The Relationship of Age to Ten Dimensions of Job Performance

Thomas W. H. Ng
The University of Hong Kong

Daniel C. Feldman
The University of Georgia

Previous reviews of the literature on the relationship between age and job performance have largely focused on core task performance but have paid much less attention to other job behaviors that also contribute to productivity. The current study provides an expanded meta-analysis on the relationship between age and job performance that includes 10 dimensions of job performance: core task performance, creativity, performance in training programs, organizational citizenship behaviors, safety performance, general counterproductive work behaviors, workplace aggression, on-the-job substance use, tardiness, and absenteeism. Results show that although age was largely unrelated to core task performance, creativity, and performance in training programs, it demonstrated stronger relationships with the other 7 performance dimensions. Results also highlight that the relationships of age with core task performance and with counterproductive work behaviors are curvilinear in nature and that several sample characteristics and data collection characteristics moderate age–performance relationships. The article concludes with a discussion of key research design issues that may further knowledge about the age–performance relationship in the future.

Keywords: age, aging, older workers, job performance, meta-analysis

According to the Bureau of Labor Statistics, the median age of the American workforce has been increasing over the last 30 years—35 years old in 1980, 37 years old in 1990, 39 years old in 2000, and 41 years old in 2006. This trend is also evident worldwide. For instance, International Labor Organization (2005) statistics indicate that young adults between the ages of 20 and 24 were the largest segment of the working population in 1980. However, by 1990 the 30–34 age group was the largest segment of the working population, and today the largest segment of the world’s working population is the age 40–44 cohort.

Older workers are becoming an increasingly important concern for organizations for reasons beyond their sheer numbers. The shift to an older workforce has caused many organizations to spend more money on succession planning, pension benefits, health insurance, and medical benefits (Bech & Bowling, 2002; Paul & Townsend, 1993). In addition, numerous organizations have concerns (and/or stereotypes) that older workers may exhibit lower productivity (Avolio & Waldman, 1994; Greller & Simpson, 1999; Hassell & Perrewé, 1995; Lawrence, 1996). For instance, compared with younger workers, older workers are stereotyped as being less physically capable, as more likely to have problems getting along with coworkers, as preferring to invest more time in their families than in their jobs (Fung, Lai, & Ng, 2001; Paul & Townsend, 1993), as less technologically savvy, and as less willing to adapt quickly in volatile environments (Isaksson & Johansson, 2000; Riolli-Saltzman & Luthans, 2001).

Previous research has produced mixed results, however, regarding the precise relationship between age and job performance. In the three most-cited quantitative reviews of this literature, one found a moderate-sized positive relationship between age and performance (Waldman & Avolio, 1986), one found that age was largely unrelated to performance (McEvoy & Cascio, 1989), and the third found that the age–performance relationship took an inverted-U shape (Sturman, 2003). We believe that one reason for these mixed results is that much of the previous research on the age–performance relationship has focused rather narrowly on the performance of core task activities. As a result, past research has not closely examined the broad spectrum of behaviors that comprise “job performance” and the multiple ways in which age is related to work effectiveness.

Over the past 2 decades, organizational researchers have been examining numerous other job-related behaviors that also legitimately fall under the rubric of job performance. These include the following: creativity, performance in training programs, organizational citizenship behaviors (OCBs), safety performance, counterproductive work behaviors, on-the-job substance use, workplace aggression, tardiness, and absence. Although most of these job behaviors could not be called core task activities per se (Organ, 1988), they do significantly affect organizational productivity by shaping the organizational cultures and environments in which core task performance takes place (Borman & Motowidlo, 1997). As such, examining a broader and more inclusive set of job performance measures may help clarify the complex relationship between age and performance.

Mixed results on the age–performance relationship may also be partially attributable to the differing nature of research samples and data collection characteristics (Lawrence, 1996; S. R. Rhodes, 1983). For instance, research samples may vary in terms of the types of jobs workers perform, and as such, results may vary depending upon which skills older workers are required to utilize.
Similarly, because the nature of the work environment has changed substantially over the past 30 years, studies on the age–performance relationship conducted in the 1970s may have yielded very different results than studies conducted more recently. In addition, there may be differences in results depending upon whether data were collected cross-sectionally or longitudinally. For example, the effect of intraindividual aging on performance observed in longitudinal studies may be smaller in magnitude than the effect of broad age group differences observed in cross-sectional studies at any one point in time.

Thus, examining the potential moderating effects of sample and data collection characteristics is not only important for research methodology purposes but for theoretical and practical reasons too. It allows us to identify the conditions under which age is likely to have positive, zero, or negative associations with various components of job performance.

In the following section of the article, then, we briefly address some definitional issues, consider the results of previous quantitative reviews of the age–performance relationship, and discuss how the current study extends these previous reviews. Next, we present the results of an extended meta-analysis and provide evidence on the relationships between age and 10 performance dimensions. In the subsequent section, we examine the moderating effects of sample and data collection characteristics and also explore potential curvilinear relationships between age and performance dimensions. Finally, in the concluding section, we discuss the implications of our findings for future research and the management of older workers.

Theoretical Background

Definitional Issues

Age versus aging. Age is a continuous variable and is used as such in our analyses. When we refer to age differences, we are referring to group-level differences between individuals at one age and individuals at another age.

It is important to note here that, similar to previous quantitative reviews in this area of research (McEvoy & Cascio, 1989; Waldman & Avolio, 1986), the goal of the current meta-analysis is not to isolate the effects of aging per se. That is, we are not directly examining the intraindividual aging process itself and how it relates to job performance. Instead, our goal is to examine the relationships between age and job performance dimensions across different cohorts and research contexts. For instance, is age, on average, related to job performance after taking into consideration different sample characteristics and research conditions? Is the relationship between age and core task performance stronger or weaker than the relationship between age and citizenship behavior? These are the kinds of questions we attempt to address here.

Older workers. Who is considered an “older worker” has been debated in the literature for quite some time. In the retirement literature, older workers are often identified by having reached retirement age or by years until reaching retirement age (Beehr, 1986; Doeringer, 1990). Moreover, as Cleveland and Shore (1992) have noted, age can be defined in terms of an employee’s chronological age, the employee’s subjective age (the individual’s self-perception of age), the employee’s social age (others’ perceptions of the employee’s age), and the employee’s relative age (the degree to which the individual is older than others in the work group). Thus, the meaning of “old” depends, to some extent, on the demographic profiles of an organization or occupation (Shore, Cleveland, & Goldberg, 2003).

Another definition that is frequently used in this literature is the legal definition of “older worker” provided by the U.S. Age Discrimination in Employment Act of 1967 (ADEA). This act prohibits discrimination against workers who are 40 years old or above. Although our data analyses use continuous measures of age wherever possible, in our discussion of “older workers” in the text, we generally rely on the ADEA definition for a variety of reasons.

First, previous meta-analyses of age in the applied psychology literature have also utilized 40 years old as the cutoff age (Thornton & Dumke, 2005). Second, unlike authors in disciplines like gerontology and sociology (e.g., Lindenberger & Baltes, 1997) who use high cutoff ages to make fine distinctions between the “young elderly” and “old elderly” (e.g., under and over age 85), scholars in the organizational sciences are particularly attuned to the fact that the age range in the active workforce is typically 16–65 years old (International Labor Organization, 2005). Thus, at least in terms of making a dichotomous split in the workforce, 40 years old appears to be an acceptable cutoff to distinguish between younger and older workers. Third, careers researchers have observed that age 40 typically marks the end of career establishment stage and the start of career maintenance stage (Super, 1980). As such, the chronological age of 40 often represents a major transition in career stages as well. Finally, defining older workers as 40 or above has some practical benefits because it directly aligns research findings regarding older workers to management implications regarding ADEA compliance in hiring, termination, performance evaluation, and promotion decisions.

Undoubtedly over time, the definition of “older worker” will change. For example, the amendment of ADEA (by the Older Workers Benefit Protection Act of 1990 and the Civil Rights Act of 1991) to prohibit mandatory retirement ages may ultimately push back the age at which people retire (although labor statistics have not indicated any increases in retirement age as of yet). Even more likely, gains in life expectancy will change our conceptions of who is “middle-aged” and who is “old.” For our current purposes, though, the definition of older workers as being age 40 or older is consistent with both previous research and legal definitions and has the benefit of being “objectively” determined and verifiable across researchers and contexts.

Previous Reviews of the Age–Performance Relationship

Three major quantitative reviews of this literature have been published in the last 25 years. Waldman and Avolio’s (1986) review is the earliest meta-analysis in this area. The authors identified 13 empirical studies containing 40 samples. They found that age was positively related to productivity measures of job performance (.27). On the other hand, age was weakly but negatively related to supervisor ratings of job performance (−.14). Furthermore, Waldman and Avolio found that the relationship between age and supervisor-rated job performance was stronger for nonprofessionals (−.18) than for professionals (−.05). Peer ratings of job performance were related to age at .10.

In sum, Waldman and Avolio (1986) illustrated that the sign of the relationship between age and task performance varies depend-
ing upon which measure of performance is being utilized, who
does the performance ratings, and what kinds of jobs workers hold.
To the extent that there is a drawback to this meta-analysis, it is the
lack of availability of a large number of studies at that time.
Meta-analyses that include too few cumulative studies may contain
second-order sampling errors (Hunter & Schmidt, 1990).

Addressing this drawback, McEvoy and Cascio (1989) identi-
ified 65 empirical studies (containing 96 samples) conducted on the
relationship between age and performance. Across these 96 sam-
ple cases, the authors found that the mean correlation between age and
job performance was quite low (.06) and that the confidence
intervals contained the value of zero. Unlike Waldman and Avolio
(1986), McEvoy and Cascio found that type of performance rating
(productivity vs. supervisory rating) and job type (professional vs.
nonprofessional) did not moderate the relationship between age
and performance. The different results observed in these two
meta-analyses may be attributable to the different sets of studies
the researchers considered, as McEvoy and Cascio located a wider
range of empirical studies than Waldman and Avolio had included.

Sturman’s (2003) meta-analysis hypothesized that the relation-
ships of performance with three age-related variables (chronolog-
ical age, job experience, and organizational tenure) were in the
form of an inverted-U shape. With respect to the age–performance
relationship, Sturman found that the corrected effect size across
115 empirical studies was .03. Although this effect size is very
small, he did find that this relationship was indeed an inverted-U
shape. That is, age was positively related to job performance when
age was low but was negatively related to job performance when
age was high (>49 years old).

