The Use of Person–Organization Fit in Employment Decision Making: An Assessment of Its Criterion-Related Validity

Winfred Arthur Jr., Suzanne T. Bell, and Anton J. Villado
Texas A&M University

Dennis Dooverspike
University of Akron

Because measures of person–organization (P-O) fit are accountable to the same psychometric and legal standards used for other employment tests when they are used for personnel decision making, the authors assessed the criterion-related validity of P-O fit as a predictor of job performance and turnover. Meta-analyses resulted in estimated true criterion-related validities of .15 (k = 36, N = 5,377) for P-O fit as a predictor of job performance and .24 (k = 8, N = 2,476) as a predictor of turnover, compared with a stronger effect of .31 (k = 109, N = 108,328) for the more commonly studied relation between P-O fit and work attitudes. In contrast to the relations between P-O fit and work attitudes, the lower 95% credibility values for the job performance and turnover relations included zero. In addition, P-O fit’s relations with job performance and turnover were partially mediated by work attitudes. Potential concerns pertaining to the use of P-O fit in employment decision making are discussed in light of these results.

Keywords: person–organization fit, personnel selection, criterion-related validity, job performance, turnover

Using meta-analytic procedures, the primary objective of this article is to investigate the criterion-related validity of person–organization (P-O) fit as a predictor of job performance and turnover. The construct of P-O fit originates from interactional psychology, which assumes that behavior is caused by the continuous interaction between the person and the environment (Pervin, 1968; Terborg, 1981). Within this framework, an individual’s organizational behavior is posited to result from the interaction between the person and the organization. Schneider, Smith, and Paul (2001) identified two approaches to person–environment theory and research. One approach focuses on the environment as a moderator of the relation between the person and some individual-level criterion. The second approach views the fit between the individual and the environment as a predictor of specified outcomes. Our focus is on the second approach—in particular, the use of P-O fit as a predictor of job performance and turnover.

Although there are many types of fit, including person–team (group) fit, person–vocation fit, and person–job fit (see Kristof, 1996, for a discussion of the distinctions among these different types of fit), our focus is limited to P-O fit. We focus on P-O fit because of advocations for its use in personnel selection (e.g., Adams, Elacqua, & Collarelli, 1994; Bowen, Ledford, & Nathan, 1991; cf. Rynes, Brown, & Colbert, 2002) and the unique issues that surround the use of P-O fit in employment decision making. The issues we discuss may not be applicable to other forms of fit. For example, although person–job fit is also used for employment decision making, it is analogous to the standard selection model in which the focus is on matching an individual’s attributes (i.e., knowledge, skills, and abilities) to the demands of the job, and, thus, the concerns we outline are not applicable to person–job fit. In reference to the criteria, as we elaborate later, our focus is on job performance and turnover because, in contrast to work attitudes, these are criteria that are recognized by the Civil Rights Act (CRA; 1964, 1991) and the Equal Employment Opportunity Commission (EEOC; 1978) as appropriate employment (selection) test validation criteria. In addition, job performance is the most widely used criterion in personnel selection in particular and employment decision making in general.

In brief, we argue that although P-O fit may predict work attitudes (Verquer, Beehr, & Wagner, 2003), the nature of its relation to job performance and turnover may make its use in employment decision making questionable. In particular, we posit that P-O fit may not display direct relations with performance and turnover and that, subsequently, its indirect relations via work attitudes may lessen its desirability as an employment decision-making tool. In addition, the relation between P-O fit and job performance, turnover, and work attitudes may depend on its operationalization, the dimensions of fit, and the research validation design. Consequently, we conducted a meta-analysis to investigate the boundary conditions in which the use of P-O fit in employment decision making may be appropriate and those in which it may not.

P-O Fit and Work Attitudes

Kristof (1996) defined P-O fit as “the compatibility between people and organizations that occurs when: (a) at least one entity provides what the other needs, or (b) they share similar fundamental characteristics, or (c) both” (p. 45). Schneider’s (1987) attraction–seleciton–attrition (A-S-A) theory is an example of the strong theoretical foundation that serves as the basis for the hy-
pothesized relation between P-O fit and work attitudes. For example, the efficacy of P-O fit is predicated on the idea that the congruence between an organization and the individual’s values, interests, beliefs, and needs is related to the outcomes of interest. Thus, as a specific example, A-S-A theory posits that individuals are attracted to organizations that match their values and interests. Organizations, in turn, tend to select candidates who are most similar to the organization. After entry into the organization, individuals whose values are incongruent with the organization tend to leave, either voluntarily or involuntarily. With the attrition of these “different” individuals, those retained tend to be similar to one another, which thus increases the homogeneity of the organization (Schneider, 2001). Therefore, the outcome of three interrelated processes—attraction, selection, and attrition—determines the types of people in organizations. Other processes via which fit occurs include mentoring and socialization (Cable & Parsons, 2001; Chao, O’Leary-Kelly, Wolf, Klein, & Gardner, 1994; Chatman, 1991; Cooper-Thomas, van Vianen, & Anderson, 2004). In both instances, attempts are made to modify and alter individuals’ values, beliefs, interests, and behaviors to bring them in line with the organization’s culture, norms, and expectations.

Regardless of how fit is achieved, because the remaining individuals in the organization display a higher level of congruence with the organization, it is posited that they will also display more favorable work attitudes. Consequently, attitudinal outcomes, such as job satisfaction, organizational commitment, and turnover intentions, have been the most frequently used criteria in P-O fit studies. The viability of these hypothesized relations has been empirically demonstrated, as reflected in the results of Verquer et al.’s (2003) meta-analysis of the relations between P-O fit and work attitudes. Theoretically, the relation between fit and attitudes is predicated on the reasoning that when there is fit, the environment affords individuals the opportunity to fulfill their needs (Pervin, 1992; Rounds, Davis, & Loquist, 1987; Schneider, Kristof-Brown, Goldstein, & Smith, 1997). Need fulfillment results in favorable attitudes, such as job satisfaction and organizational commitment. In addition, social–psychological theories about similarity in attitudes (e.g., Byrne, 1971; Newcomb, 1961) can also be used to explain why fit is posited to be related to favorable attitudes. That is, people find it more desirable to interact with others who have similar psychological characteristics because the interaction verifies and reinforces their own beliefs, expressed behaviors, and affect (Swann, 1987; Swann, Stein-Seroussi, & Giesler, 1992). Thus, high levels of fit provide individuals opportunities to interact with similar others, and this, in turn, results in favorable attitudes, such as job satisfaction and organizational commitment.

P-O Fit and Personnel Selection

A reading of the P-O fit literature indicates that, in the course of its history, this literature initially focused more on organizational-level outcomes, such as organizational structure, culture, and effectiveness (Schneider et al., 2001), and later focused on individual-level outcomes, such as affect and behavior. More recently, there has been a migration of P-O fit from its historical origins in the posthire arena to prehire prescriptive use, specifically, in personnel selection. Indeed, the use of P-O fit for selection purposes appears to be on the increase (Rynes et al., 2002). For example, Adams et al. (1994) suggested that the focus on employment interviews should shift from the prediction of performance approach to a focus that stresses the interview’s usefulness for assessing P-O fit. In a similar vein, Judge, Higgins, and Cable (2000) reviewed recent developments in the employment interview research that focus on the use of the interview as a means of assessing P-O fit. Huffcutt, Conway, Roth, and Stone’s (2001) meta-analysis indicated that high-structure interviews that assess organization fit may be related to job performance (however, this was based on only four studies). Additional research focusing on the use of P-O fit in personnel selection includes C. L. Adkins, Russell, and Werbel (1994); Cable and Judge (1997); Chatman (1991); Hambleton, Kalili, and Taylor (2000); Karren and Graves (1994); and Powell (1998). This trend is most aptly summarized by Bowen et al.’s (1991) advocacy of “a new approach to selection in which employees are hired to fit the characteristics of an organization” (p. 35) and their description of “the basic steps of this new selection model” (p. 35).

