*Work, Aging and Retirement*, 2022, **8**, 208–223
https://doi.org/10.1093/workar/waab039
Advance Access publication 26 April 2022
**Review Article**

OXFORD

Review Article

# Age, Experience, and Business Performance: A Meta-Analysis of Work Unit-Level Effects

**Richard A. Guzzo[1], Haig R. Nalbantian[2], and Nicholas L. Anderson[3]**

[1]Workforce Sciences Institute, Mercer, Washington, DC, USA
[2]Workforce Sciences Institute, Mercer, New York, NY, USA
[3]Mercer, New York, NY, USA

Correspondence concerning this article should be addressed to Richard A. Guzzo, Workforce Sciences Institute, Mercer, 1150 Connecticut Avenue NW, Washington, DC 20036, USA. E-mail: rick.guzzo@mercer.com

*Decision Editor: Donald M. Truxillo, PhD*

## Abstract

Adopting an interdisciplinary perspective, this article reports new evidence on the impact of age and experience on work unit performance. Two types of experience that increase with age are "general" and "firm-specific." The focus here is on the influence of general human capital (which increases with time spent in the workforce) and firm-specific human capital (which increases with tenure with the current employer) on work unit performance. Although age–performance relationships have been investigated extensively in two research literatures, psychology and economics, neither addresses such relationships at the unit-within-organization level of analysis, concentrating instead on age–performance relationships at individual, organizational, or national levels. Using a unique data set comprised of large-sample, long-duration, multivariate studies of unit performance within firms this meta-analysis synthesizes partial effect sizes for the effects of age and tenure. A key finding is that tenure positively affects unit performance whereas age has no effect. Work unit leaders' tenure but not age was found to positively affect unit performance. The lack of evidence of an age–performance relationship is consistent with psychological research at the individual level but contravenes economics research literature which, at all levels of analysis, generally reports negative relationships between age and performance. Neither the heterogeneity of tenure nor age was related to performance nor was there evidence of nonlinearities in relationships. Practical implications of the findings are discussed regarding ageism and employers' use of gig or contract workers. Implications for future research and theory focus on interdisciplinary theory development and the scientific contribution of organizationally based research.

The relationship between age and work performance has been extensively investigated in two empirical literatures, psychology and economics, with little convergence. Psychological research findings primarily concern age–performance relationships at the individual level of analysis whereas economics research addresses individual as well as other levels of analysis, such as organizational and national. Further, the two literatures' patterns of findings mostly diverge. Negative relationships between age and performance – manifest in an inverted U-Shape – predominate at all levels of analysis in economics research while in psychological research relatively weak age–performance relationships are reported at the individual level, with certain exceptions related to type of task and aspect of performance. One thing common to the two literatures is the absence of evidence about age–performance relationships at the level of analysis of *work unit within organizations*, the focus of the present study. This study adopts a human capital theoretical perspective and applies it to unit members as well as unit leaders in assessing the impact of age-based experience on performance, thus contributing new findings to two empirical literatures. In addition to discussing the findings' implications for future research and practice, the article also addresses two broader issues: How best to create more integrative interdisciplinary theory and ways of enhancing the scientific contributions of organizationally based big-data research.

## Work units as a level of analysis

Work units, as defined in this research, are well-bounded, naturally occurring entities created by organizations to get work done. All such units are accountable for producing outcomes and have measurable results that reflect the collective performance of their members. Schneider and Pulakos (in press) define work units the same way and emphasize their distinctiveness from other social entities in workplaces (sometimes called groups or teams). Because this research analyzes in situ works units over extended periods of time, a unit can experience changes in membership due to typical workplace events (e.g., a new employee joins, an incumbent quits or transfers to another unit), throughout which the unit remains intact as an organizational entity. Examples of the units studied in this research include fast food restaurants, small stores, departments within a large store, retail bank branches, building maintenance teams, wards within a hospital, production lines within a factory, and distribution centers.