Below, we highlight the main evidence supporting these three
different perspectives on the age–performance relationship.
Whereas earlier research on older workers largely focused on the
negative relationship between age and task performance (S. R.
Rhodes, 1983), more recently researchers have been examining the
ways in which age can facilitate task performance or, at the
minimum, not adversely affect it (Ebner, Freund, & Baltes, 2006;
Kanfer & Ackerman, 2004). As a result, we have a much richer
picture now of how age is positively or negatively related to core
functions, that is, monitoring and controlling attention, suppressing
irrelevant information, utilizing analytical reasoning, and updating
information in working memory. Older individuals were found to
perform much more poorly on this test battery than their younger
counterparts. In general, M. G. Rhodes’s results suggest that older
individuals may have more difficulties with complex tasks that
require a high level of executive functioning. Indeed, there is also
cumulative empirical evidence to indicate that older individuals do
not do as well as younger individuals when performing multiple
complex tasks simultaneously (Verhaeghen, Steitz, Sliwinski, &
Cerella, 2003).

Another area in which age appears to have negative association
with performance is memory capacity. Previous meta-analyses have
demonstrated a significant negative relationship between age and
memory. For instance, older adults were found to have poorer
recognition and recall memory than younger adults (La Voie &
Light, 1994; Spencer & Raz, 1995; Verhaeghen, Marcoen, &
Goosens, 1993). Moreover, as a result of these memory differ-
ences, employees are less likely to trust the memories of older
coworkers. In an experimental study of attribution theory, for
example, Erber and Danker (1995) found that participants ex-
pected memory-related performance problems of older workers to
continue longer than those of younger workers and were less likely
to recommend training when “problem” employees were older.

Above and beyond these differences in aptitudes and short-term
memory, researchers have also found that older individuals may
have less intense work motivation than their younger colleagues
(S. R. Rhodes, 1983). For instance, Ebner et al. (2006) found that
younger individuals are more likely to frame their goal orientations
in terms of striving for gains (e.g., I want to improve my fitness),
whereas older individuals are more likely to frame their goal
orientation around maintaining the status quo or preventing loss
(e.g., I do not want my fitness to deteriorate). These changes in
motivation may also be linked to lower productivity on the job.

Evidence Supporting Positive or Neutral Relationships of
Age With Performance

The above literature paints a rather pessimistic view of the
relationship of age with job performance. Nonetheless, a strong
case can also be made that older workers may exhibit at least the
same, if not greater, job performance as their younger colleagues
(Greller & Simpson, 1999). The rationale most frequently cited to
support this case is that older workers substitute lengthy job
experience and greater general expertise for speed of information
acquisition and information recall. This wisdom and expertise,
accumulated over the course of a career, may be sufficient to
compensate for productivity losses due to any changes in cognitive
and physical abilities (Baltes, Staudinger, Maercker, & Smith,
1995).

Kanfer and Ackerman (2004) have emphasized that older age is
often accompanied by increases in “crystallized” intelligence (ex-
periential knowledge). Indeed, experimental evidence provides
robust support for Kanfer and Ackerman’s assertion. For instance,
Allen, Lien, Murphy, Sanders, and McCann (2002) found that
older participants could multitask as effectively as younger partic-
pants, albeit at a slower pace. Artístico, Cervone, and Pezzuti
(2003) found that older adults’ performance in solving problems
exceeded that of younger adults when the problems were familiar

Evidence Supporting Negative Relationships of Age With Performance

Numerous studies have found support for the proposition that
age negatively relates to cognitive functioning. For instance, in a
large sample of 20,000 American workers across multiple occu-
pational groups, Avolio and Waldman (1994) found that age was
negatively related to several types of aptitudes, including general
intelligence, verbal aptitude, numerical aptitude, spatial aptitude,
form perception, clerical perception, motor coordination, finger
dexterity, and manual dexterity. Thus, for jobs in which general
cognitive abilities, visual-perceptual abilities, and psychomotor
abilities are important components for superior job performance,
Avolio and Waldman have suggested that age is negatively related
to job performance.

Furthermore, M. G. Rhodes (2004) found that there was a strong
and significant difference between older and younger adults in
performance on a test battery measuring individuals’ executive
and representative of tasks frequently encountered. Colonia-Willner (1998) found that the best performing older employees had higher levels of tacit knowledge than their younger employees. Studies using different research designs have also found that professional expertise, developed over years of practice and experience, can attenuate potential negative relationships between age and performance dimensions (Hess & Auman, 2001; Lindenberg, Kliegl, & Baltes, 1992; Morrow, Leirer, Altieri, & Fitzsimmons, 1994; Thornton & Dumke, 2005; Wilson, Li, Bienias, & Bennett, 2006).

Taken together, the research described above suggests that, after a slower pace of initial learning, older workers can reach the same performance levels as those of their younger colleagues and can multitask effectively. Moreover, when older workers are asked to solve familiar problems, higher self-efficacy beliefs are activated, and these beliefs can accelerate performance. Thus, although fluid intelligence, short-term working memory, and cognitive speed may decrease with age, deductive reasoning and professional expertise are likely to increase (Masunaga & Horn, 2001). Moreover, increased wisdom and judgment gained over years of service may increase older workers’ effectiveness in contextual performance activities as well.

Expanding the Domain of the Performance Construct

Previous reviews of the age–performance relationship have primarily focused on the performance of core tasks. According to Borman and Motowidlo (1997), core task performance is concerned with “the effectiveness with which job incumbents perform activities that contribute to the organization’s technical core” (p. 99). In addition to core task performance, however, we also examine nine other performance dimensions that comprise the two broad categories of job behaviors identified by Hunt (1996) as independent of the core job role, namely, citizenship behaviors and minimum performance behaviors.

Citizenship behaviors are those extra behaviors engaged in by employees, over and above their core task requirements, that actively promote and strengthen the organization’s effectiveness (Organ, 1988). In this study, the first category is represented by such dimensions as employee creativity, performance in training programs, citizenship behaviors geared to different beneficiaries, and safety performance. In contrast, minimum performance behaviors are those that employees have to engage in (like attending work) or refrain from engaging in (like theft) to keep their jobs (Hunt, 1996). This second category is represented by such dimensions as general counterproductive work behaviors, workplace aggression, on-the-job substance use, tardiness, and absenteeism.

Conceptually, there are certainly other behavioral dimensions that could be included under these two categories. For example, working long hours and demonstrating effective leadership are examples of additional citizenship behaviors that might be considered, whereas refraining from sexual harassment and manipulating stock prices are additional examples of minimum performance variables that could be considered. Empirically, however, a meta-analysis is constrained by the number of previous studies conducted on a specific relationship, and here we have included the nine dimensions identified above as the ones on which the most empirical studies are available. We discuss each of these additional nine dimensions of job performance in more detail below.

Creativity

Creativity is the extent to which employees generate new and useful ideas for improving organizational productivity (Anderson, De Dreu, & Nijstad, 2004). For many jobs, creativity might be considered as a separate element of job performance, particularly when creativity relates to organizational adaptability and flexibility. For instance, some researchers have emphasized the importance of employees’ creativity as a critical component of an organization’s ability to adapt to rapidly changing business environments (A. De Jonge & De Ruyter, 2004; Johnson, 2001). Consistent with this view of creativity as a key element of job effectiveness, a major study of performance evaluation systems found that some organizations evaluated employees on their innovativeness as well as on their core task performance (Welbourne, Johnson, & Erez, 1998).

Performance in Training Programs

Older workers are often stereotyped as being somewhat resistant to change and slow in learning new material. For this reason, researchers have examined older workers’ performance in training programs as an element of job effectiveness (Martocchio, 1994). As Tracey, Tannenbaum, and Michael (1995) have suggested, organizations typically provide training to employees on the basis of the assumption that the short-run costs of the design and execution of training can be recouped through employees’ increased productivity in the long run. However, if employees do not perform well in training programs, it is highly unlikely that they will transfer that new knowledge to real work settings. Furthermore, when employees fail to learn in training programs, the expenses associated with training are wasted as well (Winfred, Bennett, Edens, & Bell, 2003).

OCBs

Researchers have documented the importance of OCB for organizational functioning over the last 2 decades (LePine, Erez, & Johnson, 2002; Organ, 1988). OCBs (sometimes called prosocial behaviors or extrarole behaviors) are not job-specific but rather support the broader organizational environment in which core performance takes place (Motowidlo & Van Scotter, 1994; Organ, 1988). Examples of OCB are compliance with organizational norms, not complaining about trivial matters, and helping coworkers. Employees’ aggregated OCBs frequently benefit group, unit, and organizational productivity (Podsakoff, MacKenzie, Paine, & Bachrach, 2000).

Safety Performance

Safety performance is the extent to which employees comply with safety rules and demonstrate safe behaviors in the workplace (Parker, Axtell, & Turner, 2001). This performance dimension is particularly important in industries that require employee contact with hazardous materials, operation of heavy machinery, and extensive highway driving (Clarke & Robertson, 2005). Poor safety performance can have two distinct negative consequences for firms’ effectiveness. At the individual level, inattention to safety behaviors contributes to employee injuries; these injuries result in lower worker productivity and time lost from work. At the orga-
nizational level, poor safety practices contribute to potentially costly litigation. These claims create substantial financial burdens for firms in the event of serious employee accidents, dismemberments, and death (Hofmann & Morgeson, 1999). Both directly and indirectly, then, employee safety behaviors are an important component of job performance.

**General Counterproductive Work Behaviors**

Whereas research investigating OCB concentrates on what workers can do to promote smooth organizational functioning, research on workplace deviance examines how the lack of counterproductive work behaviors is essential to maintaining smooth organizational functioning (Neuman & Baron, 1998). Counterproductive work behaviors are intentional employee acts that harm organizations’ legitimate business interests (Bennett & Robinson, 2000). Examples of counterproductive work behaviors include working on personal matters instead of assigned tasks, neglecting supervisors’ instructions, stealing property, starting or repeating rumors and gossip, and using unprofessional language. It is easy to see the multiple ways in which these counterproductive work behaviors can reduce both individual and group performance. Moreover, Dalal (2005) found that employees who frequently engaged in counterproductive work behaviors were also less likely to demonstrate OCBs.

**Specific Counterproductive Work Behaviors**

In addition to general counterproductive work behaviors, four specific forms of counterproductive work behavior have been discussed separately and extensively in the literature. We therefore examine these four specific counterproductive work behaviors—workplace aggression, on-the-job substance use, tardiness, and absenteeism—in the present meta-analysis.

**Workplace aggression.** Workplace aggression consists of employees’ efforts to harm others with whom they work, harm the reputation of their current employers, or harm former colleagues and previous employers (Lapierre, Spector, & Leck, 2005). Acts of workplace aggression can cause bodily harm to employees, pose physical danger for customers, create public relations crises, and damage the business reputation of the firm as a whole.

