstracts International, 42*, 4606B. (University Microfilms No. DE082-07909).

*Baehr ME, Froemel E. (1977). The arrow dot test as a predictor of police officer performance. *Perceptual and Motor Skills, 45*, 683–693.

Barrick MR, Mount MK. (1991). The Big Five personality dimensions and job performance: A meta-analysis. PERSONNEL PSYCHOLOGY, *44*, 1–26.

Barrick MR, Mount MK, Strauss JP. (1993). Conscientiousness and performance of sales representatives: Test of the mediating effect of goal setting. *Journal of Applied Psychology, 78*, 715–722.

*Berkley S. (1984). VII *Validation Report Corrections Officer Trainee*. Harrisburg, PA: Commonwealth of Pennsylvania, State Civil Service Commisssion.

*Borman WC, White LA, Pulakos ED, Oppler SH. (1991). Models of supervisory job performance ratings. *Journal of Applied Psychology, 76*, 863–872.

*Campion MA, Campion JE, Hudson JP. (1994). Structured interviewing: A note on incremental validity and alternative question types. *Journal of Applied Psychology, 79*, 998–1002.

*Campion MA, Pursell ED, Brown BK. (1988). Structured interviewing: Raising the psychometric properties of the employment interview. PERSONNEL PSYCHOLOGY, *41*, 25–42.

Conway JM, Jako RA, Goodman DF. (1995). A meta-analysis of interrater internal consistency reliability of selection interviews. *Journal of Applied Psychology, 80*, 565–579.

Costa PT Jr, McCrae RR. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13*, 653–665.

*Crant JM. (1995). The proactive personality scale and objective job performance among real estate agents. *Journal of Applied Psychology, 80*, 532–537.

Dalessio AT, Silverhart TA. (1994). Combining biodata test and interview information: Predicting decisions and performance criteria. PERSONNEL PSYCHOLOGY, *47*, 303–316.

*Day DA, Silverman SB. (1989). Personality and job performance: Evidence of incremental validity. PERSONNEL PSYCHOLOGY, *42*, 25–36.

*Delery JE, Wright PM, Tolzman K. (1992, April). *Employment tests and the situational interview: A test of incremental validity*. Paper presented at the 7th Annual Conference of the Society for Industrial and Organizational Psychology, Montreal, Quebec, Canada.

*Dicken C. (1969). Predicting the success of Peace Corps community development workers. *Journal of Consulting and Clinical Psychology, 33*, 597–606.

Dipboye RL. (1989).Threats to the incremental validity of interviewer judgments. In Eder RW, Ferris GR (Eds.), *The employment interview: Theory, research, and practice* (pp. 45–60). Newbury Park, CA: Sage.

Dipboye RL, Gaugler BB. (1993). Cognitive and behavioral processes in the selection interview. In Schmitt N, Borman WC (Eds.), *Personnel selection in organizations* (pp. 135–170). San Francisco: Jossey-Bass.

*Dipboye RL, Gaugler B, Hayes T. (1990, April). *Differences among interviewers in the incremental validity of their judgments*. Paper presented at the 5th Annual Conference of the Society for Industrial and Organizational Psychology, Miami, FL.

*Exxon. (1973). *Validity of a selection program for process trainees*. Houston, TX: Exxon Research and Development.

*Exxon. (1974). *Validity of a selection program for commercial drivers*. Houston, TX: Exxon Research and Development.

*Exxon. (1978). *Selecting laboratory technicians*. Houston, TX: Exxon Research and Development.

*Friedland D. (1973). *Selection of police officers*. Los Angeles, CA: City of Los Angeles, Personnel Department.

*Friedland D. (1976). *Section 11—City of Los Angeles Personnel Department* (Junior Administrative Assistant Validation Study.) Los Angeles, CA: City of Los Angeles, Personnel Department.

*Friedland D. (1980). *A predictive study of selected tests for police officer selection*. Los Angeles, CA: City of Los Angeles, Personnel Department.

*Gully SM, Payne SC, Kiechel KL, Whiteman JK. (2000). *The impact of error training and individual differences on training outcomes: An attribute-treatment interaction perspective*. Manuscript submitted for publication.