Although there is nothing inherently objectionable in the use of P-O fit in employment decision making, in the absence of evidence demonstrating its job relatedness, we are somewhat wary of its migration into personnel selection for a variety of reasons. First, when used in this manner, P-O fit serves and functions as a test. Under these circumstances, it must be held to the same psychometric and legal standards expected of other personnel tests and predictors. In particular, P-O fit should be job related, as typically conceptualized in standard employment testing contexts; that is, better fit should result in higher outcomes that are acknowledged as appropriate test validation criteria by the CRA (1964, 1991) and the EEOC (1978). Yet, as early as 1997, Schneider et al. observed that the overwhelming majority of research on fit (both person–environment and P-O) has used various indices of individual affect as outcome criteria: adjustment, satisfaction, commitment, and turnover. Little of the research on P-O fit has concerned productivity or other indicators of work performance. (p. 396)

Despite voicing this nearly a decade ago, Schneider et al.’s observation still accurately characterizes the current literature. Thus, in the absence of a substantive body of evidence documenting the empirical relation between P-O fit and appropriate test validation criteria, its use as a selection device may be presumptuous, because it appears to be a carryover from its well-documented relation with work attitudes. In summary, if P-O fit is to be used to make employment-related decisions, then the extent to which it is related to appropriate test validation criteria (e.g., performance, turnover) is of interest. At the present time, we do not think this has been sufficiently established.

Second, in contrast to the work attitudes literature, in which the theoretical and conceptual bases for the fit–attitudes relation are well developed and direct, the conceptual and theoretical bases for hypothesized relations between P-O fit and job performance tend to be indirect and typically involve a relation mediated by work attitudes. For example, the implications of the need fulfillment (Pervin, 1992; Schneider et al., 1997), work adjustment (Loquist & Davis, 1969; Rounds et al., 1987), and similarity in attitudes (e.g., Byrne, 1971; Newcomb, 1961) theories presented earlier to explain why fit is posited to be related to favorable work attitudes can be extended to job performance. A conceptual argument for the P-O fit and job performance relation could be that P-O fit may have an indirect effect on job performance through job satisfaction. For instance, recent meta-analytic research by Judge, Thoresen, Bono, and Patton (2001) suggests a moderate relation between job satisfaction and job performance (p = .30). This moderate job satisfaction–job performance relation may suggest that because P-O fit is related to worker attitudes such as job satisfaction (Verquer et al., 2003) and job satisfaction is related to job perfor-
performance (Judge et al., 2001), P-O fit may have an indirect effect on job performance through job satisfaction. Similar arguments could be made for other attitudes, such as organizational commitment (Meyer, Stanley, Herscovitch, & Topolnytsky, 2002). It is also conceivable that the relation between P-O fit and job performance could be via a reduction in job stressors, such as role ambiguity and a lack of role clarity, such that fit attenuates these stressors (Parkington & Schneider, 1979), which, in turn, facilitates increased job performance (Tubre & Collins, 2000).

In summary, the mediational role of attitudes in the fit–performance relation suggests that the relation between fit and performance will be weak at best after the effects of work attitudes are taken into account. Thus, if there is no independent contribution of P-O fit to the prediction of performance, then the complex relation based on intermediary variables may limit the concept’s use as an acceptable employment decision-making tool. We note that it is not our intention to imply that if work attitudes mediate the relation between P-O fit and performance or turnover, then work attitudes should be used in employment decision making. To the contrary, although one could posit that some organizations may consider employee satisfaction to be an important enough organizational value that they want to hire on the basis of who is likely to be satisfied on the job, we are yet to encounter such a selection scenario in our research and applied experience. In addition, should such a selection system display adverse impact, its defense should such a selection system display adverse impact, its defense is likely to be tenuous at best, because, as we discuss later, the EEOC (1978) guidelines do not list work attitudes as appropriate test validation criteria. Finally, it is obvious that the use of work attitudes as predictors in selection or hiring decision making is limited by the fact that individuals first have to be hired and on the job to have these work attitudes.

Length of service (i.e., turnover or tenure) is recognized as a legitimate test validation criterion by the EEOC (1978, Section 14, B.3). Thus, turnover could be of particular interest in the selection of employees (Barrick & Zimmerman, 2005). For researchers, this interest is motivated by a desire to understand critical motivated and volitional behavior, and for managers, it is motivated by a desire to reduce the myriad personnel and human resource management costs associated with voluntary turnover. From a theoretical perspective, the attrition phase of A-S-A theory provides the rationale for a posited relation between P-O fit and turnover. In addition, one would also expect work attitudes to play a mediational role in the P-O fit–turnover relation, such that individuals who are less satisfied or committed as a result of poor fit are more likely to leave the organization (Griffeth, Hom, & Gaertner, 2000; Tett & Meyer, 1993). However, as with job performance, if, after the effects of work attitudes are taken into account, there is a limited or no independent contribution of P-O fit to the prediction of turnover, then P-O fit may have limited use as a selection tool.

Third, a close reading of the P-O fit–performance literature indicates that although there is some disagreement as to whether increased P-O fit is always desirable (e.g., because of detrimental effects of groupthink; Schneider et al., 2001), several discussants of the use of fit in selection have focused on organizational effectiveness or other organizational-level outcomes, not individual performance, as the criterion of interest (e.g., Argyris, 1957; Bowen et al., 1991; Hambrick & Brandon, 1988; Hambrick & Mason, 1984; Schneider, 1987; Schneider et al., 1997). Indeed, this focus on organizational-level outcomes is consistent with the historical origins of P-O fit (Schneider et al., 2001). Concomitantly, although they did not present any empirical data to support their proposition, Bowen et al.’s (1991) emphasis on selection for the organization and on the importance of looking beyond jobs was based on the premise that a good fit of individuals to the organization is beneficial for organizational effectiveness. Similar suggestions to extend P-O fit–related criterion development and validation designs to include organizational levels of analyses, such as organizational effectiveness, have been made by Schneider et al. (1997). More recently, Schneider, Smith, and Sipe (2000) proposed a multilevel model of personnel selection that expands the traditional approach to validating selection systems to include the impact that these systems have on the broader organizational system. Although these extended and broader validation models that encompass organizational-level criteria may be long overdue in personnel psychology, the current legal and professional standard and practice are that employment tests must be validated against individual-level criteria—a specific focus on validation against what is performed on the job (CRA, 1964, 1991; EEOC, 1978). In addition, just because relations are observed at the organizational level is no reason to assume that they are homologous at the individual level.

In summary, in the absence of empirical evidence demonstrating its relation to individual-level performance, the position that the use of P-O fit in personnel selection results in increased organizational effectiveness appears to be a legally weak justification for the use of P-O fit in employee selection.

In summary, given the interest in P-O fit as a selection tool and the potential limitations of its use that we have outlined, we sought to use meta-analytic procedures to investigate the criterion-related validity of P-O fit as a predictor of job performance and turnover. Consequently, the objectives of the present study were to

1. comparatively evaluate the extent to which P-O fit is related to work attitudes, job performance, and turnover;
2. investigate the extent to which P-O fit’s relations with job performance and turnover are mediated by work attitudes; and
3. examine the influence of potential moderator variables to clarify the boundary conditions germane to the observed overall pattern of relations between P-O fit and the specified criteria.