**On-the-job substance use.** On-the-job substance use involves drinking alcohol or taking illegal drugs at work or during work time (Fronc, 2003). Researchers have found that on-the-job substance use hampers individuals’ decision-making abilities; increases the frequency of dysfunctional job behaviors; and puts coworkers, supervisors, and customers at increased risk of injury (Lehman & Simpson, 1992).

**Tardiness.** Tardiness is lateness for work (Blau, 1994; Koslowsky, Sagie, Krausz, & Singer, 1997). Employee tardiness is likely to create both direct financial costs to organizations (e.g., decreased time on productive activities) and indirect financial costs (e.g., time lost by coworkers waiting for late colleagues). Left unchecked, numerous cases of tardiness can lead to a “culture of tardiness” (Koslowsky et al., 1997) in which employees come to see being late as an acceptable behavior rather than as a deviant one.

**Absenteism.** Skipping work has also been conceptualized as a form of employee counterproductive behavior (Bennett & Robinson, 2000; Martocchio, 1989). Researchers have documented numerous negative effects of employee absence on organizational productivity (Harrison & Martocchio, 1998). When employees are absent from work, the completion of their own work is slowed down. Coworkers are often called upon to cover for absent employees, thereby distracting them from completing their own assignments. In cases in which task interdependence among a group of workers is high, the whole team’s progress may be affected when an employee is chronically absent or absent for extended periods of time. As with tardiness, frequent absences can also adversely affect organizational productivity by creating an “absence culture” in which more and more employees consider being absent acceptable (rather than counterproductive) behavior (Johns & Xie, 1998).

**Summary.** Before we can draw strong conclusions about the relationship of age to job performance, then, it is important to consider citizenship behaviors and minimum performance behaviors in addition to core task performance. In light of the above literature review, we include 10 dimensions of job performance in the current meta-analysis: core task performance, creativity, performance in training programs, OCB, safety performance, general counterproductive work behaviors, workplace aggression, on-the-job substance use, tardiness, and absenteeism. From this point on, we use the phrase “job performance” to represent these 10 dimensions collectively.

**Moderator Relationships**

Another way in which the present study contributes to the literature is by investigating how different sample and design characteristics moderate the relationship between age and job performance. Many of these characteristics have been discussed in the literature as variables that can affect age–performance relationships (Lawrence, 1996; S. R. Rhodes, 1983; Shore et al., 2003; Sturman, 2003) and as potential explanations for inconsistent research findings in the area.

**Sample Characteristics**

In this study, we examine the potential moderating effects of the average age, age dispersion, job tenure, and organizational tenure of research samples. Testing for average sample age as a moderator essentially examines whether the form of the relationship between age and performance is linear or curvilinear (see Sturman, 2003). Testing the moderating role of age dispersion associated with the sample (operationalized as the standard deviation of age in the sample) assesses whether the age–performance relationship varies across samples with different degrees of age homogeneity. Testing for average job tenure and average organizational tenure examines whether the age–performance relationship varies across samples with different (average) tenures. In general, these four sample characteristics might influence authors’ definitions of “old” in a particular setting.

We also examine the potential moderating effects of job complexity. The job complexity of the sample is important to consider here because it influences the extent to which the specific job skills required on jobs are associated with age-related performance problems. For the same reason, we also consider two related variables, namely, proportion of college degree holders and proportion of
managers in the sample. It is likely that older workers with more education or who hold managerial roles may have greater ability to substitute accumulated knowledge and judgment for precise technical skills. As such, the results of research on age–performance relationships may vary across samples with different mixes of college graduates and managers.

In a more exploratory fashion, we also examine the effects of proportion of women and proportion of Caucasians in the sample. Gender and race have frequently been studied as independent or control variables in previous research on aging, but here we consider their potential moderating effects instead. Examining these two “proportion” moderators assesses whether the age–performance relationship varies across samples with different proportions of women and Caucasians.

**Data Collection Characteristics**

We examine the moderating effects of two data collection characteristics in particular: (a) whether data were collected longitudinally or cross-sectionally, and (b) the publication year of articles. Both these moderators address the role of time in the age–performance relationship.

Cross-sectional data collection allows researchers to examine the extent to which there are group-level age differences (between different age cohorts) in job performance across different samples and research contexts. In contrast, longitudinal data collection helps shed light on the extent to which intraindividual aging is related to job performance (Baltes, Schaeie, & Nardi, 1971; Schaeie & Hofer, 2001). Results may vary, then, depending upon how the data in a study were collected. For instance, the effect of intraindividual aging on performance may be subtle and gradual, but the differences in performance between broad age groups at one point in time can be quite large. Thus, we examine the distinction between cross-sectional and longitudinal data collection as a potential moderator here.

It should be noted that to truly examine the effect of aging on performance, it would be ideal to focus the meta-analysis on longitudinal studies alone. However, there are too few articles available to conduct a meta-analysis on only longitudinal studies; in fact, only 12% of the articles on this topic have been longitudinal in nature. Further, even if we tried to examine only longitudinal studies, the time period examined in those studies does not exceed 5 years, and the average period of time examined in those studies is only about 1 year. Thus, these studies generally do not use long enough time spans to capture intraindividual aging either.

Thus, our database here largely consists of cross-sectional studies, allowing us to get an overall picture of the strength of various age–performance links, considering the variation in the specific performance measures, sample characteristics, and data collection characteristics. We revisit the limitations of cross-sectional data collection in conducting research involving age in more detail later in the article.

The year a particular study was published is used here as a rough proxy for the nature of the work environment at the time a study was conducted. For instance, the last 2 decades have seen increased reliance on career self-management (Arthur & Rousseau, 1996), which may have loosened workers’ attachment to long-time employers and increased workers’ willingness to change careers later in life. The introduction of the Family and Medical Leave Act of 1993 may have changed the ways in which organizations distinguish between “absence” and “leave” before and after that year. Similarly, the amended ADEA prohibits mandatory retirement ages, which might have changed the relationship between age and performance over time because of a potential change in the proportion of older workers in the labor force. Collectively, then, these changes may have affected the context in which older workers perform and thus the relationship of age to performance as well.

**Method**

**Literature Search**

We performed a comprehensive search for those articles published during or before December 2006 that examined the relationship between age and job performance. Other published studies that did not aim at investigating this specific relationship but reported their effect sizes nonetheless were also included. We also searched for unpublished studies and dissertations to reduce the “file-drawer problem” (Rosenthal, 1979).

We began our literature search by using the following keywords: age, job performance, task performance, productivity, creativity, innovation, training effectiveness, training performance, contextual performance, citizenship behavior, prosocial behavior, extrarole behavior, safety, injuries, accidents, counterproductive behavior, deviance, deviant behavior, aggression, aggressive behavior, violence, violent behavior, substance use, drinking, alcohol, drug, tardiness, lateness, absenteeism, absence, and job complexity. We searched in numerous research databases, including Dissertation Abstracts International, EBSCOHost, Emerald, Factiva, JSTOR, Oxford Journals, Proquest, PsycINFO, ScienceDirect, Sage Full-Text Collections, and several Wiley InterScience databases.

Furthermore, the reference lists of recent meta-analyses that focused on criterion variables of interest were examined carefully to locate other relevant articles. These included, for instance, reference lists from meta-analyses on OCB (LePine et al., 2002), absence (Martocchio, 1989), task performance (Judge, Thoresen, Bono, & Patton, 2001), work injuries (Clarke & Robertson, 2005), and counterproductive work behavior (Dalal, 2005). Such meta-analyses contained comprehensive and current reference lists in their respective research fields.

Four inclusion and exclusion criteria were set prior to the start of the article search. First, we included field studies in which the age–job performance relationship occurred naturally, whereas laboratory studies were excluded. It is much more difficult to gather measures on performance dimensions besides core task performance in experimental settings (e.g., lateness and absence). In addition, we had some concern that core task performance observed in laboratory studies might be partially attributable to the strength and form of the experimental manipulations themselves. Second, studies that examined age or performance only at the team level or organization level were excluded because our focus in this research is on individual behavior rather than group-level performance. Third, to avoid double counting, we excluded those studies in which authors used the same data set and reported the same correlations as in their other published studies. Fourth, we included studies that involved four types of performance ratings: ratings by
supervisors, ratings by others (peers, subordinates, and customers), self-ratings, and objective measures. This allowed us to examine the moderating role of sources of rating in the age–performance relationship.

With the above search criteria, our search yielded a total of 380 empirical studies, which collectively contain 438 independent samples. Thirteen studies were unpublished dissertations. Of the studies, 14% were published before 1990, 30% of the studies were published between 1990 and 1999, and the remaining 56% were published between 2000 and 2006. The average age across all samples that provided age information was 36.6 years old (ranged between 17 and 59 years old with a standard deviation of 8.8 years). The list of studies is provided in the Appendix.

Measures of Key Constructs

Creativity. Previous studies have measured creativity either via self-ratings or ratings by others; we used these two categories in the meta-analysis as well. Sample items of typical Likert-scale measures of creativity include the following: “creating new ideas for improvements”; “searching out new working methods, techniques, or instruments”; and “generating original solutions to problems” (Janssen, 2001). It should be noted that we did not differentiate between creativity and innovation (Anderson et al., 2004); both types of measures are included in our study.

Performance in training programs. Here, we focused on studies that involved training of adults on tasks that have at least some relevance in organizational contexts. Furthermore, only studies that had an explicit training intervention and had measured post-training performance, competence, or learning were included. Measures of training performance consisted of either ratings given by supervisors or performance on post-training tests.

Studies that assessed employees’ participation in computer usage training are representative of the kinds of research articles included in this meta-analysis (e.g., Martocchio, 1994). Other example studies include training programs that teach specific competences or subject-related knowledge (e.g., law or health care). Although our decision to include only studies that measured post-training performance gave us a consistent perspective on the effects of training on performance, as we discuss later, the resulting set of studies was likely overrepresented by technology training programs in which older workers might be less likely to excel.

OCBs. We included two types of OCB in the meta-analysis. The first set of studies examined general OCBs and did not differentiate among beneficiaries of those OCBs. The second set of studies examined OCBs geared to three specific beneficiaries: other people on the job, the employer organization as a whole, and the tasks themselves. These subtypes have been identified by previous researchers as reasonable groupings of behaviors in this domain (LePine et al., 2002).

According to LePine et al. (2002), examples of OCB directed to others are helping colleagues with their work and orienting newcomers. Examples of organization-directed OCB are compliance with organizational norms and not complaining about trivial matters. Examples of task-directed OCB are spending extra effort and persistence on the job and trying hard to improve personal and group performance. It should be noted, too, that within each of these three subtypes, we further differentiated self-ratings from ratings by others.