Studying units is important because of the practical reality that organizations deliver work through them. Beyond practical realities, there is a need for more psychologically based research on unit-level outcomes in the workplace to advance theory and relevance as pointed out by Schneider (2018) and Schneider and Pulakos (in press), a need exemplified very well in the age–performance research domain. Research evidence is needed to guide theory development as there is ample reason to believe that age and experience can affect unit performance in many possible ways. Members who "have seen it all" can make their units more effective by drawing on knowledge lacking in less-experienced coworkers. Conversely, unit performance may be impaired if older members struggle to meet changing job requirements (e.g., are slow to adopt new technologies) or because their presence obstructs, for reasons of social standing or deference, younger members' opportunities to contribute. The impact of age on performance can be direct (how much each employee produces) and indirect, such as when an older worker raises a younger coworker's effectiveness through coaching or when the lower turnover propensity of older workers stabilizes work units and contributes to their effectiveness. These indirect effects are sometimes called "crossover" or "spillover" effects (Nalbantian, 2014). Additionally, studying work units directly overcomes the pitfalls of the "exception fallacy" of assuming what is true at the unit level on the basis of individual-level data and the "ecological fallacy" of making inferences about work units from organizational-level data. Lastly, a particular advantage of unit-within-organization analyses is the ability to control for the influence of many contextual factors that, if left uncontrolled, may confound observed relationships in the workplace. Accounting for such contextual factors not only is important for accuracy of estimation of relationships but also for developing robust theories.

## Age and performance

### *Psychological research*

Much of the foundational psychological research on age and performance focused on individual-level physical and cognitive functions devoid of work contexts. That research, according to Charles and Carstensen (2010, p. 383), was characterized by "unidimensional decline models" emphasizing decrements that come with aging, some of which start in mid-life. More recent research, however, emphasizes both age-related decrements and gains. For example, while individuals' speed of processing new information declines with age, their stock of knowledge continues to grow (Salthouse, 2012). Likewise, more recent evidence shows that some aspects of cognitive functioning (e.g., vigilance) decline with age, some stay stable, and some (e.g., avoiding distractions) improve (Verissimo et al., 2021). Further, Charles and Carstensen (2010) find that with age comes greater sensitivity to emotional cues. Extrapolating from research findings on basic processes to workplace performance is extremely hazardous because jobs require bundles of cognitive, behavioral, and social capabilities. Consider an identical customer service job being performed in the same circumstances by a 20 year old and a 70 year old: Would we predict the older worker to be more likely to create negative customer experiences because of less vigilance and slower information processing or to be more likely to create positive customer experiences because of greater stock of knowledge and heightened sensitivity to customers' emotional cues?

Reflecting the decline models of the earlier era and more directly related to performance in the workplace, Rhodes (1983) identified a widespread belief that work performance deteriorates with age. Since then a series of cumulative meta-analyses have summarized a growing body of research reports on this topic. Waldman and Avolio's meta-analysis (1986; 13 studies, 40 samples) found that age is modestly, positively related to individual performance when performance is assessed with objective measures but not so when subjective ratings are the measures. McEvoy and Cascio (1989; 65 studies, 96 samples) concluded that age and performance are unrelated, whatever the type of performance measure. Both meta-analyses report substantial heterogeneity in effect sizes. Sturman (2003; 115 samples) found that age–performance relationships are moderated by job complexity. For low-complexity jobs evidence supported an inverted-U relationship between age and performance was detected but for high-complexity jobs greater age was associated with better performance, linearly. A comprehensive meta-analysis by Ng and Feldman (2008; 380 studies, 438 samples) concluded that age is unrelated to core task performance. Their analysis also expanded the performance construct to include dimensions such as organizational citizenship behaviors, safety behaviors, and workplace aggression and found, for example, that older workers engage in more citizenship and safety behaviors and less aggression. Each of the meta-analyses of age and individual performance reports considerable heterogeneity in effect sizes. A more recent narrative review concludes that "on average, age is a weak predictor of job performance" (Hedge & Borman, 2019, p. 137), although these authors note that beyond age 50 there may be a negative relationship between age and performance in jobs that are less cognitively demanding, thus pulling forward a thread from Sturman's (2003) work. One other stream of research with psychological roots has investigated the relationship between the age of one employee, the CEO, and performance of the entire organization. Wang et al.'s meta-analysis (2016; 308 studies, 315 samples) finds no relationship between CEO age and organizational performance. Economics research, described below, also investigates organizational performance but as a function of age of the overall workforce (e.g., average employee age).

### *Economics research*

In contrast to psychological research, economics research is organized around a small number of key theoretical propositions, analyzes age–performance relationships at multiple levels of analysis, and at the individual level often relies on pay as a proxy for performance. Much of the empirical literature is driven by the study of the "age-earnings profile." The reasoning is simple: If, in competitive labor markets, pay reflects the marginal product of labor—a mainstay proposition of neoclassical (textbook) microeconomics—then the trajectory of pay should mirror the trajectory of productivity (performance) over the employee lifecycle.