Moderators of the Relation Between P-O Fit and the Outcome Variables

Previous research suggests that the relation between P-O fit and specified outcome variables may differ in the context of different boundary conditions. For instance, Verquer et al.’s (2003) meta-analysis indicated that the dimension and operationalization of fit moderate the relation between P-O fit and work attitudes. Consequently, because our primary focus is on behavioral outcomes, we also sought to investigate the role of these moderators within the context of P-O fit’s relation with job performance and turnover. We also investigated the role of two additional moderators—the type of validation design and the method of calculating fit. In the present study, we refer to the three commonly used operationalizations of fit as indirect–actual, indirect–perceived, and direct–perceived fit. Researchers (e.g., Verquer et al., 2003) often have used the respective labels of objective, perceived, and subjective fit. With indirect–actual fit, a target individual’s ratings of his or her own characteristics are compared with ratings or descriptions of the organization on the same
dimensions obtained from a different source. With indirect–perceived fit, ratings of the target individual and the organization on the same dimensions or characteristics, obtained from the same source, are compared. Finally, direct–perceived fit measures simply ask the rater (either the target individual or another rater) to rate the extent to which fit between the target individual and the organization exists. As has been noted, the preceding are quite different ways of assessing fit and thus are potentially viable moderators of observed effects. For instance, Verquer et al.’s results suggested that measures of direct–perceived fit displayed the strongest relations with work attitudes. However, this finding may be reflective of common source bias, as both the fit and the attitudinal data were obtained from the same source—a situation that is less likely to be the case when the criterion is a behavioral variable (e.g., job performance), which is typically obtained from another source.

Dimensions of fit pertain to what is being measured or the content of fit. Common dimensions present in the extant literature include values, goals, and personality–climatic congruence. We sought to investigate dimensions of fit as a moderator because it is increasingly being acknowledged in the personnel selection literature that it is important to recognize the method–construct distinction (Arthur, Day, McNelly, & Edens, 2003; Arthur & Doverspike, 2005; Huffcutt, Conway, Roth, & Stone, 2001). In addition, this recent stream of research has demonstrated that the specific constructs or dimensions being measured moderate the criterion-related validity of specified predictors (i.e., methods). Consequently, it is plausible that the criterion-related validity of P-O fit could vary as a function of the specific dimensions of fit in conjunction with the appropriateness of the match to specified criterion variables. So, for example, whereas value congruence may not predict how well someone will perform the job, it may be predictive of whether an individual will leave the organization (i.e., turnover).

Previous research has suggested that, under certain circumstances, the validation design can moderate the criterion-related validity of predictors (see G. V. Barrett, Phillips, & Alexander, 1981; Schmitt, Gooding, Noc, & Kirsch, 1984; cf. Ones, Viswesvaran, & Schmidt, 1993). Indeed, the predictive versus concurrent design distinction (which typically covaries with the use of applicant vs. incumbent samples, respectively) might be even more important in the present case because of the A-S-A process. For instance, one would expect weaker relations for concurrent designs because of lowered variability in the personal characteristics of the study participants. Consequently, we sought to investigate the efficacy of the validation design as a potential moderator. In coding the articles on this variable, we found that all the job performance studies used incumbents as the study participants. Therefore, although we were able to code for and assess the effects of validation design, contrary to what is typically found in the personnel selection literature, for both types of designs, the research participants were all incumbents. The implications of this, which we consider to be an idiosyncrasy of the P-O fit–job performance literature, are addressed in the Discussion.

The fourth potential moderator was the method used to calculate fit. Whereas direct–perceived measures of fit do not require a calculation of fit, both indirect–actual and indirect–perceived measures of fit require some calculation of the discrepancy between the person and the organization on the characteristic of interest. Different techniques have been used, including difference scores (e.g., Vigoda, 2000), correlations (e.g., O’Reilly, Chatman, & Caldwell, 1991), and polynomial regressions (e.g., Van Vianen, 2000), and each technique has its strengths and weaknesses. One major weakness associated with the use of difference scores to calculate P-O fit is the attenuation of reliability (see Edwards, 1994, for additional critiques of the use of difference scores). Thus, P-O fit–outcomes relations may be stronger when P-O fit is calculated by polynomial regression or correlations compared with when fit is calculated via difference scores. However, this particular critique may not be an issue for the present meta-analysis. In particular, we address the attenuation of reliability because the estimated true criterion-related validity reflects a correction for unreliability. The correction for predictor unreliability for studies using a difference-score technique to calculate fit was based on the product of the two predictor measures to more accurately represent the reliability of the scores.

The use of correlations has also been critiqued on the basis that these indexes reflect similarity in profile shape but not additional potentially important information, such as the distance between the profiles; consequently, they may result in an attenuation of the P-O fit–outcomes relations. However, some researchers (e.g., C. L. Adkins et al., 1994; Verquer et al., 2003) have suggested that the rank order of values is more important than their absolute levels in determining reactions to the organization; thus, P-O fit calculated via correlations may still be a better predictor of specified outcomes compared with that calculated by difference scores. Therefore, given its potential to influence the observed relations between P-O fit and specified outcomes, we tested the method of calculating fit as a potential moderator. In summary, because of the absence of a clear demonstration of P-O fit’s job relatedness (i.e., as a predictor of job performance and turnover), in the present study we sought to use meta-analytic procedures to investigate the criterion-related validity of P-O fit and examine the influence of specified potential moderators.

Method

Literature Search

We conducted an extensive literature search to identify relevant P-O fit studies. The literature search encompassed studies published in journals, books or book chapters, conference papers and presentations, and dissertations or theses that pertained to P-O fit and its relations to the outcome variables of interest (e.g., job performance). The search process started with a search of PsycINFO, with search terms such as person–organization fit, person–climate fit, person–culture fit, organizational fit, person–environment fit, person–organization congruence, person–organization fit + job performance, organization fit + job performance, and organization fit + selection. This electronic search was complemented with a manual search of the reference lists from Kristof’s (1996) review of the P-O fit literature, Verquer et al.’s (2003) meta-analysis, and conference programs from the Society of Industrial and Organizational Psychology and the Academy of Management from 2000 to 2003. We also made an attempt to obtain unpublished research. We sent letters requesting unpublished research to a comprehensive list of consulting and research organizations. The organizations came from a list that had been generated to request funds and support for the 24th annual Industrial/Organizational and Organizational Behavior Graduate Student Conference in 2003. In addition, we supplemented this list by contacting consulting firms that had advertisements in the 2002–2003 issues of The Industrial–Organizational Psychologist and other firms known to us. This resulted in a total of 43 unique organizational requests (some organizations might have been sent multiple letters through more than one contact). The letter explained the purpose of the current research as well as the characteristics required of the requested data or research. Three months later, we sent a follow-up letter to the organizations that had not responded to the first letter. This letter repeated the request for research and emphasized that it was not too late to submit available data or
research. There were few responses to both letters, and none of the organizations had research or data that met the inclusion criteria.

We reviewed and evaluated the abstracts of the articles and papers obtained as a result of the initial search for the appropriateness of content (i.e., empirical studies that evaluated the relation between P-O fit and some individual-level outcome measure, such as job performance). This review, along with a decision to retain only English-language articles, resulted in a list of 110 articles and papers. Next, we reviewed the reference lists of these 110 articles and papers to identify additional sources. The cumulative result of our search was a total list of 197 sources. We then reviewed each of these for possible inclusion in the meta-analysis.

Inclusion Criteria

We used a number of decision rules to determine which studies would be included in the meta-analysis. First, a study had to be an empirical investigation of P-O fit in which the organization (i.e., not the supervisor, work group, or job) was the comparison or referent and the criterion was an individual-level outcome variable. Because of the study’s objectives, we limited the outcome variables to job performance, turnover, and three attitudinal criteria—organizational commitment, job satisfaction, and turnover intentions. Although attitudes were not the focal criteria of the present study (and were recently the topic of investigation in Verquer et al.’s, 2003, meta-analysis), we chose to code for attitudinal criteria to avoid potential shortcomings associated with making comparative statements across meta-analyses. We note that, in general, our results for the attitudinal data were similar to those reported by Verquer et al. (2003).1 Second, to be included in the meta-analysis, studies had to report sample sizes and a correlation or enough information for us to compute a correlation. For the studies that reported other univariate test statistics, we used the appropriate conversion formulas.