Safety performance. With respect to the measurement of safety performance, three major indicators have frequently been used in previous research and were included in the present meta-analysis. First, some studies have measured the frequency of work injuries via archival industrial health records. For instance, Hofmann and Morgeson (1999) measured workplace injuries by using the company’s archival records of injuries or accidents that occurred in the previous 12-month period. Second, other studies have utilized self-report measures of frequency of work injuries. A third set of studies have measured self-reported levels of compliance with safety procedures and practices. A sample item from this kind of measure is “Occasionally I bend the safety rules when I know it’s safe to do so (reverse coded)” (Parker et al., 2001).

General counterproductive work behaviors. Most studies have measured general counterproductive work behaviors without differentiating targets, that is, without specifying the target of the counterproductive behavior. A few studies have differentiated between counterproductive work behaviors directed at specific others and those directed at the organization as a whole (e.g., Liao, Joshi, & Chuang, 2004). However, because of the small number of studies making this distinction, the current meta-analysis did not differentiate between these two categories of studies. In those few studies that reported both interpersonal and organizational counterproductive work behaviors, we averaged the correlations to obtain an estimate of general counterproductive work behavior.

Sample Likert items of measures of general counterproductive work behaviors include the following: “I keep important information away from my boss” and “I openly compromise with others but delay implementing the compromise until my own objectives are accomplished” (Duffy, Ganster, & Shaw, 1998). Here, too, we differentiated between self-ratings of counterproductive work behavior and ratings by others.

Workplace aggression. As mentioned above, four specific types of counterproductive work behaviors have frequently been measured in previous research on this topic, and therefore we examined these four specific types of counterproductive behavior in more depth. Measures of workplace aggression typically ask respondents to indicate the frequency of occurrence of aggressive behaviors, such as yelling, swearing at others, damaging others’ property, and fighting (Glomb & Liao, 2003). All the studies we located utilized self-reported measures.

On-the-job substance use. These measures typically ask respondents to indicate the frequency of on-the-job use of alcohol or drugs (Frone, 2003). Here, too, all the studies identified utilized self-ratings.

Tardiness. Tardiness is typically measured in two ways. In the first, employees are asked self-report questions like “How often are you late from work? (never to constantly)” (Hanisch & Hulin, 1990). In the second, archival measures of lateness are obtained directly from personnel records (Conte & Jacobs, 2003). We included both self-report and archival measures of tardiness in the meta-analysis.

Absenteism. Absenteism has been measured in three different ways in previous research. The first group of studies measure general absenteism; these studies do not differentiate between when employees are absent because of sickness or for purely discretionary reasons (Xie & Johns, 2000). Other studies in this line of research measure either the number of days absent from work in a given period (absence duration) or the frequency of
absence spells in a given period (absence frequency). Because these indices are all closely related (Conte & Jacobs, 2003), they are aggregated together in the present meta-analysis.

A second research stream includes studies that measure sickness absenteeism. As an example, J. De Jonge, Reuvers, Houtman, and Kompier (2000) computed sickness absence as the number of separate spells of sickness absence during 1 full calendar year. Researchers have traditionally viewed absenteeism due to sickness as involuntary absence (Dalton & Todor, 1993).

The third, and last, group of absence studies consists of those that measure nonsickness-related absenteeism. For instance, Vignoda (2001, p. 1499) asked respondents to report “their estimates of days missed work (during the previous year) for reasons other than sickness.” As another example, Deery, Erwin, and Iverson (1999) obtained personnel records of frequency of nonmedically certified absences during the prior 12-month period. In contrast to sickness-related absence, researchers have generally viewed nonsickness-related absence as an indicator of voluntary withdrawal behavior (Dalton & Todor, 1993).

Meta-Analytical Procedures

Hunter and Schmidt’s (1990) meta-analysis technique, which requires corrections for both measurement error and sampling error, was used. Because there is no theoretical reason to believe the measurement of age would contain measurement error, we followed the practice of previous researchers (Martocchio, 1989) and did not disattenuate the measurement of age in individual studies. On the other hand, the behavioral measures of job performance did require disattenuation to remove the influence of measurement errors.

Disattenuation of core task performance. Some researchers have reasoned that it is more appropriate to use intrarater reliability to correct for imperfect measurement when task performance is rated by others (i.e., not by oneself; Judge et al., 2001; Schmidt & Hunter, 1996; Viswesvaran, Ones, & Schmidt, 1996). In contrast, other researchers argue that measures of intrarater reliability (that is, alpha coefficients or internal consistency estimates) are more appropriate in this regard (Murphy & De Shon, 2000). In light of these different perspectives on disattenuation, we disattenuated the observed correlations for either imperfect intrarater reliability or imperfect interrater reliability.

We first corrected studies for the lack of perfect intrarater reliability in the measurement of core task performance. Because different studies may measure the same construct using different scales, this disattenuation process also adjusts for the different levels of measurement error contained in different scales (Hunter & Schmidt, 1990). This type of correction requires the use of alpha coefficients (i.e., internal consistency estimates) reported in individual studies. If no alpha value was reported for a particular scale in a study, an average alpha value calculated from the rest of the studies using the same scale was taken as a substitute (Judge et al., 2001). The above procedure for correction for imperfect intrarater reliability was performed on all self-report measures and on all measures provided by supervisors, peers, and customers.

We corrected studies for imperfect intrarater reliability when core task performance was measured via ratings by others (i.e., not via self-reports). This type of correction required the use of interrater reliability estimates. Unfortunately, many studies included in the meta-analysis did not report these estimates. Therefore, following Judge et al. (2001), for studies that used supervisors as the rating source, we disattenuated the correlations using the meta-analytical estimate of the reliability of supervisor-rated job performance provided by Viswesvaran et al. (1996). For studies that used peers to rate job performance, we disattenuated the correlations using the meta-analytical estimate of the reliability of peer-rated job performance. Because there is a lack of research on interrater reliability associated with less frequently-used raters (i.e., customers, subordinates, students, or spouses), in these cases we utilized as a proxy estimate the average of the meta-analytical estimate of the reliability of supervisor-rated performance and that of peer-rated job performance.

As noted earlier, several studies utilized objective measures of core task performance (e.g., quantity of output, sales volume, errors made). Although the notion of intrarater and interrater reliability does not apply to objective measures, it should be noted that these objective measures are still likely to contain some measurement error and might vary depending upon the job complexity of the sample (Sturman, Cheramie, & Cashen, 2005). Because studies seldom reported any kind of reliability estimates for objective measures of core task performance, as a substitute we adopted the meta-analytical estimate of the test–retest reliability of the objective measure of job performance provided by Sturman et al. (2005). Sturman et al. calculated two separate estimates of the reliability of objective task performance, one for jobs of low complexity (e.g., secretary) and one for jobs of high complexity (e.g., engineer). For those studies that sampled mixed job complexity types, we used the average value of Sturman et al.’s two estimates as the proxy.

Disattenuation of other performance measures. The disattenuation of the observed correlations between age and the remaining dimensions of job performance largely followed the protocols outlined above. For instance, for self-ratings of creativity, OCB, counterproductive work behaviors, workplace aggression, tardiness, and absence, we corrected for imperfect interrater reliability using either alpha coefficients (if provided) or an average alpha value based on the remaining studies that used the same scales (if alpha was not provided). For nonself ratings of these performance dimensions, in addition to correcting for imperfect intrarater reliability using the above procedure, we also corrected for imperfect interrater reliability using Viswesvaran et al.’s (1996) meta-analytical estimates of the reliability of job performance. Finally, for performance measures that were objective in nature (e.g., training test scores and company records of work injuries, lateness, and absence), we used Sturman et al.’s (2005) meta-analytical estimate of test–retest reliability of the objective performance measure as a substitute.

Corrections for sampling errors. The second step in the meta-analysis was correction for sampling errors. Here, we calculated the sample size–weighted corrected correlation. A corrected correlation was judged to be significant at $\alpha = .05$ when its 95% confidence interval did not include the value of zero.

Moderator Analyses

Subgroup moderator analyses. Two approaches to moderator testing were used. The first one was subgroup analysis; this ap-
proach was taken when the moderator variables were categorical in nature (e.g., longitudinal vs. cross-sectional data collection).

The coding of most of these categorical moderator variables is self-explanatory. The coding for the job complexity and age moderators, though, requires some further explanation.

In the case of job complexity, two researchers were responsible for the coding. The coding process was guided by previous meta-analyses that also coded job complexity (e.g., Avolio & Waldman, 1990; Salgado et al., 2003; Wood, Mento, & Locke, 1987). Specifically, the two raters classified each sample occupation into high and low job complexity according to the general intelligence, verbal ability, and numerical ability required to perform the job (Avolio & Waldman, 1990). Interrater agreement was 93%. In situations in which there was disagreement, discussion was used to reach consensus. The Dictionary of Occupational Titles (1991) was used to assist in these discussions, too, because jobs in the Dictionary of Occupational Titles are coded and classified according to several dimensions (e.g., data, people, and things) that reflect job complexity (Avolio & Waldman, 1990; Salgado et al., 2003). Examples of “high complexity” jobs are researchers, accountants, business consultants, psychiatrists, engineers, managers/executives, financial analysts, nurses, IT professionals, and teachers. “Low complexity” jobs include clerks, restaurant workers, highway maintenance workers, truck drivers, and receptionists.

In those cases in which we needed to examine the differences across different age subgroups, we sorted studies into four groups on the basis of the average age of the sample: less than 30 years old, 31–35, 36–40, and over 40 years old. The reason underlying this categorization is four-fold. First, it approximately equalized the number of studies contained in each group. Second, previous research on life-span development has suggested that the time before 30 years old is often a period within which individuals explore careers, whereas the time after 40 years old roughly marked the end of establishment years (Super, 1980). Third, ADEA also uses 40 as the “cutoff age.” Last, much of the previous research on age and job performance has used decade or half-decade benchmarks.

Regression-based moderator testing. In other cases, the original studies did not provide sufficient sample information with which to categorize moderators. For these moderator variables, therefore, regression-based moderator testing was used. To illustrate the statistical procedures, we utilize the example of “proportion of women” (i.e., gender). We tested the plausibility of gender as a moderator by investigating the effect of the percentage of women in the sample (which was a continuous variable) on the observed relationship between age and job performance (cf. Ng, Eby, Sorensen, & Feldman, 2005). In essence, we used the percentage of women in the sample as an independent variable, in a weighted least squares multiple regression, to predict the correlation coefficients for the age–job performance relationship. If the percentage of women in the sample was a significant predictor of a relationship between age and a dimension of job performance, then it would suggest that gender moderated that relationship.

It should be noted that, among the eight sample characteristics on which we applied this regression approach, only average age, age dispersion, average job tenure, and average organizational tenure of the sample were truly continuous variables per se. The remaining characteristics (proportion of degree holders, managers, women, and Caucasians) were proxies for sociodemographic variables that are typically measured as categorical variables at the individual level. Nonetheless, this regression technique for testing for moderators in meta-analyses has been found to be more reliable and robust than alternative methods for dealing with this analysis challenge (Steel & Kammeyer-Mueller, 2002).