*Gully SM, Phillips JM, Beaubien JM, Payne SC. (1998, April). Effects of individual differences in goal orientation and self-regulatory tendencies on learning. In Gully SM, Mathieu JE (Chairs), *Individual differences, learning, motivation, and training outcomes*. Symposium presented at the 13th Annual Conference of the Society for Industrial and Organizational Psychology, Dallas, TX.

*Handyside JD, Duncan DC. (1954). Four years later: A follow-up of an experiment in selecting supervisors. *Occupational Psychology, 28*, 9–23.

*Hansen CP. (1989). A causal model of the relationship among accidents, biodata, personality, and cognitive factors. *Journal of Applied Psychology, 74*, 81–90.

Harris MM. (1989). Reconsidering the employment interview: A review of recent literature and suggestions for future research. PERSONNEL PSYCHOLOGY, *42*, 791–726.

Hogan R, Hogan J. (1992). *The Hogan Personality Inventory manual*. Tulsa, OK: Hogan Assessment Systems.

Hough LM, Eaton NK, Dunnette MD, Kamp JP, McCloy RA. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology, 75*, 581–595.

Huffcutt AI, Arthur W. (1994). Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology, 79*, 184–190.

Huffcutt AI, Roth PL, McDaniel MA. (1996). A meta-analytic investigation of cognitive ability in employment interview evaluations: Moderating characteristics and implications for incremental validity. *Journal of Applied Psychology, 81,* 459–473.

Hunter JE, Hunter RF. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96,* 72–98.

Hunter JE, Schmidt FL. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

*Judge TA, Higgins CA, Thoresen CJ, Barrick MR. (1999). The Big Five personality traits, general mental ability, and career success across the life span. PERSONNEL PSYCHOLOGY, 52, 621–652.

Landy FJ. (1976). The validity of the interview in police officer selection. *Journal of Applied Psychology, 61,* 193–198.

Landy FJ, Shankster LJ, Kohler SS. (1994). Personnel selection and placement. In Porter LW, Rosenzweig MR (Eds.), *Annual review of psychology* (Vol. 45, pp. 261–296). Palo Alto, CA: Annual Reviews.

Latham GP, Sue-Chan C. (1999). A meta-analysis of the situational interview: An enumerative review of reasons for its validity. *Canadian Psychology, 40* (1), 56–67.

*Lopez FM. (1966). Current problems in test performance of job applicants: I. PERSONNEL PSYCHOLOGY, 19, 10–18.

Marchese MC, Muchinsky PM. (1993). The validity of the employment interview: A meta-analysis. *International Journal of Selection and Assessment, 1,* 18–26.

*Martocchio JJ, Judge TA. (1997). Relationship between Conscientiousness and learning in employee training: Mediating influences of self-deception and self-efficacy. *Journal of Applied Psychology, 82,* 764–773.

Mayfield EC. (1964). The selection interview: A reevaluation of published research. PERSONNEL PSYCHOLOGY, 17, 239–260.

McDaniel MA, Whetzel DL, Schmidt FL, Maurer SD. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology, 79,* 599–616.

*McHenry JJ, Hough LM, Toquam JL, Hanson ML, Ashworth S. (1990). Project A validity results: The relationship between predictor and criterion domains. PERSONNEL PSYCHOLOGY, 43, 335–354.

*Motowidlo SJ, Carter GW, Dunnette MD, Tippins N, Werner S, Burnett JR, Vaughn MJ. (1992). Studies of the structured behavioral interview. *Journal of Applied Psychology, 77,* 571–587.

*Motowidlo SJ, Schmit MJ. (1996). *Performance assessment and interview procedures for store manager and store associate positions.* Technical Report. Human Resource Research Center, University of Florida, Gainesville, FL.

Mount MK, Barrick MR. (1995). The Big Five personality dimensions: Implications for research and practice in human resources management. In Ferris G (Ed.), *Research in personnel and human resources management* (pp. 153–200) Greenwich, CT: JAI Press.

Mount MK, Barrick MR, Strauss JP. (1994). Validity of observer ratings of the Big Five personality factors. *Journal of Applied Psychology, 79,* 272–280.

*Neuman GA, Wright J. (1999). Team effectiveness: Beyond skills and cognitive ability. *Journal of Applied Psychology, 84,* 376–389.