Data Sets

As a result of the inclusion criteria, we obtained an initial set of 288 correlations from 46 sources; however, some of these correlations were non-independent or computed from data collected from the same sample of participants. Correlations that were based on the same sample were retained as independent correlations only if they assessed a different level of a specified moderator; otherwise, the correlations were represented in the data set by a linear composite (Hunter & Schmidt, 1990, pp. 454–457; Viswesvaran & Ones, 1995, p. 873). Because we conducted five separate moderator analyses (i.e., criteria, operationalization of fit, dimension of fit, validation design, calculation of fit), we constructed an independent data set for each analysis. We computed Huffcutt and Arthur’s (1995; cf. Beal, Corey, & Dunlap, 2002) sample-adjusted meta-analytic deviance statistic to detect outliers in each of the five data sets. On the basis of these analyses, 7 unique correlations were identified as outliers across the five data sets. A detailed review of the correlations indicated 1 suspect correlation, which was probably the result of a typographical error in the study because it was a large negative correlation that ran counter to the theoretical prediction and was neither noted nor discussed by the author. Therefore, we deleted this correlation from further analyses. Consequently, the final data sets consisted of 153 independent correlations from 46 sources for the analyses investigating criterion type (i.e., job performance, turnover, attitudes), 107 independent correlations from 46 sources for the analyses investigating operationalizations of P-O fit (i.e., indirect–actual, indirect–perceived, and direct–perceived), 109 independent correlations from 46 sources for the analyses investigating dimensions of fit (e.g., value congruence), 153 independent correlations from 44 sources for the analyses investigating validation design (i.e., predictive, concurrent), and 86 independent correlations from 28 sources for the analyses investigating calculation of fit (i.e., difference scores, correlations).

Description of Variables

We coded for the following variables in the meta-analysis. We coded criterion type as being either turnover, job performance, or attitudinal. For job performance, we differentiated between task and contextual performance. We coded job performance as task performance if the behaviors appeared to be role prescribed (i.e., could be identified as part of the job; Campbell, McCloy, Oppler, & Sager, 1993) and as contextual performance if they were behaviors that supported the broader organizational, psychological, and social environment in which the technical core operated (Borman & Motowidlo, 1993, 1997; Motowidlo & Van Scotter, 1994; Van Scotter & Motowidlo, 1996). When we could not code performance as task or contextual performance, either because there was not enough information to make the distinction or because the job performance measure included both, we coded it as “job performance—unspecified.” We further coded attitudinal criteria as job satisfaction, organizational commitment, and turnover intentions.

We coded the operationalization of fit as one of three categories—indirect–actual, indirect–perceived, or direct–perceived fit. We coded for several dimensions of fit, including value congruence, goal congruence, and personality–climate congruence. However, because of a general focus of the literature on values and a small number of correlations for many of the facets, for the analyses and presentation of results, we grouped the dimensions of fit into value congruence (i.e., the fit between the values of the individual and the values of the organization), value + other congruence (for instances in which value congruence was measured along with other types of congruence), and other congruence (for measures of any type of congruence other than value congruence).

We also coded the validation design of each study (i.e., predictive and concurrent) by recording the length of time that had elapsed between the assessment of the predictor and criterion measures. We coded predictive studies as those studies in which 1 or more days had elapsed between the assessment of the predictor and criterion (mean number of days between predictor and criterion measures = 276, SD = 138, minimum = 6, maximum = 720). Concurrent studies were those studies in which the predictor and criterion were assessed on the same day.

We coded for the techniques used for the calculation of fit. That is, for measures of fit that required a calculation of fit (i.e., indirect–actual and indirect–perceived), we coded calculation of fit as either a difference score, a correlation, or a polynomial regression. We note that we were unable to locate any studies that assessed P-O fit–job performance or P-O fit–turnover relations using polynomial regression techniques. In addition, studies that used polynomial regressions for the calculation of fit to assess the P-O fit–attitudes relation did not include enough information for the conversion of the effect to the common effect size metric.

Coding Accuracy and Interrater Agreement

The coding process was as follows. First, Suzanne T. Bell and Anton J. Villado were provided with a coding sheet and a reference guide. Each coder used the reference guide to independently code one article. Next, they attended a follow-up meeting during which they discussed problems encountered using the coding sheet and reference guide and made changes to the coding sheet as deemed necessary. They then coded one additional article. After they coded this second article, we assessed the degree of convergence. We resolved discrepancies and disagreements related to the article via discussion and consensus. After the second meeting, Suzanne T. Bell and Anton J. Villado subsequently independently coded all articles used in the meta-analysis. An empirical comparison of their independent data sets indicated less than 2% disagreement; these differences were resolved via discussion.

Artifact Distribution and Meta-Analytic Procedures

We used Arthur, Bennett, and Huffcutt’s (2001) SAS PROC MEANS program to conduct a psychometric meta-analysis (Hunter & Schmidt, 1990).

1 A reference list that presents a comparison of our primary articles that included attitudinal criteria and Verquer et al.’s (2003) primary studies is available from Winfred Arthur Jr.
The psychometric meta-analysis procedure allows for the correction of statistical artifacts, such as sampling and measurement error. We used an artifact distribution approach (Hunter & Schmidt, 1990, pp. 159–160) because predictor and criterion reliability data were not available for every study. We computed a separate artifact distribution for each analysis such that the artifact distribution only included artifact information for the specific analysis of interest. For example, when estimating the relation between P-O fit and task performance, we used all P-O fit reliability estimates to create the predictor reliability distribution \( \rho_{xx} \), whereas we used only those reliability estimates pertaining to task performance to create the criterion reliability distribution \( \rho_{yy} \). Alternatively, when estimating the relation between value congruence measures of P-O fit and overall job performance, we used only those reliability estimates pertaining to value congruence measures of P-O fit to create the predictor reliability distribution \( \rho_{xx} \), whereas we used all reliability estimates based on measures of job performance to create the criterion reliability distribution \( \rho_{yy} \). Descriptive information for the artifact distributions is available from Winfred Arthur Jr.

As we have previously mentioned, we analyzed the potential impact of several moderators. When one is testing for moderators, if correcting for statistical artifacts does not account for all or nearly all of the observed variation in correlations, the standard deviation of the estimated true validities is large, or both, then there is reason to believe that the validity is dependent on the situation or moderators. Conversely, if all or a major portion of the observed variance in validities is attributed to statistical artifacts, one can conclude that the validities are constant or nearly so (Hunter & Schmidt, 1990). We also computed lower 95% credibility values to assess whether the validities were positive across situations (i.e., whether validity generalized) (Hunter & Schmidt, 1990). However, the lower 95% credibility value can be greater than zero (validity generalizes) and still have sizable variance in the validities after correction for statistical artifacts. Under these conditions, one can conclude that the validities are positive, although the actual magnitude may vary as a function of specified moderators (i.e., situational specificity). Finally, we also computed 95% confidence intervals as a measure of the accuracy of the effect size (Whitener, 1990). The confidence interval reflects the extent to which sampling error remains in the effect size estimate (Whitener, 1990) and suggests the range of plausible values for the true effect size (Cumming & Finch, 2005). The confidence interval is generated via the standard error for the mean correlation and is applied to the mean sample-weighted correlation before the correction for attenuating artifacts (Whitener, 1990).