Results

The meta-analysis results for the relationships between age and the 10 performance dimensions are presented in Table 1.
related to workplace aggression at \(-.08\) and negatively related to on-the-job substance use at \(-.07\). Age was particularly strongly and negatively related to tardiness (as rated by supervisor or measured objectively) at \(-.26\) (\(-.28\) when corrected for interrater reliability). The effect was smaller, though, when tardiness was measured by self-ratings (effect size of \(-.12\)).

The fourth specific counterproductive work behavior we examined was absenteeism. Here, the effect sizes of age depended on measured by self-ratings (effect size of \(.19\)).

The fourth specific counterproductive work behavior we exam-

Table 1

\begin{tabular}{lrrrrrr}
\hline
Variable & \(N\) & \(k\) & \(r_c\) & \(SD_c\) & 95\% LCI & 95\% UCI & \(Q\) \\
\hline
Core task performance & & & & & & & \\
Supervisor-rated & 52,048 & 118 & .02 & .10 & .02 & .03 & 576.82* \\
(Corrected for interrater reliability) & & & .03 & .13 & .02 & .04 & 987.57* \\
Rated by peers and others & 1,555 & 7 & -.02 & .09 & -.08 & .04 & 15.02 \\
(Corrected for interrater reliability) & & & -.04 & .15 & -.09 & .03 & 32.69* \\
Objective measures & 8,970 & 28 & .03 & .16 & .01 & .05 & 221.93* \\
Self-rated & 17,807 & 52 & .06 & .11 & .05 & .08 & 252.47* \\
Creativity & & & & & & & \\
Rated by supervisor or objective measures & 1,662 & 9 & .01 & .12 & -.05 & .07 & 27.36* \\
(Corrected for interrater reliability) & & & .02 & .18 & -.04 & .07 & 52.15* \\
Self-rated & 1,537 & 8 & -.01 & .12 & -.07 & .05 & 26.98* \\
Performance in training programs & & & & & & & \\
Rated by supervisor or objective measures & 9,228 & 16 & -.04 & .21 & -.07 & -.02 & 277.10* \\
OCB & & & & & & & \\
General (undifferentiated targets) & & & & & & & \\
Rated by supervisors, peers, or others & 5,404 & 18 & .06 & .09 & .04 & .09 & 56.53* \\
(Corrected for interrater reliability) & & & .08 & .13 & .05 & .11 & 96.64* \\
Self-rated & 5,755 & 23 & .08 & .15 & .05 & .11 & 144.46* \\
Directed at others & & & & & & & \\
Rated by supervisors, peers, or others & 10,565 & 42 & .05 & .08 & .03 & .07 & 112.84* \\
(Corrected for interrater reliability) & & & .06 & .12 & .04 & .08 & 191.28* \\
Self-rated & 5,728 & 24 & .07 & .08 & .04 & .10 & 56.97* \\
Directed at organization & & & & & & & \\
Rated by supervisors, peers, or others & 9,308 & 34 & .06 & .10 & .04 & .09 & 124.40* \\
(Corrected for interrater reliability) & & & .08 & .14 & .06 & .10 & 204.83* \\
Self-rated & 10,398 & 37 & .14 & .15 & .12 & .16 & 262.45* \\
Directed at tasks & & & & & & & \\
Rated by supervisors, peers, or others & 1,205 & 3 & .21 & .07 & .09 & .33 & 5.31 \\
(Corrected for interrater reliability) & & & .27 & .10 & .15 & .40 & 9.92 \\
Self-rated & 1,761 & 7 & .13 & .04 & .07 & .19 & 8.26 \\
Safety performance & & & & & & & \\
Self-rated compliance with safety rules & 612 & 5 & .10 & .07 & .02 & .17 & 8.48 \\
Objective frequency of work injuries & 4,205 & 12 & -.08 & .32 & -.12 & -.05 & 369.45* \\
Self-rated frequency of work injuries & 22,949 & 13 & -.03 & .11 & -.04 & -.01 & 133.60* \\
General counterproductive work behavior (undifferentiated targets or types) & & & & & & & \\
Rated by supervisor or peers & 1,151 & 6 & -.09 & .02 & -.17 & -.02 & 5.33 \\
(Corrected for interrater reliability) & & & -.12 & .08 & -.20 & -.05 & 10.66 \\
Self-rated & 7,072 & 28 & -.12 & .08 & -.15 & -.10 & 74.17* \\
Self-rated workplace aggression & 3,641 & 15 & -.08 & .15 & -.12 & -.05 & 85.76* \\
Self-rated on-the-job substance use & 5,182 & 14 & -.07 & .08 & -.10 & -.04 & 46.13* \\
Tardiness & & & & & & & \\
Rated by supervisor or objective measures & 1,763 & 7 & -.26 & .14 & -.32 & -.20 & 35.74* \\
(Corrected for interrater reliability) & & & -.28 & .14 & -.34 & -.22 & 35.64* \\
Self-rated & 1,657 & 7 & -.12 & .12 & -.19 & -.06 & 23.39* \\
Absence & & & & & & & \\
General absence (undifferentiated causes) & & & & & & & \\
Objective measures & 72,631 & 54 & -.26 & .12 & -.27 & -.25 & 551.81* \\
Self-rated & 8,867 & 17 & -.01 & .11 & -.03 & .02 & 105.16* \\
Sickness absence & & & & & & & \\
Objective measures & 44,465 & 18 & .02 & .12 & .01 & .03 & 346.79* \\
Self-rated & 9,830 & 8 & .04 & .07 & .02 & .06 & 40.17* \\
Nonsickness-related absence & & & & & & & \\
Objective measures & 2,508 & 12 & -.10 & .20 & -.14 & -.06 & 102.03* \\
Self-rated & 3,024 & 6 & -.01 & .11 & -.05 & .04 & 29.22* \\
\hline
\end{tabular}

Note. \(N\) = cumulative sample size; \(k\) = number of studies cumulated; \(r_c\) = sample-size weighted corrected correlation; \(SD_c\) = standard deviation of \(r_c\); 95\% LCI = lower bound of confidence interval; 95\% UCI = upper bound of confidence interval; \(Q\) = \(Q\) statistic; OCB = organizational citizenship behavior.

\(^* p < .05.\)
the type of absence measure utilized. When general measures of absence were used (i.e., absence measures that did not differentiate among causes for absence), age was negatively related to objective measures at −.26. However, the effect size of age on general absenteeism by self-ratings was only −.01. A similar pattern of results emerged for nonsickness-caused absence, that is, age was negatively related to objective measures at −.10 but to self-ratings only at −.01. In contrast, age demonstrated a very weak but positive relationship with sickness absence. The effect size in the case of objective measures was only .02, and the effect size for self-ratings was only .04.

Subgroup Moderator Analyses

As evidenced by significant Q statistics (Hedges & Olkin, 1985), there is considerable variability in the relationships we examined in this meta-analysis. Table 2 presents subgroup analyses of the age–performance relationships based on two moderators—longitudinal versus cross-section design and year of study published. It should be noted that some subgroups contain fewer than three original studies \( (k < 3) \), and for the sake of parsimony, those results are not presented in Table 2.

Longitudinal versus cross-sectional design. In our pool of studies, longitudinal studies adopted different lagged time frames, ranging from 2 months to 5 years (mean time period = 11 months). Perhaps because of the limited number of longitudinal studies available and the rather short time frame adopted in these few studies, we found that many of the longitudinal age–performance relationships reported in Table 2 were not statistically significant. Another possible reason for this pattern of results is that the mean age of the samples was only 35; there were relatively few “older workers” included even in these longitudinal studies.

However, we did find that age was more closely related to older workers’ contextual performance in longitudinal studies. For example, age was significantly related to OCB directed at others (ratings by others) at .10, to self-rated OCB directed at organization at .22, and to the objective measure of general absence at −.23. These results suggest that age in and of itself does not necessarily result in performance declines. Instead, as individuals age, they may change the performance areas to which they contribute more fully.

Furthermore, with the exception of three cases, confidence intervals associated with longitudinal studies and those with cross-sectional studies overlapped, suggesting that these two types of research designs generate largely similar effect sizes. The caveat here again is that these two types of designs address different research questions, and only longitudinal studies can truly demonstrate the effects of aging per se.

The three differences found were in the areas of OCB and absence. For self-rated OCB directed at organization, cross-sectional designs revealed an effect size of .09, whereas longitudinal designs revealed an effect size of .22. For sickness absence (objective measures), cross-sectional designs revealed an effect size of −.05, whereas longitudinal designs revealed an effect size of .04. For nonsickness-related absence (objective measures), cross-sectional designs revealed an effect size of −.22, whereas longitudinal designs revealed an effect size of .05 \( (n.s) \).

Year of study published. We divided year of study published into three groups by decade of publication: before 1990, 1990–1999, and 2000 or after. This categorization allowed us to examine whether the age–performance relationship was affected by changes in the work environment over the past few decades. We found that four relationships (the relationships of age with work injuries and absence) demonstrated significant differences across time periods.

For both objective and self-rated measures of work injuries, we observed significantly stronger effect sizes of age for earlier time periods. That is, the relationship between age and injuries was found to be less negative in studies published more recently. One potential explanation for this pattern of results is that, as the technology associated with physical labor has improved and the awareness of occupational health has heightened over the years, there may be fewer differences in injury rates across age groups today than there were 20 years ago.

We also found that the negative relationship between age and nonsickness-related absence (objective measures) has become significantly weaker over the years. On the other hand, we found that the negative relationship between age and general absence (objective measures) has become stronger after the year 2000. One possible explanation here is that the definition of what counts as absence has changed over the years. After the Family and Medical Leave Act was enacted in 1993, missed days of work that used to be counted as absences may now be counted as “leave” instead. As a result, younger workers may be better able to take advantage of leave policies (e.g., for childbirth) than older workers and, therefore, have fewer missed days attributed to “nonsickness absence” for them.

The stronger negative association between age and general absence over the years may be partially attributable to improved health of older workers. As older workers have become more health conscious and life spans have increased, the physical conditioning of the average 60-year-old today may be much better than that of the average 60-year-old 20 years ago.

Regression-Based Moderator Analyses

Table 3 presents the results of the regression-based moderator testing. Three points, in particular, are worth noting about these results.

First, regression-based moderator searches were conducted on only those subsets of studies that we used for calculating the corrected correlations presented in Table 1 and had complete sample descriptions. In these cases, we used a regression-based moderator search solely for the purpose of detecting the existence and direction of moderator effects. The explained variance of these moderators in the age–performance relationships is generally satisfactory.