Ones DS, Viswesvaran C, Reiss AD. (1996). The role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology, 81,* 660–679.

Paunonen SV, Jackson DN, Oberman SM. (1987). Personnel selection decisions: Effects of applicant personality and the letter of reference. *Organizational Behavior and Human Decision Processes, 40,* 96–114.

*Phillips JM, Gully SM. (1997). Role of goal orientation, ability, need for achievement, and locus of control in the self-efficacy and goal-setting process. *Journal of Applied Psychology, 82*, 792–802.

*Pulakos ED, Schmitt N. (1995). Experience-based and situational interview questions: Studies of validity. PERSONNEL PSYCHOLOGY, *48*, 289–307.

*Pulakos ED, Schmitt N. (1996). An evaluation of two strategies for reducing adverse impact and their effects on criterion-related validity. *Human Performance, 9*, 241–258.

*Pulakos ED, Schmitt N, Whitney D, Smith M. (1996). Individual differences in interviewer ratings: The impact of standardization, consensus discussion, and sampling error on the validity of a structured interview. PERSONNEL PSYCHOLOGY, *49*, 85–102.

Quiñones MA, Ford JK, Teachout MS. (1995). The relationship between work experience and job performance: a conceptual and meta-analytic review. PERSONNEL PSYCHOLOGY, *48*, 887–910.

Ree MJ, Carretta TR, Doub TW. (1998/1999). A test of three models of the role of g and prior job knowledge in the acquisition of subsequent job knowledge in training. *Training Research Journal, 4*, 135–150.

Ree MJ, Carretta TR, Teachout MS. (1995). Role of ability and prior job knowledge in complex training performance. *Journal of Applied Psychology, 80*, 721–730.

Ree MJ, Earles JA, Teachout MS. (1994). Predicting job performance: Not much more than g. *Journal of Applied Psychology, 79*, 518–524.

*Reeb M. (1969). A structured interview for predicting military adjustment. *Occupational Psychology, 43*, 193–199.

Reilly RR, Chao GT. (1982). Validity and fairness of some alternative employee selection procedures. PERSONNEL PSYCHOLOGY, *35*, 1–62.

Roth PL, Campion JE. (1992). An analysis of the predictive power of the panel inteview and pre-employment tests. *Journal of Organizational and Occupational Psychology, 65*, 51–60.

*Ryan AM, Ployhart RE, Gregaras GJ, Schmit MJ. (1998). Test preparation programs in selection contexts: Self-selection and program effectiveness. PERSONNEL PSYCHOLOGY, *51*, 599–621.

Schmidt FL, Hunter JE. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262–274.

Schmidt FL, Ones DS, Hunter JE. (1992). Personnel selection. *Annual Review of Psychology, 43*, 627–670.

*Schmit M. (1996). *Development of a situational interview and personality test for the selection of store associates.* Technical Report. Human Resource Research Center, University of Florida, Gainesville, FL.

Schmitt N. (1976). Social and situation determinants of interview decisions: Implications for the employment interview. PERSONNEL PSYCHOLOGY, *29*, 79–101.

Schmitt N, Gooding RZ, Noe RA, Kirsch M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. PERSONNEL PSYCHOLOGY, *37*, 407–422.

Schneider B, Schmitt N. (1986). *Staffing organizations* (2nd ed.). Glenview, IL: Scott, Foresman.

Shahani C, Dipboye RL, Gehrlein TM. (1991). The incremental validity of an interview to college admissions. *Educational and Psychological Measurement, 51*, 1049–1061.

Stewart GL, Carson KP, Cardy RL. (1996). The joint effect of Conscientiousness and self-leadership training on employee self-directed behavior in a service setting. PERSONNEL PSYCHOLOGY, *49*, 143–164.

Tett RP, Jackson RN, Rothstein M. (1991). Personality measures as predictors of job performance: A meta-analytic review. PERSONNEL PSYCHOLOGY, 44, 742.

Tett RP, Jackson DN, Rothstein M, Reddon JR. (1994). Meta-analysis of personality-job performance relations: A reply to Ones, Mount, Barrick, and Hunter (1994). PERSONNEL PSYCHOLOGY, 47, 157–172.