Results

Extent to Which P-O Fit Has Been Validated as a Predictor of Job Performance and Turnover

We first assessed the extent to which P-O fit has been investigated as a predictor of job performance or turnover in the extant literature relative to attitudinal outcomes. There were a total of 153 independent correlations that represented the relation between P-O fit and job performance, turnover, and attitudinal criteria. Figure 1

---

2 All relations were corrected for predictor and criterion unreliability except for the relations between P-O fit and turnover, which were corrected for predictor unreliability only.
Table 1
Meta-Analysis Results for the Relation Between Person–Organization (P-O) Fit and Job Performance, Turnover, and Attitudinal Criteria

<table>
<thead>
<tr>
<th>Variable</th>
<th>$k$</th>
<th>$n$</th>
<th>Sample-weighted mean $r$</th>
<th>Sample-weighted SD</th>
<th>% var. sampling error</th>
<th>95% CI</th>
<th>% var. acc. for</th>
<th>Lower 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Criteria</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job performance—overall</td>
<td>36</td>
<td>5,377</td>
<td>.12</td>
<td>.15</td>
<td>31.09</td>
<td>.08 .17</td>
<td>.15 .16</td>
<td>32.32</td>
</tr>
<tr>
<td>Task performance</td>
<td>12</td>
<td>2,195</td>
<td>.08</td>
<td>.12</td>
<td>39.83</td>
<td>.01 .14</td>
<td>.10 .12</td>
<td>40.70</td>
</tr>
<tr>
<td>Contextual performance</td>
<td>9</td>
<td>1,426</td>
<td>.17</td>
<td>.16</td>
<td>22.60</td>
<td>.07 .28</td>
<td>.22 .18</td>
<td>24.53</td>
</tr>
<tr>
<td>Unspecified</td>
<td>15</td>
<td>1,756</td>
<td>.14</td>
<td>.14</td>
<td>39.60</td>
<td>.07 .21</td>
<td>.17 .13</td>
<td>40.73</td>
</tr>
<tr>
<td>Turnover—overall$^a$</td>
<td>8</td>
<td>2,476</td>
<td>.21</td>
<td>.17</td>
<td>9.74</td>
<td>.09 .34</td>
<td>.24 .19</td>
<td>11.24</td>
</tr>
<tr>
<td>Attitudinal criteria—Overall$^{ab}$</td>
<td>109</td>
<td>108,328</td>
<td>.12</td>
<td>.12</td>
<td>5.78</td>
<td>.22 .27</td>
<td>.31 .14</td>
<td>10.79</td>
</tr>
<tr>
<td>Turnover intention$^{ab}$</td>
<td>33</td>
<td>34,938</td>
<td>.19</td>
<td>.12</td>
<td>6.14</td>
<td>.15 .24</td>
<td>.25 .14</td>
<td>9.83</td>
</tr>
<tr>
<td>Job satisfaction$^{b}$</td>
<td>33</td>
<td>38,999</td>
<td>.29</td>
<td>.13</td>
<td>5.66</td>
<td>.25 .33</td>
<td>.36 .16</td>
<td>11.22</td>
</tr>
<tr>
<td>Organizational commitment$^{b}$</td>
<td>30</td>
<td>35,291</td>
<td>.25</td>
<td>.10</td>
<td>8.19</td>
<td>.22 .29</td>
<td>.31 .11</td>
<td>16.57</td>
</tr>
<tr>
<td><strong>Operationalization of fit</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indirect–actual</td>
<td>13</td>
<td>2,260</td>
<td>.09</td>
<td>.12</td>
<td>39.03</td>
<td>.03 .16</td>
<td>.12 .12</td>
<td>40.73</td>
</tr>
<tr>
<td>Indirect–perceived</td>
<td>10</td>
<td>1,055</td>
<td>.13</td>
<td>.11</td>
<td>74.93</td>
<td>.06 .19</td>
<td>.16 .16</td>
<td>77.05</td>
</tr>
<tr>
<td>Direct–perceived</td>
<td>10</td>
<td>1,688</td>
<td>.17</td>
<td>.15</td>
<td>24.55</td>
<td>.08 .27</td>
<td>.21 .21</td>
<td>26.16</td>
</tr>
<tr>
<td>Turnover</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indirect–actual</td>
<td>6</td>
<td>1,856</td>
<td>.20</td>
<td>.19</td>
<td>7.93</td>
<td>.05 .36</td>
<td>.23 .21</td>
<td>9.88</td>
</tr>
<tr>
<td>Indirect–perceived</td>
<td>1</td>
<td>361</td>
<td>.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct–perceived</td>
<td>3</td>
<td>596</td>
<td>.30</td>
<td>.21</td>
<td>9.27</td>
<td>.06 .54</td>
<td>.32 .22</td>
<td>9.47</td>
</tr>
<tr>
<td>Attitudinal criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indirect–actual$^b$</td>
<td>17</td>
<td>30,324</td>
<td>.21</td>
<td>.05</td>
<td>17.14</td>
<td>.18 .23</td>
<td>.26 .06</td>
<td>45.24</td>
</tr>
<tr>
<td>Indirect–perceived</td>
<td>26</td>
<td>9,855</td>
<td>.36</td>
<td>.14</td>
<td>10.71</td>
<td>.31 .42</td>
<td>.45 .16</td>
<td>16.03</td>
</tr>
<tr>
<td>Direct–perceived</td>
<td>21</td>
<td>4,872</td>
<td>.52</td>
<td>.17</td>
<td>8.34</td>
<td>.45 .59</td>
<td>.62 .19</td>
<td>11.19</td>
</tr>
<tr>
<td><strong>Dimensions of fit</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value congruence</td>
<td>15</td>
<td>2,098</td>
<td>.11</td>
<td>.14</td>
<td>34.94</td>
<td>.04 .18</td>
<td>.14 .14</td>
<td>35.58</td>
</tr>
<tr>
<td>Value + other congruence</td>
<td>5</td>
<td>1,637</td>
<td>.15</td>
<td>.15</td>
<td>12.20</td>
<td>.01 .29</td>
<td>.18 .17</td>
<td>13.20</td>
</tr>
<tr>
<td>Other congruence</td>
<td>14</td>
<td>1,488</td>
<td>.15</td>
<td>.16</td>
<td>35.48</td>
<td>.06 .23</td>
<td>.18 .16</td>
<td>37.19</td>
</tr>
<tr>
<td>Turnover</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value congruence</td>
<td>5</td>
<td>942</td>
<td>.34</td>
<td>.19</td>
<td>12.03</td>
<td>.18 .51</td>
<td>.38 .19</td>
<td>13.07</td>
</tr>
<tr>
<td>Value + other congruence</td>
<td>2</td>
<td>1,107</td>
<td>.15</td>
<td>.20</td>
<td>4.11</td>
<td>−.13 .44</td>
<td>.17 .22</td>
<td>4.31</td>
</tr>
<tr>
<td>Other congruence</td>
<td>3</td>
<td>764</td>
<td>.23</td>
<td>.09</td>
<td>47.84</td>
<td>.13 .32</td>
<td>.26 .06</td>
<td>57.11</td>
</tr>
<tr>
<td>Attitudinal criteria$^b$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value congruence</td>
<td>33</td>
<td>11,716</td>
<td>.39</td>
<td>.16</td>
<td>8.22</td>
<td>.34 .45</td>
<td>.48 .18</td>
<td>11.36</td>
</tr>
<tr>
<td>Value + other congruence</td>
<td>12</td>
<td>3,297</td>
<td>.48</td>
<td>.16</td>
<td>8.65</td>
<td>.39 .57</td>
<td>.57 .18</td>
<td>13.62</td>
</tr>
<tr>
<td>Other congruence$^b$</td>
<td>20</td>
<td>30,513</td>
<td>.22</td>
<td>.08</td>
<td>9.44</td>
<td>.18 .25</td>
<td>.27 .09</td>
<td>20.69</td>
</tr>
<tr>
<td><strong>Validation design</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predictive</td>
<td>9</td>
<td>2,174</td>
<td>.10</td>
<td>.08</td>
<td>63.60</td>
<td>.05 .15</td>
<td>.12 .06</td>
<td>64.32</td>
</tr>
<tr>
<td>Concurrent</td>
<td>24</td>
<td>2,694</td>
<td>.12</td>
<td>.18</td>
<td>27.47</td>
<td>.04 .19</td>
<td>.14 .19</td>
<td>28.03</td>
</tr>
<tr>
<td>Turnover$^c$</td>
<td>8</td>
<td>2,476</td>
<td>.21</td>
<td>.17</td>
<td>9.74</td>
<td>.09 .34</td>
<td>.24 .19</td>
<td>11.24</td>
</tr>
<tr>
<td>Attitudinal criteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predictive$^b$</td>
<td>28</td>
<td>4,646</td>
<td>.33</td>
<td>.17</td>
<td>17.35</td>
<td>.27 .40</td>
<td>.41 .18</td>
<td>23.20</td>
</tr>
<tr>
<td>Concurrent</td>
<td>84</td>
<td>103,820</td>
<td>.24</td>
<td>.12</td>
<td>5.20</td>
<td>.21 .26</td>
<td>.30 .14</td>
<td>11.23</td>
</tr>
<tr>
<td><strong>Calculation of fit</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference score</td>
<td>8</td>
<td>2,128</td>
<td>.04</td>
<td>.07</td>
<td>73.98</td>
<td>−0.01 .09</td>
<td>.05 .05</td>
<td>74.99</td>
</tr>
<tr>
<td>Correlation</td>
<td>12</td>
<td>1,297</td>
<td>.11</td>
<td>.16</td>
<td>34.07</td>
<td>.02 .20</td>
<td>.13 .16</td>
<td>34.44</td>
</tr>
<tr>
<td>Turnover$^e$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference score</td>
<td>1</td>
<td>914</td>
<td>.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>4</td>
<td>798</td>
<td>.38</td>
<td>.18</td>
<td>11.55</td>
<td>.20 .55</td>
<td>.45 .20</td>
<td>13.52</td>
</tr>
</tbody>
</table>