Second, the moderator search focused on four relationships that had the largest number of cumulative samples \( (k) \): (a) age-core task performance (rated by supervisors), (b) age-general OCBs (rated by supervisors or others), (c) age-counterproductive work behaviors (combining self-rated general counterproductive work behaviors, workplace aggression, on-the-job substance use, and tardiness), and (d) age-general absence (objective measures).

Third, to test for moderating effects of sample characteristics, we ran separate regression models for each of the following nine sample characteristics: average age, sample age dispersion, average age job tenure, average organizational tenure, job complexity, proportion of college degree holders, proportion of managers,
proportion of women, and proportion of Caucasians. This strategy was adopted because there were relatively few studies that provided information on all nine sample characteristics. Running separate regression models enabled us to include a minimum of at least 15 studies in each moderator search.

**Relationship between age and core task performance.** With respect to the relationship between age and supervisor-rated task performance, we found that average age, age dispersion, average job tenure, average organizational tenure of sample, proportion of managers in the sample, and proportion of Caucasians were all significant moderators.

The significant result associated with the moderating role of average sample age suggests that the relationship between age and core task performance may be curvilinear (Sturman, 2003). We found that the relationships between age and core task performance for four age groups—less than 30 years old, 31–35, 36–40, and over 40 years old—were .04, .09, .06, and -.05, respectively. Thus, there appears to be an inverted-U shape relationship between age and task performance for people in different age groups. That is, the relationship between age and performance is strongest and most positive for 31–35 years old but weaker for employees under age 30 and negative for workers over age 40.

The moderating role of age dispersion is negative; when age dispersion in the sample is large, the age–task performance relationship is less positive. The moderating effects of job tenure and organizational tenure on the age–task performance relationship yielded contrasting results. The age–task performance relationship is more positive when job tenure is high but less positive when organizational tenure is high.

The moderating role of proportion of managers in the age–task performance relationship is negative in form; the age–task performance relationship is less positive for managers than it is for nonmanagers. Similarly, the proportion of Caucasians in the sample has a negative moderating effect on the age–task performance relationship. That is, the age–task performance relationship is less positive for Caucasians than for other racial groups.

**Relationship between age and OCB.** We found that the relationship between age and OCB (as rated by supervisors or others) was negatively moderated by job complexity, the proportion of college degree holders, the proportion of managers, and the proportion of Caucasians in the sample. The relationship between age and OCB was more positive for low-complexity jobs, for nondegree holders, for nonmanagers, and for non-Caucasians.

**Relationship between age and counterproductive work behavior.** Interestingly, we found that the relationship between age and counterproductive work behaviors is moderated by the average age of the sample. We divided the samples into three groups of roughly equal size and treated workers aged 40 and older as a separate group: (a) average age less than 25 years old, (b) 25–39 years old, and (c) average age of 40 years and older. For the youngest group (less than 25 years old), the average effect size is only -.01. However, the effect size of the second group (25–39 years old) is -.12, and the effect size of the third group (40 years old or above) is -.17. Even though all three age groups had negative slopes, the slopes of the older groups are steeper than the slopes of the younger groups. The form of the relationship between age and counterproductive work behavior, then, is a negative, concave downward curve.

We also found that the relationship between age and counterproductive work behavior is moderated by the age dispersion in the sample, the job tenure of the sample, and the proportion of managers in the sample. The age–counterproductive work behavior relationship is more negative when the age dispersion is greater, when job tenure is higher, and when there are more managers in the sample.

**Relationship between age and general absenteeism.** Last, we found that age dispersion is the only significant moderator of the age–general absenteeism relationship. The relationship is more negative when the age dispersion of the sample is greater.

**Discussion**

By and large, the results here support the proposition that older workers contribute effectively to noncore domains of job performance. For example, the meta-analysis suggests that older workers tend to demonstrate more citizenship behaviors and greater safety-related behavior. At the same time, older workers appear to engage in fewer counterproductive work behaviors in general and exhibit less workplace aggression, on-the-job substance use, tardiness, and voluntary absence in particular. These observations suggest either that older workers are as motivated as younger workers to contribute to their organizations or that they more consciously engage in discretionary behaviors to compensate for any losses in technical core performance.

Indeed, many empirical studies of older individuals in cognitive and social psychology support these meta-analysis results. For instance, researchers have observed that older adults are more motivated to volunteer in general (Okun, Barr, & Herzog, 1998). Furthermore, in a series of studies (Barefoot, Beckham, Haney, Siegler, & Lipkus, 1993; Barefoot et al., 1991), researchers have observed that individuals over age 60 exhibit much less hostility and anger than younger adults. Other researchers have found that older adults are better able to regulate their moods and display higher levels of emotional intelligence as well (Chapman & Hayslip, 2006; Gross, Carstensen, Tsai, Skorpen, & Hsu, 1997; Siu, Spector, Cooper, & Donald, 2001). Taken together, the pattern of findings suggests that older employees are good citizens, are more likely to control their emotions at work, and are less likely to engage in counterproductive behaviors. The stereotype of older workers as difficult colleagues, then, seems largely unfounded. In contrast, age is only minimally related to job behaviors revolving around core task performance and creativity and is slightly negatively related to performance in training programs.

The null relationship between age and core task performance is similar to what two previous reviews have found (McEvoy & Cascio, 1989; Sturman, 2003). One possible explanation is that the demands of core job performance create a "strong" situation in which individuals are motivated to conform to performance expectations whatever their individual differences may be. Another possibility is that poor performers, regardless of age, are weeded out of their jobs over time so that there are no major differences in task performance and creativity among those employees who remain (Kanfer, Crosby, & Brandt, 1988).

The negative relationship between age and training performance might be partially explained by the fact that older adults are often less confident in their abilities to learn new material and retain new skills (Touron & Herzog, 2004). Moreover, this meta-analysis
Table 2
Longitudinal Versus Cross-Sectional Design and Year of Study Published as Moderators

<table>
<thead>
<tr>
<th>Variable</th>
<th>( N )</th>
<th>( k )</th>
<th>( r_c )</th>
<th>( SD_c )</th>
<th>95% LCI</th>
<th>95% UCI</th>
<th>( Q )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core task performance (supervisor-rated)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal</td>
<td>5,169</td>
<td>14</td>
<td>.00</td>
<td>.09</td>
<td>−.07</td>
<td>.07</td>
<td>15.62</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>46,879</td>
<td>104</td>
<td>.03</td>
<td>.10</td>
<td>.02</td>
<td>.03</td>
<td>521.06</td>
</tr>
<tr>
<td>Before 1990</td>
<td>3,334</td>
<td>15</td>
<td>.04</td>
<td>.10</td>
<td>.01</td>
<td>.08</td>
<td>42.94</td>
</tr>
<tr>
<td>1990–1999</td>
<td>11,563</td>
<td>28</td>
<td>.03</td>
<td>.13</td>
<td>.01</td>
<td>.05</td>
<td>174.20</td>
</tr>
<tr>
<td>2000 or after</td>
<td>37,151</td>
<td>75</td>
<td>.02</td>
<td>.09</td>
<td>.01</td>
<td>.03</td>
<td>358.40</td>
</tr>
<tr>
<td>Core task performance (objective measures)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal</td>
<td>822</td>
<td>3</td>
<td>.07</td>
<td>.17</td>
<td>−.08</td>
<td>.22</td>
<td>15.53</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>16,684</td>
<td>46</td>
<td>.07</td>
<td>.11</td>
<td>.05</td>
<td>.08</td>
<td>242.90</td>
</tr>
<tr>
<td>Before 1990</td>
<td>719</td>
<td>4</td>
<td>.08</td>
<td>.24</td>
<td>−.04</td>
<td>.20</td>
<td>28.70</td>
</tr>
<tr>
<td>1990–1999</td>
<td>2,084</td>
<td>9</td>
<td>.02</td>
<td>.13</td>
<td>−.08</td>
<td>.05</td>
<td>71.57</td>
</tr>
<tr>
<td>2000 or after</td>
<td>9,935</td>
<td>15</td>
<td>.03</td>
<td>.14</td>
<td>.01</td>
<td>.06</td>
<td>116.94</td>
</tr>
<tr>
<td>Core task performance (self-rated)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal</td>
<td>1,123</td>
<td>6</td>
<td>.01</td>
<td>.04</td>
<td>−.06</td>
<td>.09</td>
<td>6.62</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>16,684</td>
<td>46</td>
<td>.07</td>
<td>.11</td>
<td>.05</td>
<td>.08</td>
<td>242.90</td>
</tr>
<tr>
<td>Before 1990</td>
<td>719</td>
<td>4</td>
<td>.08</td>
<td>.24</td>
<td>−.04</td>
<td>.20</td>
<td>28.70</td>
</tr>
<tr>
<td>1990–1999</td>
<td>2,084</td>
<td>9</td>
<td>.02</td>
<td>.13</td>
<td>−.08</td>
<td>.05</td>
<td>71.57</td>
</tr>
<tr>
<td>2000 or after</td>
<td>9,935</td>
<td>15</td>
<td>.03</td>
<td>.14</td>
<td>.01</td>
<td>.06</td>
<td>116.94</td>
</tr>
<tr>
<td>Performance in training programs (rated by supervisor or objective measures)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal</td>
<td>1,291</td>
<td>7</td>
<td>−.05</td>
<td>.34</td>
<td>−.12</td>
<td>.02</td>
<td>103.79</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>7,937</td>
<td>9</td>
<td>−.04</td>
<td>.19</td>
<td>−.07</td>
<td>−.02</td>
<td>173.20</td>
</tr>
<tr>
<td>OCB directed at others (rated by others)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal</td>
<td>1,139</td>
<td>5</td>
<td>.10</td>
<td>.10</td>
<td>.02</td>
<td>.18</td>
<td>11.94</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>9,426</td>
<td>30</td>
<td>.07</td>
<td>.11</td>
<td>.05</td>
<td>.08</td>
<td>55.48</td>
</tr>
<tr>
<td>Before 1990</td>
<td>979</td>
<td>4</td>
<td>.06</td>
<td>.19</td>
<td>−.04</td>
<td>.13</td>
<td>25.59</td>
</tr>
<tr>
<td>1990–1999</td>
<td>1,700</td>
<td>5</td>
<td>.01</td>
<td>.06</td>
<td>−.06</td>
<td>.08</td>
<td>7.84</td>
</tr>
<tr>
<td>2000 or after</td>
<td>4,028</td>
<td>19</td>
<td>.10</td>
<td>.08</td>
<td>−.06</td>
<td>.13</td>
<td>40.46</td>
</tr>
<tr>
<td>OCB directed at organization (rated by others)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal</td>
<td>1,118</td>
<td>4</td>
<td>.03</td>
<td>.00</td>
<td>−.06</td>
<td>.13</td>
<td>2.47</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>8,127</td>
<td>30</td>
<td>.07</td>
<td>.11</td>
<td>.05</td>
<td>.09</td>
<td>120.58</td>
</tr>
<tr>
<td>Before 1990</td>
<td>2,648</td>
<td>8</td>
<td>.03</td>
<td>.11</td>
<td>−.02</td>
<td>.08</td>
<td>35.27</td>
</tr>
<tr>
<td>1990–1999</td>
<td>6,071</td>
<td>24</td>
<td>.07</td>
<td>.09</td>
<td>−.04</td>
<td>.09</td>
<td>72.72</td>
</tr>
<tr>
<td>OCB directed at tasks (self-rated)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal</td>
<td>943</td>
<td>3</td>
<td>.14</td>
<td>.11</td>
<td>−.02</td>
<td>.29</td>
<td>8.04</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>4,028</td>
<td>19</td>
<td>.12</td>
<td>.00</td>
<td>−.02</td>
<td>.23</td>
<td>0.15</td>
</tr>
<tr>
<td>Safety performance (objective frequency of work injuries)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal</td>
<td>1,233</td>
<td>6</td>
<td>−.37</td>
<td>.44</td>
<td>−.44</td>
<td>−.29</td>
<td>174.69</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>788</td>
<td>3</td>
<td>.10</td>
<td>.07</td>
<td>−.05</td>
<td>.25</td>
<td>4.20</td>
</tr>
<tr>
<td>Before 1990</td>
<td>2,184</td>
<td>3</td>
<td>.01</td>
<td>.19</td>
<td>−.08</td>
<td>.11</td>
<td>46.14</td>
</tr>
<tr>
<td>1990–1999</td>
<td>3,002</td>
<td>4</td>
<td>−.10</td>
<td>.10</td>
<td>−.16</td>
<td>−.05</td>
<td>33.49</td>
</tr>
<tr>
<td>2000 or after</td>
<td>19,947</td>
<td>9</td>
<td>−.02</td>
<td>.09</td>
<td>−.03</td>
<td>−.01</td>
<td>80.62</td>
</tr>
<tr>
<td>General counterproductive work behavior (self-rated)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 1990</td>
<td>1,315</td>
<td>4</td>
<td>−.13</td>
<td>.08</td>
<td>−.22</td>
<td>−.04</td>
<td>8.60</td>
</tr>
<tr>
<td>1990–1999</td>
<td>736</td>
<td>5</td>
<td>−.15</td>
<td>.09</td>
<td>−.26</td>
<td>−.05</td>
<td>7.76</td>
</tr>
<tr>
<td>2000 or after</td>
<td>5,021</td>
<td>19</td>
<td>−.12</td>
<td>.09</td>
<td>−.15</td>
<td>−.09</td>
<td>56.91</td>
</tr>
<tr>
<td>Self-rated on-the-job substance use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before 1990</td>
<td>1,315</td>
<td>4</td>
<td>−.09</td>
<td>.03</td>
<td>−.18</td>
<td>−.01</td>
<td>3.73</td>
</tr>
<tr>
<td>2000 or after</td>
<td>2,797</td>
<td>9</td>
<td>−.02</td>
<td>.09</td>
<td>−.06</td>
<td>.03</td>
<td>28.04</td>
</tr>
<tr>
<td>General absence (objective measures)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longitudinal</td>
<td>2,121</td>
<td>9</td>
<td>−.23</td>
<td>.20</td>
<td>−.28</td>
<td>−.18</td>
<td>79.30</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>70,510</td>
<td>45</td>
<td>−.26</td>
<td>.12</td>
<td>−.27</td>
<td>−.25</td>
<td>470.81</td>
</tr>
<tr>
<td>Before 1990</td>
<td>4,451</td>
<td>33</td>
<td>−.20</td>
<td>.22</td>
<td>−.23</td>
<td>−.16</td>
<td>225.46</td>
</tr>
<tr>
<td>1990–1999</td>
<td>13,440</td>
<td>11</td>
<td>−.18</td>
<td>.16</td>
<td>−.20</td>
<td>−.16</td>
<td>89.97</td>
</tr>
<tr>
<td>2000 or after</td>
<td>54,740</td>
<td>10</td>
<td>−.28</td>
<td>.14</td>
<td>−.29</td>
<td>−.27</td>
<td>111.38</td>
</tr>
</tbody>
</table>
only included studies in which the measures of training performance were posttraining scores. Consequently, the results may be different when pretest scores are taken into consideration. Also, because a large proportion of the training performance studies were conducted on technology training, the age–training performance relationship may be confounded by different attitudes toward new technology held by “Baby Boomers” and the “digital generation” (Morris & Venkatesh, 2000).