Thumin FJ, Barclay AG. (1980). Sequential validity: A new model of evaluating judgmental prediction in personnel selection. *Professional Psychology, 11*, 48–54.

*Tubiana JH, Ben-Sitakhar G. (1982). An objective group questionnaire as a substitute for a personal interview in prediction of success in military training in Israel. PERSONNEL PSYCHOLOGY, 35, 349–357.

Tziner A, Dolan S. (1982a). Evaluation of a traditional selection system in predicting success of females in officer training. *Journal of Occupational Psychology, 55*, 269–275.

*Tziner A, Dolan S. (1982b). Validity of an assessment center for identifying future female officers in the military. *Journal of Applied Psychology, 67*, 728–736.

Ulrich L, Trumbo D. (1965). The selection interview since 1949. *Psychological Bulletin, 63*, 100–116.

Walters LC, Miller MR, Ree MR. (1993). Structured interviews for pilot selection: No incremental validity. *The International Journal of Aviation Psychology, 3*(1), 25–38.

Wagner R. (1949). The employment interview: A critical summary. PERSONNEL PSYCHOLOGY, 2, 17–46.

Wiesner WH, Cronshaw SF. (1988). A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational Psychology, 61*, 275–290.

Williamson LG, Campion JE, Malos SB, Roehling MV, Campion MA. (1997). Employment interview on trial: Linking interview structure with litigation outcomes. *Journal of Applied Psychology, 82*, 900–912.

Wright PM, Lichtenfels PA, Pursell LD. (1989). The structured interview: Additional studies and a meta-analysis. *Journal of Occupational Psychology, 62*, 191–199.

Appendix A

*Meta-Analytic Information From Previously Conducted Studies
Included in the Meta-Correlation Matrix*

Variable	N	K	Observed mean r	ρ	SD_ρ	% of variance due to artifacts
Cognitive ability–Performance[a]	32,125	425	–	.45	.14[b]	–
Conscientiousness–Performance[c]	31,275	172	.18	.31	.09	39%
Level 1 interview–Performance[d]	7,308	15	.11	.20	.04	55%
Level 2 interview–Performance[d]	4,621	39	.20	.35	.11	39%
Level 3 interview–Performance[d]	4,358	27	.34	.56	.12	25%
Level 4 interview–Performance[d]	2,365	33	.34	.57	.17	28%

[a] Hunter and Hunter (1984).

[b] .14 = mean of SD_ρ(values reported in Hunter and Hunter (1984) for ability composite and Schmitt, Gooding, Noe, and Kirsch (1984) for their ability composite.

[c] Mount and Barrick (1995).

[d] Huffcutt and Arthur (1994); Levels 3 and 4 were combined in the present study.

Appendix B

Study	r	N	Nature of the job
Cognitive ability–Level 1 interviews			
Exxon, 1973	−0.050	170	Industrial process operator
	−0.050	170	
	−0.010	170	
	0.000	170	
	0.060	170	
Friedland, 1973	0.140	406	Police officers
Handyside & Duncan, 1954	0.070	110	Industrial supervisors
Cognitive ability–Level 2 interviews			
Dicken, 1969	−0.040	53	Peace Corps workers
	0.160	53	
Dipboye et al., 1990	0.070	446	Entry-level corrections officers
Exxon, 1974	0.210	69	Commercial drivers
	0.230	69	
	0.230	69	
	0.230	69	
	0.240	69	
	0.280	69	
	0.280	69	
	0.300	69	
	0.300	69	
	0.310	69	
	0.340	69	
	0.350	69	
	0.360	69	
	0.360	69	
	0.460	69	
	0.480	69	
Exxon, 1978	−0.130	75	Laboratory technicians
	−0.020	75	
	−0.010	75	
	0.000	75	
	0.020	75	
	0.020	75	
	0.020	75	
	0.040	75	
	0.040	75	
	0.050	75	
	0.050	75	
	0.050	75	
	0.060	75	
	0.060	75	
	0.070	75	
	0.080	75	
	0.080	75	
	0.100	75	
	0.130	75	
	0.140	75	
	0.160	75	
	0.170	75	
	0.180	75	
	0.190	75	
	0.200	75	