792 ARTHUR, BELL, VILLADO, AND DOVERSPIKE
presents the distribution (histogram) of independent correlations included in the meta-analysis, broken down by job performance, turnover, and attitudinal criteria. The correlations in Figure 1 are grouped in .10 intervals. Of 153 correlations, 24% (k = 36) represented the relation between fit and job performance. These 36 correlations were extracted from 10 journal articles, 2 conference presentations, and 5 dissertations. There were 8 turnover correlations (5%), and they were obtained from 5 journal articles, 2 conference presentations, and 1 dissertation. In contrast, 71% (k = 109) of the total number of correlations in the meta-analysis represented the relation between fit and an attitudinal outcome variable. These 109 correlations were extracted from 19 journal articles, 4 conference presentations, and 19 dissertations. A proportionally larger number of the attitudinal outcome correlations represented the relation between fit and job satisfaction (k = 46; 42%). There were approximately equal numbers of organizational commitment and turnover intention correlations (k = 30, 28%; and k = 33, 30%, respectively).

In summary, although the P-O fit–job performance relation has received some research attention, consistent with what would be expected given the historical origins of P-O fit, the attention has been considerably less than that paid to P-O fit’s relation with attitudinal outcomes. Furthermore, the P-O fit–attitudes relation was represented by almost twice as many peer-reviewed publications as the fit–job performance relation. In addition, very few of the P-O fit–performance studies used job performance as the focal criterion, and often the P-O fit–job performance relation was not even hypothesized and was presented in the source in a supplementary, tangential manner. The P-O fit–turnover relation was rarely studied.

Meta-Analytic Estimate of the Relation Between P-O Fit and Job Performance and Turnover

The main objective of the present study was to use meta-analytic procedures to estimate the true criterion-related validity of P-O fit as a predictor of job performance and turnover and to compare these relations with the P-O fit–attitudes relation. The results of the meta-analysis are presented in Table 1. The estimated true criterion-related validities (ρ) for P-O fit were .15, .24, and .31 for job performance, turnover, and work attitudes, respectively. Thus, the job performance effect was small and half the magnitude of the medium effect for work attitudes (Cohen, 1992). The criterion-related validity for turnover was more similar to the medium effect observed for work attitudes than to the effect observed for job performance. A further breakdown of the job performance analyses by criterion type revealed a stronger effect for contextual (ρ = .22) compared with task performance (ρ = .10). However, for both facets of performance as well as the overall performance, the lower 95% credibility value was less than zero. This indicates that estimates of the true P-O fit–job performance criterion-related validities include zero, so the validity of P-O fit as a predictor of job performance does not generalize. In contrast, the breakdown of the attitudinal variables indicated the strongest effects for job satisfaction (ρ = .36) and organizational commitment (ρ = .31), both of which had lower 95% credibility values that were greater than zero.

Operationalizations of fit. As indicated in Table 1, only a small percentage of variance could be attributed to statistical artifacts, and the standard deviations of the population estimates were fairly large, indicating that there may be subpopulations present for the P-O fit–outcome relations. Consequently, we next investigated the effect of the previously specified moderators of the P-O fit relations. As the results in Table 1 indicate, the strongest fit–job performance relation was obtained when P-O fit was operationalized as direct–perceived fit (ρ = .21), compared with indirect–perceived (ρ = .16) and indirect–actual fit (ρ = .12). However, of these, only the indirect–perceived fit operationalization had a lower 95% credibility value above zero. A similar rank order of the operationalizations was obtained for the P-O fit–attitudinal criteria relations. The strongest effect was for direct–perceived fit (ρ = .62), followed by indirect–perceived (ρ = .45) and indirect–actual fit (ρ = .26). However, unlike the job performance effects, all the lower 95% credibility values were greater than zero. There were too few turnover correlations to permit a meaningful interpretation of the moderator analyses associated with this variable; nevertheless, they are presented in Table 1 for the sake of completeness. We note that, across all criteria, for the direct–perceived operationalizations analyses, only a small amount of variance could be attributed to artifacts, which suggests that there may be subpopulations (i.e., additional moderators) present.

Dimensions of fit. The moderator analyses of dimensions indicated that the relation between value congruence and work attitudes was stronger (ρ = .48) than that between value congruence and job performance (ρ = .14). Although it was based on a
limited number of correlations, the value congruence–turnover relation was fairly strong ($\rho = .38$), as A-S-A theory would suggest. In addition, the relative magnitude of the effects of the dimensions of fit were dependent on the specific criterion. Therefore, for example, attitudes were best predicted in instances in which value congruence was measured along with other types of congruence (value + other), whereas turnover was best predicted by value congruence.

Validation design. The moderator analysis of validation design indicated that although the magnitude of the P-O fit–job performance relation was quite similar for both predictive ($\rho = .12$) and concurrent designs ($\rho = .14$), only the lower 95% credibility value for predictive designs was greater than zero. However, as previously noted, none of the job performance effect sizes identified for the present meta-analysis was based on applicant samples. Thus, contrary to what one might expect, even the predictive designs used incumbents. These data may therefore provide a weak test of the question at hand because, given that all study participants were incumbents, they represent a restricted pool of individuals who might have already been attracted, selected, and trimmed through attrition. Consequently, our results may represent a conservative estimate of the P-O fit–job performance relation. However, this possible threat may be mitigated by the fact that, although larger effects were obtained for predictive designs ($\rho = .41$) than for concurrent designs ($\rho = .30$) for the P-O fit–attitudes relation, when the predictive designs were disaggregated by sample type, larger effects were observed for incumbent samples ($\rho = .44$; $k = 16$) compared with applicant samples ($\rho = .37$; $k = 12$). Thus, the use of incumbent samples in the P-O fit–attitudes studies does not appear to have attenuated the observed relations.

Calculation of fit. The final moderator we investigated was the calculation of fit. For both the P-O fit–job performance and the P-O fit–attitude relations, stronger relations were observed when fit was calculated via correlations ($\rho = .13$ and $\rho = .35$ for performance and attitudes, respectively) compared with difference scores ($\rho = .05$ and $\rho = .30$ for performance and attitudes, respectively). However, the lower 95% credibility values for both job performance relations included zero.