**Limitations of Current Research and Implications for Research Design**

The meta-analytical results presented highlight the need for new approaches to studying age–performance relationships in the future. From researchers’ initial design decisions through their conclusions about their findings, we identify seven specific issues that need to be addressed before robust conclusions about the age–performance relationship can be drawn. They are as follows: (1) the use of longitudinal designs, (2) the incorporation and measurement of mediating mechanisms, (3) the operationalization of age, (4) the operationalizations of performance criteria for older workers, (5) the use of multiple sources and types of performance measurements, (6) the interpretation of findings in light of sample characteristics, and (7) the use of meta-analysis.

**Use of longitudinal designs.** In the present meta-analysis, we focused on cross-sectional studies. This strategy allowed us to get an overall picture of the strength of various age–performance links across studies with different sample and data collection characteristics. This strategy, although also dictated in part by practical limitations of data availability (only 12% of extant studies in this area are longitudinal in nature), has several drawbacks. Unlike other demographic variables, such as gender and race, age is unique in that it steadily changes over time. Put another way, all individuals, given reasonable health, will eventually move into the category of “older workers.” Sociologists have long recognized that to fully understand the effect of aging, it is important to use longitudinal designs that track individuals over time (Baltes et al., 1971; Schaie & Hofer, 2001). That is, only longitudinal designs truly allow us to study the intraindividual process of aging and its effect on job performance.

An additional advantage of using longitudinal designs to examine age–performance relationships would be their ability to appraise the effectiveness of planned interventions (e.g., changes in company policy on medical benefits) or the consequences of external events. In the case of planned interventions, longitudinal research is important because the effects of the intervention may not be immediate in nature. For example, the effects of training interventions with older workers may not be dramatic in the short run but may be more substantive over time as older workers integrate the new material into their behavior repertoires. As another example, results obtained before the passage of the ADEA amendment in 1991 might be different from those obtained after 1991. This amendment prohibited mandatory retirement age and therefore might have increased the upper bound for age in samples. Likewise, the passage of the Family and Medical Leave Act in 1993 may have changed the ways organizations counted days missed from work as absence or leave.

However, in perhaps no area of organizational research is the legitimate barrier to longitudinal research greater than it is in the area of aging. If we truly want to study intraindividual aging, we have to collect data over several years—and the pressures of academic publishing make that task both daunting and formidable. In the small pool of longitudinal empirical studies we identified, the low-end of the time frame range was only 2 months, hardly long enough to identify meaningful aging effects anyway. Even the high end of the time range in our set of studies (5 years) falls short relative to time frames used in sociological and gerontological research (Baltes et al., 1971). Thus, we believe that generic calls for more longitudinal research are not likely to influence researchers’ design decisions in the case of age research and that more nuanced judgment calls are needed.

One possibility we suggest is gathering data longitudinally across important transition periods in individuals’ careers. For example, although it may not be realistic to plan on a 20-year study
of aging, it is more feasible to study aging in 5-year periods corresponding to three important career stage transitions (Super, 1980): between the exploration and the establishment stages (typically surrounding age 30), between the establishment and career maintenance stages (typically surrounding age 40), and between the career maintenance and career decline stages (typically surrounding age 50).

Also as noted earlier, the careers literature has largely focused on the early career participants; the mean age of the samples in the studies that have examined age–performance relationships is only 35. Probably the most heated debate about who an “older worker” is has centered on the difference between “middle age” and “old age.” Consequently, studying individuals as they make the transition from their late 30s to their early 40s or as they transition from their late 40s to their early 50s would especially help us better understand where and when major effects of aging take place.

Thus, even in cross-sectional studies, we need to focus more on differences between individuals in their 40s and 50s (typically the last 2 decades of full-time employment) as they make the transition from being “middle-age workers” to “older workers” and less on the differences between early career individuals in their 20s and 30s.

In short, although life-long studies of aging are highly desirable and influential in our understanding of developmental issues (cf. Valliant, 1977), moderate-length longitudinal studies—those that look at aging across critical transition periods, and those that use samples with higher mean ages—would be especially valuable at this point in the research stream.

**Mediating mechanisms.** In our review of the literature, we found that very limited attention has been given to the articulation of mediating mechanisms and the measurement of mediating variables in the age–performance relationship. Unlike gender, which is virtually immutable, the process of aging can bring on a number of physical, cognitive, and emotional changes that may help explain more concretely why age is related to job performance. For instance, aging may cause changes in self-efficacy (Artistroc et al., 2003), which may in turn affect the productivity of workers. As another example, changes in the frequency and intensity of emotional expression that accompany aging (Gross et al., 1997) may also affect the way coworkers and supervisors interact with older workers. Thus, conceptualizing and measuring mediating processes may be one of the most effective ways to help researchers explain why age matters to job performance, not only that age matters to job performance.

To this end, the longitudinal designs discussed above are particularly helpful because they can demonstrate whether intrapersonal processes of aging cause changes in levels of the mediating variables, which in turn may affect job performance. Equally importantly, researchers need to collect data on proposed mediating variables at multiple points in time as well. If the argument is being made, for example, that losses in self-efficacy are the major reason why performance declines with age, multiple measurements of the presumed mediating variable (self-efficacy) need to be gathered too. Fortunately, advances in structural equation modeling and the availability of user-friendly software in recent years have made statistically testing these mediating processes much easier.

**Operationalization of age.** In our review of the literature, we found that age is virtually always operationalized as chronological age. There have certainly been exceptions to that observation (Cleveland & Shore, 1992; Shore et al., 2003), but by and large chronological age is the dominant operationalization in this literature.

However, it is important that researchers start considering alternative measures of age as well. One strategy would be to gather information on workers’ “subjective age” (e.g., how old they feel, whether they see themselves as older, the same age as, or younger than their chronological age). On the positive side, because subjective measures of age are perceptual in nature, they may be particularly useful for explaining how biological aging contributes to changes in attitudes and perceptions that precede changes in job performance. On the other hand, self-report measures are subject to numerous measurement problems like social desirability response bias.