An early empirical finding—sometimes characterized in labor economics as a "stylized fact"—is that the age-earnings profile has an inverted U-shape: pay rises rapidly early in career, reaches an inflection point where its rate of growth declines, levels off later in career and ultimately declines as employees get closer to retirement. This pattern has been documented in empirical analyses of employees in many different countries, sectors, job families, education levels, and demographics, though there is considerable variance across these segments in the slopes of the curves and whether pay actually declines after a certain age. Mincer's (1958, 1974) work provides a theoretical explanation by drawing on Human Capital theory (Becker, 1964). Human capital refers to individuals' attributes that have value at work, consisting of both inherent characteristics of the individual (e.g., intelligence) and acquired attributes including such things as stock of knowledge and skill levels. Mincer theorizes that early in careers individuals invest substantially in enhancing the value of their human capital, first through schooling and subsequently at work (e.g., training, experience-based learning) because they anticipate future payoffs from this investment. As employees age, their investments in human capital decline both because the duration of time to enjoy returns is shorter and because the opportunity costs associated with new learning are higher. The reduced investment in human capital is followed by a reduction in productivity growth, pay growth and, ultimately, earnings. Mincer's proposition that expectations of payoffs decrease with age is consistent with Carstensen's (2006) work on age-related changes in time perspective and work of Zacher and Frese (2009) reporting correlations between age and perceptions of occupational opportunity.

Mincer's work spawned voluminous empirical and theoretical literatures that have become cornerstones of modern labor economics. Still, the empirical evidence has not always substantiated the theory. For example, Baker et al. (1994) find steeper declines in real wages for older workers than Mincer's empirical work suggests, and Casanova (2013) finds that real wages flatten but do not decline for older workers so long as they remain in full-time employment, suggesting that transitions to part-time employment may be driving the wage decline among older workers. Other empirical challenges are found in the work of Murphy and Welch (1990), Heckman et al. (2008), and Lemieux (2006).

Another challenge to the Mincer perspective comes from research calling into question the validity of drawing inferences from age-earnings profiles. Lazear (1979) shows why younger employees are paid less than their productivity early in career, when their performance characteristics are less well known to employers, but paid more than their productivity later in their career as their actual performance becomes visible. And Lazear and Rosen (1981) demonstrate how pay and performance become disconnected when employees "compete" in a "tournament" for "prizes" of higher pay that comes with advancement up the organizational hierarchy. As one moves up, pay differences between levels increase beyond the relative productivity differences among employees in those levels. Since occupying positions in the highest levels of organizations often comes at later rather than earlier ages, the earnings of these older workers—the tournament "winners"—become disconnected from contemporaneous performance, undercutting the validity of using pay to proxy for productivity.[1]

Some studies in economics, as is more typical in psychological research, use ratings as measures of individual performance. In such research negative age–performance relationships prevail, with some equivocality. Medoff and Abraham (1980, 1981) found evidence of either a negative or no relationship between age and performance among professional and managerial employees as measured by performance ratings. Frederiksen et al. (2017) find inconsistent results on the experience–performance rating relationship using data from six different firms. Analyzing data for nearly 500,000 employees in a dozen organizations representing multiple sectors, Nalbantian and Marciniak (in press) find a significant and sizeable negative relationship between age and the probability of employees receiving high performance ratings. This result holds no matter how "older worker" is defined—continuous age, age bands, or generation. Strikingly, the negative relationship holds for each of the 12 organizations.

Econometric studies that assess age–productivity relationships at the firm (rather than individual) level generally find that firms with higher mean workforce age or a large share of older workers are less productive than those who have a younger workforce. Skirbekk (2004) found that firm performance declines with greater employee age in 5 out of 7 studies with evidence indicating an inverted U-shape relationship between firm performance and employee age, with the age-related decrement declining most notably in firms with greater proportions of employees over 50. Findings of negative age–performance relationships at the firm level also are reported by Grund and Westergard-Neilsen (2008), Lallemand and Rycx (2009), and Von Bonsdorff et al. (2016). There are, however, exceptions. Backes-Gellner and Veen (2009) report a direct, negative association between average workforce age and firm-level productivity but further find that heterogeneity of age and type of work affect the relationship. Specifically, the age–productivity relationship was found to turn positive when work is creative and problem-solving in nature versus routine and when there is greater heterogeneity in workforce age. Børing (2019) finds a positive relationship between employee age and firm productivity as measured by sales per employee and Mahlberg et al. (2013) find no statistically significant relationship between larger shares of older workers and productivity in a sample of 19,633 firms over a 3-year period. At a higher level of analysis, among states in the United States, Maestas, Mullen, and Powell (2016) find that GDP growth rates decline with increases in both the shares of the general population and of the working population over age 60, with the decline attributable to an aging workforce greater than that attributable to an aging general population. At the country level of analysis Sharpe (2011) found workforce aging to have a statistically detectable but small adverse effect on productivity in Canada and Westelius and Liu (2016) found similar results in Japan.