In summary, the results of the meta-analysis indicate that P-O fit was more strongly related to work attitudes than job performance. Furthermore, the P-O fit–turnover relation, although based on a limited number of correlations, was more similar to the effect for attitudinal outcomes than the effect for job performance. In addition, all but two of the lower 95% credibility values for job performance and many of the turnover relations included zero. Comparatively, the estimated true job performance criterion-related validities for P-O fit were generally lower than those for other predictor constructs. In particular, the small P-O fit–task performance criterion-related validity was similar to those observed for years of education ($\rho = .10$) and interests ($\rho = .10$) and much lower than those of other constructs, such as general mental ability ($\rho = .51$; Schmidt & Hunter, 1998); assessment center constructs, such as problem solving ($\rho = .39$) and organizing and planning ($\rho = .37$; Arthur et al., 2003); and high-structure interview constructs, such as interpersonal ($\rho = .40$) and communication skills ($\rho = .31$; Huffcutt et al., 2001). Finally, across outcomes, the strongest effects were observed for direct–perceived operationalizations of fit and when fit was calculated via correlations. We note that, with the exception of P-O fit–performance relations from predictive designs and when fit was calculated via difference scores, for all moderator analyses only a small amount of variance could be attributed to statistical artifacts, and the standard deviations around the population estimates were relatively large, which suggests the presence of additional moderators beyond those investigated here.

The Extent to Which the P-O Fit–Job Performance and P-O Fit–Turnover Relations Are Mediated by Work Attitudes

To explore the conceptual basis for the posited relation between P-O fit and job performance, we examined the extent to which the relation was mediated by work attitudes (i.e., job satisfaction, organizational commitment, turnover intentions). We conducted a path analysis on the meta-analytically derived (corrected for unreliability of the predictor and criterion of interest) correlation matrix using the harmonic mean as the sample size ($n = 2,934$). (The correlation matrix of the meta-analytic estimates is presented in the Appendix.) Baron and Kenny’s (1986) first two steps to test mediation were satisfied, as demonstrated by the meta-analytic estimates of the relations. As reflected in the path coefficients presented in Figure 2, the P-O fit–job performance relation was reduced from .15 to .06 when we simultaneously considered job satisfaction, organizational commitment, and turnover intentions, which suggests that work attitudes partially mediated the (small) P-O fit–job performance relation.

In a similar analysis, we also explored the extent to which the P-O fit–turnover relation was mediated by work attitudes. Again, we conducted a path analysis on the meta-analytically derived corrected correlation matrix (harmonic mean $n = 1,685$). The relation between P-O fit and turnover was reduced from .24 to .12 when the effects of job satisfaction, organizational commitment, and turnover intentions were simultaneously considered (see Figure 3). These results suggest that work attitudes also partially mediated the P-O fit–turnover relation.

Discussion

The primary objective of the present study was to use meta-analytic procedures to investigate the criterion-related validity of P-O fit as a predictor of job performance and turnover. This objective was motivated by the premise that if P-O fit is going to be used for employment decision making, as increasingly appears to be the case, then measures of P-O fit must be held to the same psychometric and legal standards as are other selection tests. First, our results indicate that, contrary to the increased popularity of P-O fit in selection and other employment decision making (Rynes et al., 2002), the volume of literature investigating the criterion-related validity of P-O fit as a predictor of job performance and, particularly, turnover is limited when compared with attitudinal criteria and other constructs commonly used in personnel selection. Furthermore, although P-O fit as a predictor of job performance has received some attention, the vast majority of the job performance relations were not the focal criterion of interest in the primary studies and were presented in a tangential and supplementary manner.

Second, consistent with the absence of a strong theoretical or conceptual basis for a direct relation between P-O fit and job performance, the results of the meta-analysis suggest that P-O fit had only a small relation with job performance that did not
generalize. In addition, this small relation was also partially mediated by work attitudes, which reduces any independent contribution of P-O fit to the prediction of performance and suggests that the already small relation was most likely a function of intermediary attitudinal variables. In contrast, the relations between P-O fit and attitudinal outcomes—job satisfaction, organizational commitment, and turnover intentions—were relatively strong and generalizable. This latter finding is also consistent with the strong theoretical and conceptual precepts underlying these direct relations.

Although the P-O fit–turnover relation was more similar in magnitude to the effects observed for attitudinal criteria rather than...
performance criteria, the effect did not generalize. Results from the path analysis also indicate that the P-O fit–turnover relation was partially mediated by work attitudes. However, although it was based on a limited number of correlations, a fairly strong effect was observed for the P-O fit–turnover relation under certain conditions. For example, consistent with what would be predicted by A-S-A theory—that employees whose values do not match the organization’s are more likely to seek other employment—value-based P-O fit had stronger relations with turnover than other dimensions of fit. Promising results were obtained for P-O fit as a predictor of turnover when correlations were used to calculate fit and also when fit was operationalized according to direct–perceived measures. In relative terms, these results suggest that P-O fit may be a more efficacious predictor of turnover than job performance, and, given the currently limited research, more research with this criterion may be warranted.

Several theoretical implications and suggestions for future research can be garnered from the results of the present study. First, larger effects obtained for direct (i.e., direct–perceived) measures of P-O fit, which are judgments of how well the person fits with the organization (Kristof, 1996), are consistent with the view that it may be not one’s actual environment but rather the perception of the existence of fit that is responsible for subsequent favorable work attitudes and behaviors. However, this trend was most evident for attitudinal criteria, which may be an indication that the effects were, at least to some degree, the result of same-source bias. That is, all of the primary studies that examined P-O fit–attitudinal relations, by nature of the criteria, assessed attitudes using self-ratings. Because most studies that examined the P-O fit–attitudinal relations used self-ratings for measures of direct–perceived or indirect–perceived P-O fit, the relations may reflect some inflation due to same-source bias. Because indirect–actual measures are the target individuals’ ratings of their own characteristics compared with ratings or descriptions of the organization on the same dimensions obtained from a different source, these measures are less likely to be affected by such bias. This greater susceptibility of indirect–perceived and direct–perceived assessments of P-O fit to same-source bias may partially account for the trend of increasing P-O fit–attitudinal relations across indirect–actual, indirect–perceived, and direct–perceived measures, respectively. Consistent with this notion, the trend was less pronounced for studies that examined the P-O fit–job performance and P-O fit–turnover relations, in which the criteria were most often rated by someone other than the participant. Thus, the larger effects reported by primary studies for perceived measures of P-O fit could plausibly be overestimates of the P-O fit–attitudinal relations as a result of same-source bias.

Second, the variability in the relative magnitude of the criterion-related validities of the dimensions of fit as a function of the criteria provides additional support for the practice of conceptually matching predictor constructs to criteria. For example, although the largest P-O fit–turnover relation was obtained for value congruence, measures that included both value congruence and other congruence (e.g., goal congruence) were the best predictors of work attitudes. Thus, discussions of P-O fit as a predictor, especially of job performance, need to occur within the context of the dimensions or constructs of fit, and future research needs to investigate the potential complexity of matching the dimensions of P-O fit to specified criterion measures or work outcomes to enhance the magnitude of prediction.

Third, consistent with the notion that the rank order of ratings of P-O fit dimensions, rather than the magnitude of the discrepancy, is important for work outcomes, stronger effects were observed when P-O fit was calculated with correlations rather than difference scores. Nevertheless, researchers (e.g., Edwards, 1994; Kristof, 1996) have highlighted potential shortcomings with both of these methods of calculating congruence. Edwards and Parry (1993) and Edwards (1994) have proposed using polynomial regressions as an alternative. Despite the appeal in using these alternative procedures, we were unable to locate any P-O fit–performance studies that calculated fit using these techniques. Furthermore, many of the studies that used polynomial regressions to calculate fit and focused on attitudinal outcomes did not present enough information to calculate an effect size on the line of congruence. Thus, not only should future investigations of P-O fit and job performance use more appropriate analyses, such as polynomial regressions, they must also present sufficient information to permit conversion of the regressions to common effect size metrics and, subsequently, their inclusion in future meta-analyses.