Another alternative that shows promise and that has been gaining increased attention from researchers is the use of relative (relational) age, that is, age relative to other people in the same work environment (Shore et al., 2003). Moreover, unlike perceptual measures of age, relative age can often be collected archivaly,
and therefore researchers can avoid some of the pitfalls of self-report measures. In fact, we gathered some indirect evidence that relational age does influence the age–performance link. For instance, we found that, in studies with age-diverse environments, older workers were less likely to be rated highly on core task performance.

At the same time, the meta-analytic results here indicate that older workers in age-diverse environments are less likely to exhibit counterproductive work behaviors and to be absent from work. Although these results certainly reflect the beneficial effects of age on job performance, they are also at odds with the relational demography literature, which suggests that being different in age from the dominant group often results in less positive job behaviors (Lawrence, 1996; Tsui & O’Reilly, 1989). It may be the case here that in age-diverse environments, older workers may leave more core task activities to be performed by younger workers, whereas older workers concentrate on mentoring junior colleagues, managing external demands on the work group, and facilitating the production (service) process in indirect ways.

These results on age-diverse environments highlight once again that age does not exist completely independently of the broader social context. Although it is critical to collect data on chronological age, we also encourage researchers to gather at least one additional measure of age in their studies for potential use as an independent, control, or moderator variable.

**Operationalization of job performance.** We also urge researchers to think more broadly about how to operationalize job performance for older workers. One of the key patterns of results in the current study is that older workers are especially likely to contribute effectively to noncore activities at work. In these “weak situations” in which performance demands or role expectations are not strong, the effect of aging at work may be particularly observable.

For example, we found that the age–counterproductive work behavior relationship was curvilinear. This relationship is a negative, concave downward curve (~.01 for those less than 25 years old, ~.12 for those 25–39 years old, and ~.17 for those 40 years old or above). That is, older individuals—and particularly those over age 40—are less likely to engage in counterproductive work behaviors. If age is a reliable indicator of self-control and self-discipline, as numerous researchers have argued (Chapman & Hayslip, 2006; Gross et al., 1997; Siu et al., 2001), then it appears that older workers are good self-monitors of their emotions and social behaviors. Researchers’ operationalizations of job performance, thus, should also reflect the broader range and the greater time investments older employees put into extrarole activities.

Similarly, we found that, in samples with a high percentage of managers, there were less positive relationships of age to both core task performance and OCB. One possible explanation is that it is harder to differentiate OCB from core task performance for managers than it is for nonmanagers. For example, a typical behavior included as OCB is promoting the organization to outsiders. In the case of older workers who are managers, it is unclear whether this behavior should be counted as part of the formal leadership role, as an OCB, or both (Martell & DeSmet, 2001; Pearce, Stevenson, & Perry, 1985). Thus, even seemingly straightforward distinctions between core task performance and OCB may blur substantially for older members of the workforce.

**Use of multiple sources and types of measurement.** As Table 1 suggests, there is a great deal of variability in the strength of the age–performance relationship depending on the source of performance measurement and type of performance being examined. Thus, adopting multiple sources and types of performance measures is critical to reaching more definitive conclusions about age–performance relationships. Again, it is not possible for researchers to have multiple sources and multiple types of measurement for each and every variable in a study, but having at least two sources for performance measures is especially important in this research stream.

At the minimum, researchers need to clearly state their rationales for using various sources and types of performance measures. Age is a surface-level characteristic that readily evokes others’ stereotypes and prejudices (Harrison, Price, Gavin, & Florey, 1998), especially when compared with deep-level characteristics like personality traits or values. Thus, exclusive reliance on supervisor ratings as measures of performance may lead to very different conclusions than reliance on archival measures (Levy & Williams, 2004; Waldman & Avolio, 1986).

For instance, in this study, we found that the relationship between age and supervisor-rated task performance followed an inverted-U shape. Age is positively related to task performance for the age groups younger than 40 years old but negatively related to task performance for age groups over age 40. Given it is very unlikely that any major changes in cognitive ability would occur at age 40 (Greller & Simpson, 1999), the curvilinear relationship observed may largely reflect age bias against older adults often held by supervisors (Ferris, Yates, Gilmore, & Rowland, 1985). Comparing supervisor ratings with archival data or objective performance records would be particularly useful for identifying any positive or negative bias in supervisors’ evaluation of older workers’ performance.

**Interpreting findings in light of sample characteristics.** We also encourage researchers to statistically control for the effects of sample characteristics in their analyses and to interpret findings (in their own studies or those in the literature) about the relationship of age with performance in light of those sample characteristics. This recommendation is based on our observation that the strength of the age–performance relationship varies across samples’ sociodemographic characteristics (e.g., race, education, job level, job complexity, job tenure, and organization tenure).

For instance, we found that, in samples with fewer Caucasians, the relationships of age to task performance and to OCB were more positive. Older workers from racial minorities may feel they have to work harder and display more citizenship behavior because of their “double” minority status at work (Jang, Borenstein, Chiriboga, & Mortimer, 2005).

We also found that, in samples with a high proportion of college graduates, there was a less positive relationship between age and OCB. One possibility is that college graduates view their main responsibility as performing core activities and applying expertise learned in school to assigned tasks (Motowidlo & Van Scotter, 1994). Therefore, older workers with college degrees may be less likely to spend time on noncore task activities and be less likely to feel it is their responsibility to do so. The results from the job complexity moderator analyses present a similar picture. We found that the age–OCB relationship was less positive for occupations with relatively high job complexity. There are at least two potential
explanations here. First, older employees working in high-complexity jobs may have very little time left available for performing OCBs. Second, when faced with low complexity jobs, older workers may seek out more opportunities to prove their value to supervisors.

The variability in results is also evident in samples with different mixes of managers and tenure levels. In samples with a higher percentage of managers, the relationship between age and counterproductive behavior was more negative. Here, the causality could work either way. Older workers who engage in counterproductive behavior are more likely to be terminated over time, whereas older workers who do not engage in counterproductive behavior are more likely to be promoted to manager roles (Conway, 1999).

Along the same lines, samples with higher average job tenure yielded more positive age–task performance relationships and more negative age–counterproductive work behavior relationships. Taken in the context of the attraction–selection–attrition paradigm, these findings make considerable sense because older workers are more likely to stay at jobs in which they think they can excel, which they enjoy, or both (Kanfer et al., 1988; S. R. Rhodes, 1983). On the other hand, we found studies that used samples with low average organizational tenure revealed a more positive relationship between age and core task performance. This perhaps indicates that older newcomers feel more pressure to perform highly right from the start and to “hit the ground running” (Feldman & Brett, 1983). Collectively, the above moderator variable results provide a cautionary tale for researchers whose findings may be inadvertently distorted by the characteristics of samples from which they collect data.

Use of meta-analysis. In our review, we found that several meta-analyses had been performed to address different research questions related to aging. Although meta-analyses provide a robust picture of cumulated research findings, they have some limitations too. In our meta-analyses of age–performance relationships, four limitations in particular were noticeable.

First, because of unreported data, we were not able to search for moderators for all age–performance relationships of interest here (e.g., age–creativity). A second limitation is the relatively small number of aggregated studies for some of the relationships investigated, regardless of a comprehensive search for such studies. Even though meta-analysis can be executed with as few as two studies (Hunter & Schmidt, 1990), the cumulated effect sizes are more stable when the number of cumulative studies increases.

Third, because of the data contained in individual studies, we had to use proxy variables in our moderator search, such as the percentage of women for the effect of gender, and the year of study published for the effect of time or changing nature of work. Finally, our analyses were predominantly based on published studies. Therefore, the “file-drawer” problem, though alleviated, was not entirely eliminated.

In sum, then, the quality of the meta-analysis is constrained by the quality of the empirical studies on which it is based. As the prior research design issues get resolved, the inherent restrictions imposed by meta-analyses will be lessened as well.

Implications for Practice

Consistent with our earlier discussion of performance evaluations, we hope the present study will help managers rethink their approaches to performance appraisal. As Welbourne et al. (1998) have emphasized, performance appraisals should be largely based on the social roles that employees play in organizations. Thus, older employees’ performance in contextual activities may need to be weighed as or more heavily than their performance in core task activities. Moreover, the social context of the performance appraisal system may need to be altered to reduce age bias in the evaluations themselves (Levy & Williams, 2004; Perry & Finkelstein, 1999). Given the variance among self-evaluations, peer evaluations, and supervisor evaluations, older workers may especially benefit from the use of 360° feedback systems.

Another potential implication of this study is the need to rethink the mission and format of organizational training for older workers. In terms of training content, for instance, perhaps organizations should spend less time trying to teach older workers technical skills (e.g., computer training) on which they will be unlikely to compete effectively with their younger colleagues. Instead, it might prove more useful to provide older employees with greater supervisory and mentoring training so that they can do a better job of facilitating younger employees’ core performance. Similarly, although intense week-long training programs may be an efficient mechanism for training younger employees, training programs that are slower paced and spaced out over time might be more effective for older workers.

Managers today are well aware of the ADEA, which prohibits discrimination in employment against people of age 40 or more. Although avoiding discriminating against older workers is a widely accepted human resources (HR) strategy, our findings suggest that perhaps an even smarter HR strategy is to realize older workers’ potential contribution to organizational functioning and to make better use of their human capital. Furthermore, the current study also has implications for the utility of defining the cutoff age for age discrimination at 40, particularly because the workforce is aging, and 41 is the current median age of the workforce. Our moderator analyses highlight that the strength of the relationships between age and job performance varies considerably, depending on the work roles of the older employees, their job complexity, their education, their racial group, and most importantly, on different dimensions of performance. Thus, using one omnibus cutoff age to distinguish between “old” and “young” workers—and then to compare their respective job performance—seems to be far from ideal.

Conclusion

Given the dominant role that older workers will play in the labor market over the next 2 decades, it is critical that we understand how to shape work environments to take advantage of these cohorts’ talents and to minimize the challenges they face on the job. We hope this article provides a useful foundation upon which more theoretically rich and more methodologically diverse research can be built and through which even better HR practices targeted at the aging workforce can be developed.

References


AGE AND JOB PERFORMANCE

... theoretical and empirical literature and suggestions for future research. *Journal of Management*, 26, 513–563.

(Appendix follows)
Appendix

Empirical Studies Included in the Meta-Analysis


(Appendix continues)


(Appendix continues)


(Appendix continues)


(Appendix continues)


Tremblay, M., & Roger, A. (1993). Individual, familial, and

(Appendix continues)


Received March 7, 2007
Revision received September 3, 2007
Accepted November 13, 2007

Members of Underrepresented Groups: Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write to the address below. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.

- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.

- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In the letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, “social psychology” is not sufficient—you would need to specify “social cognition” or “attitude change” as well.

- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

Write to Journals Office, American Psychological Association, 750 First Street, NE, Washington, DC 20002-4242.