Appendix B (continued)

Study	r	N	Nature of the job
	0.200	75	
	0.200	75	
	0.210	75	
	0.210	75	
	0.220	75	
	0.250	75	
	0.280	75	
	0.290	75	
	0.300	75	
	0.300	75	
	0.300	75	
	0.340	75	
	0.340	75	
	0.350	75	
	0.350	75	
	0.350	75	
	0.420	75	
	0.440	75	
	0.440	75	
	0.500	75	
Friedland, 1976	−0.060	97	Administrative assistant
	0.090	98	
	0.140	92	
	0.180	98	
Friedland, 1980	−0.010	236	Police officers
	0.000	236	
	0.070	236	
	0.120	274	
	0.170	236	
	0.210	236	
Reeb, 1969	0.440	1000	Enlisted military
Roth & Campion, 1992	0.150	172	Industrial process technicians
Cognitive Ability–Level 3–4 interviews			
Campion et al., 1988	0.270	149	Pulp & paper mill employees
	0.370	149	
	0.500	149	
	0.540	149	
Campion et al., 1994	0.360	70	Pulp mill employees
	0.370	70	
	0.400	70	
	0.420	70	
	0.440	70	
	0.440	70	
	0.460	70	
	0.460	70	
	0.490	70	
	0.490	70	
	0.510	70	
	0.550	70	
	0.570	70	
	0.570	70	
	0.600	70	
	0.610	70	

Appendix B (continued)

Study	r	N	Nature of the job
Berkley, 1984	0.070	335	Correction officer trainees
	0.070	335	
	0.090	335	
	0.090	335	
	0.120	335	
	0.130	335	
	0.170	335	
	0.170	335	
Delery et al., 1992	0.110	47	Factory service technicians
	0.200	47	
	0.230	47	
	0.300	47	
Lopez, 1966	0.010	182	Female toll collectors
Motowidlo et al., 1992	−0.090	164	Management incumbents at a telecommunications company
	0.070	107	Applicants for entry-level management positions at a telecommunications company
	0.130	176	Marketing incumbents at a telecommunications company
	0.170	875	Applicants for entry-level management positions at a telecommunications company
Pulakos & Schmitt, 1995	0.090	464	Federal employees
Pulakos & Schmitt, 1996	0.120	464	Federal employees
Tubiana & Ben-Sitakhar, 1982	0.204	459	Male Israeli soldiers
	0.334	459	
Tziner & Dolan, 1982b	0.220	193	Female military officers
	0.300	193	
	0.380	193	
Cognitive ability–Conscientiousness			
Baehr & Froemel, 1977	−0.347	191	Police officers
	−0.327	191	
	0.396	194	
Borman et al., 1991	0.010	4362	Project A: First term Army soldiers
	0.090	4362	
Crant, 1995	−.14	131	Real estate agents
Day & Silverman, 1989	−0.090	43	Accountants
	0.150	43	
Gully et al., 2000	−.133	160	Undergraduates
Gully et al., 1998	−.032	345	Undergraduates
Hansen, 1989	0.090	362	Production and maintenance workers
	0.100	362	
Judge et al., 1999	0.290	194	Random samples of CA residents (Intergenerational studies)
Martocchio & Judge, 1997	−.24	97	Clerical and administrative employees
McHenry et al., 1990	−0.030	4039	Project A: Army soldiers
	−0.020	4039	
	0.010	4039	

Appendix B (continued)

Study	r	N	Nature of the job
	0.050	4039	
	0.050	4039	
	0.070	4039	
	0.080	4039	
	0.080	4039	
	0.080	4039	
	0.110	4039	
Neuman & Wright, 1999	.10	316	Human resource representatives
Phillips & Gully, 1997	−.03	330	Undergraduates
Pulakos et al., 1996	0.040	434	Federal employees
Ryan et al., 1998	0.120	1074	Firefighter applicants
Conscientiousness–Interview			
Motowidlo & Schmit, 1996	.17	140	Customer service
	.13	143	
Ones et al., 1996	.45	86	Manufacturing line
	.24	86	
	.38	86	
	.23	86	
	.21	86	
Schmit, 1996	.09	167	Customer service