Fourth, the study design (e.g., predictive, concurrent) appeared to moderate the magnitude of the relation between P-O fit and attitudes such that larger effects were observed for the P-O fit–attitudes relation when predictive designs were used. However, the same effect was not observed for the P-O fit–performance relation. Although predictive designs are typically associated with applicant samples and concurrent designs with incumbents, a close inspection of the included studies revealed an absence of any P-O fit–job performance effect sizes that were based on applicant samples. The extant P-O fit–job performance research, at least on the basis of the studies we identified, has exclusively used incumbents, even with predictive designs. Thus, in all of these studies, P-O fit measures were taken after the hiring decisions had already been made. The exclusive reliance on incumbent samples could result in range restriction and subsequently attenuate the P-O fit–outcome relations. This could have been addressed with a correction for range restriction, but we were unable to do so because of the absence of the necessary information in the primary studies. Thus, although the postulated range restriction effects may not be as severe because similar effects were not observed for the P-O fit–attitudes relation, it is conceivable that the P-O fit effects reported here may be conservative estimates. However, regardless of one’s position on this, the use of applicant samples to validate the P-O fit–job performance relation is a critical area of research if P-O fit continues to migrate into the prehire arena.

We note that for most P-O fit–outcome relations, a large amount of variance could not be attributed to artifacts, which indicates that the effects may represent several subpopulations. This may suggest that the P-O fit–outcome relations are moderated by additional variables not investigated here or that the relations are complex and depend on the operationalization, dimension of fit, and type of design. Therefore, future investigations that consider the role of new moderators and theoretically driven combinations of known moderators are likely to be useful in furthering our understanding of these P-O fit–outcome relations.

Practical Implications

In general, our findings raise concerns about the appropriateness of using P-O fit in hiring or selection decisions in the absence of local validation studies that demonstrate its criterion-related valid-
ity. First, the mediational results raise questions about the use of P-O fit in hiring or selection decision making, as the role of work attitudes in the fit–performance relation introduces some complexity to what is already a rather weak relation. Among other concerns, it may suggest that in using P-O fit to make selection decisions, organizations may (inadvertently) be selecting individuals on basis of subsequent employee well-being (e.g., satisfaction) instead of job performance. Whereas many organizations may deem employee well-being to be an important and desirable outcome, it appears, in our opinion, to be a rather tenuous basis for making selection decisions. Finally, although our focus has been on the use of fit in entry-level selection, our findings have implications for other employment-related decision making, including promotions, appointments to leadership positions, transfers, terminations, and even the formation of work teams.

Second, although several key discussants of the use of P-O fit in selection have focused on organizational effectiveness or other organizational-level outcomes, not individual performance, as the criteria of interest (e.g., Argyris, 1957; Bowen et al., 1991; Hambrick & Brandon, 1988; Hambrick & Mason, 1984; Schneider, 1987; Schneider et al., 1997), this position, in spite of its potential theoretical and scientific merit, is contrary to the current legal and professional standards and practice of validating employment tests against individual-level criteria (CRA, 1964, 1991; EEOC, 1978). In a similar vein, in the validation of employment tests and systems, the CRA (1964, 1991) and the EEOC (1978) speak to the use of criteria that “represent important or critical work behavior(s) or work outcomes” (EEOC, 1978, Section 14, B.3). These include “work performance” and other criteria, such as “production rate, error rate, tardiness, absenteeism, and length of service” (i.e., turnover or tenure) and “performance in training” (EEOC, 1978, Section 14, B.3). Indeed, the CRA and EEOC have not recognized work attitudes (e.g., job satisfaction, organizational commitment) as appropriate criteria for test validation in employment decision making. However, we recognize that this may technically be a nonissue if P-O fit does not display adverse impact and consequently does not violate the EEOC guidelines. Nevertheless, should it happen to do so, in spite of being relatively strong, the fit–work attitudes relation will not substantiate the use of P-O fit as an employment selection device. Furthermore, this critique is still warranted even when P-O fit is used early in the selection process by recruiters (e.g., see Brezn, Rynes, & Gerhart, 1993), because although it may not be the final decision-making tool or selection test, it still serves as an employment-related decision-making tool.

In addition to the relatedness to specified outcomes, another metric that is commonly used to evaluate the efficacy of predictors is the presence or absence of subgroup differences. Our review and search of the literature indicate that current information on potential subgroup differences on P-O fit is almost nonexistent. One study, R. S. Barrett (1995), reported an observed difference between African Americans and Whites of 0.39 standard deviations for their measure of P-O fit, compared with 1.05 for a traditional multiple-choice test in a fire fighter sample; however, the dimensions (constructs) of fit were not specified. In contrast to R. S. Barrett’s findings, in a comparison of the perceptions of organizational fit of White female, African American, and Hispanic managers with those of White male managers, Lovelace and Rosen (1996) found that African American managers reported achieving significantly poorer levels of fit compared with the other sub-

groups. Notwithstanding the presence of these two studies, one can reasonably conclude that subgroup differences on measures of P-O fit appear to be unstudied or not reported in the extant literature. It is also probably worth noting that there is little or no theoretical or conceptual a priori basis to expect EEOC (1978) designated subgroups to differ on the dimensions, such as values, typically assessed in the P-O fit literature (e.g., Harrison, Price, Gavin, & Florey, 2002). However, there may be some potential concerns regarding subgroup differences if subjective measures of fit (e.g., direct–perceived operationalizations) are particularly susceptible to contamination (see Paetzold, 2005, pp. 346–347; Rynes & Gerhart, 1990) and consequently reflect idiosyncratic interviewer preferences.

Conclusion

In summary, a review of the extant literature indicates that the majority of P-O fit studies focused on the relation between P-O fit and attitudinal criteria. Results of our meta-analysis of the criterion-related validity of P-O fit suggest that P-O fit is not a good predictor of job performance, although it may hold more promise as a predictor of turnover. Furthermore, additional analyses suggest that much of the small relation between P-O fit and job performance was due to a mediated effect of work attitudes. P-O fit’s weak relation with job performance but strong relation with work attitudes suggests that the use of P-O fit in organizations may best be limited to posthire use, such as placement, and not prehire use, such as selection. Therefore, on the basis of the results obtained in this study, we recommend that organizations should exercise caution when using P-O fit to make employment-related decisions (e.g., selection) in the absence of local validation studies or until new research refutes the findings obtained here.

References

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

*Autry, C. W. (2001). The person–organization fit of warehouse opera-


## Appendix

### Meta-Analytic Estimates of the Intercorrelations Among Study Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. P-O fit</td>
<td>—</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Job satisfaction</td>
<td>.36</td>
<td>—</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(k)</td>
<td>46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n)</td>
<td>38,099</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Organizational</td>
<td>.31</td>
<td>.69</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>commitment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(k)</td>
<td>30</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n)</td>
<td>35,291</td>
<td>17,729</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Intent to turnover</td>
<td>.25</td>
<td>.56</td>
<td>.65</td>
<td>—</td>
</tr>
<tr>
<td>(k)</td>
<td>25</td>
<td>21</td>
<td>15</td>
<td>—</td>
</tr>
<tr>
<td>(n)</td>
<td>34,938</td>
<td>17,981</td>
<td>16,209</td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>.15</td>
<td>.22</td>
<td>.34</td>
<td>.24</td>
</tr>
<tr>
<td>(k)</td>
<td>36</td>
<td>9</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>(n)</td>
<td>5,377</td>
<td>1,366</td>
<td>656</td>
<td>1,410</td>
</tr>
<tr>
<td>Turnover</td>
<td>.24</td>
<td>.36</td>
<td>.05</td>
<td>.33</td>
</tr>
<tr>
<td>(k)</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>(n)</td>
<td>2,476</td>
<td>786</td>
<td>387</td>
<td>706</td>
</tr>
</tbody>
</table>

**Note.** Table entries are estimated population (corrected) correlations. The mediation analyses were based on the correlation matrix of meta-analytic estimates. The estimates were based on the same corrections used for the overall analyses (corrected for unreliability of the predictor and criterion of interest). P-O = person–organization.