In summary, evidence from economics research at the individual level indicates that performance declines with age, generally by the mid-fifties. This stands in contrast to psychological research which suggests no such clear inverse relationship between age and performance, except perhaps in the least complex jobs, and which also points to aspects of performance that increase with age, such as citizenship behaviors. At more aggregate levels of analysis (firm, state, nation) findings from economic research are a bit more mixed but mostly show negative age–performance relationships. One thing that economics and psychological research have in common, though, is a dearth of studies assessing age–performance relationships within organizations at the work unit level of analysis.

[1]See Lazear (1995) for a thorough treatment of incentive and selection challenges that emerge when the informational assumptions of traditional models of "perfectly competitive" labor markets are violated.

### Experience, age, and tenure in organizations

Some critics characterize age as “an empty variable” (Hedge & Borman, 2019, p. 124). Age, however, is empty neither in practice nor in theory. Age is foundational to many employer practices, such when age is used as a criterion for eligibility for a benefit (e.g., pension income) or a role (e.g., an overseas assignment). Many organizations actively manage sequences of assignments for early-career (younger) employees to enable them to acquire leadership capabilities expected to be expressed at a later age (e.g., McCauley et al., 2014).

Further, age is an essential construct in human capital theory because of the experience it brings. Human capital theory distinguishes between two types of age-related experience, general and firm-specific. General work experience concerns total time spent in the workforce and the resulting acquisition of knowledge, abilities, and other personal resources through that experience. Firm-specific human capital, which accrues through time spent working in one’s current employing organization, subsumes thing such as knowledge of the proprietary processes of an organization, internalization of its culture, familiarity with (and to) customers or clients, and access to information and resources by virtue of memberships in networks unique to an organization. Firm-specific human capital is carried by current employees and is broadly relevant to performance in that organization; general human capital is transportable and relevant in any organization. In empirical research, chronological age is the key indicator of general human capital and tenure with current employer the key indicator of firm-specific human capital.

Originating in economics, basic tenets of human capital theory have been adopted by other organizational sciences. Ng and Feldman (2013a) relied on it to deduce the hypothesis that performance improves with increasing time in job due to the continued accumulation of job-specific knowledge and skills and pitted it against a competing hypothesis derived from job design theory (Hackman & Oldham, 1976) that performance declines with increasing time in job due to diminishing motivation. Ng and Feldman’s (2013a) meta-analytic findings supported the latter hypothesis. Note that their research focused on one specific type of tenure, time in job, and on individual job performance. In other work, Ng and Feldman (2013b) investigated innovation-related behavior at work (e.g., generating and disseminating new ideas), again drawing on human capital theory, and found no relationship between age, job tenure, and the extent of innovative behavior. Other researchers drawing on human capital theory include Gonzalez-Mulé et al. (2019) who explored different types of aggregate tenure (e.g., “team age” defined as duration that members of an intact team are together and “additive tenure” defined as the sum of current members’ time spent in a team, recognizing that a team may have membership changes over time). Gonzalez-Mulé et al. provide evidence that different types of tenure have distinguishable effects.

Our research draws on human capital theory’s distinction between firm-specific and general human capital. The primary objective of the research is to assess the impact of unit members’ firm-specific and general human capital on unit performance using average tenure and average age among unit members as key indicators of the two types of human capital, respectively. We also investigate whether age and tenure dispersion affects unit performance. Gonzalez-Mulé et al. (2019) and Bal and Boehm (2019) provide evidence that they may, although other research points to the opposite, that age dispersion is a likely source of performance decrements (Bezrukova et al., 2009). In this latter paradigm age dispersion contributes to the emergence of identity-based subgroups within units that give rise to problems of coordination, information flow, and other performance-relevant processes. We also explore potential nonlinearities in age and tenure effects. Further, we extend human capital theory to work unit leaders and assess the effects of leaders’ firm-specific and general human capital on their units’ performance. Work unit leaders are defined here as those individuals who occupy that position in an organization’s hierarchy to which all members of a work unit report. Prior research on the impact of leader’s general versus firm-specific experience on unit performance is sparse. One such study operationalized general human capital by the simple dichotomy of whether an R&D leader was in their first job postcollege or not and investigated the impact of this difference on patent production, finding mixed results (Choudry & Haas, 2018).

A recurring challenge when investigating age-related variables is the difficulty of isolating their impact from that of other covariates. Barnes-Farrell et al. (2019) state very clearly the importance of controlling for covariates when studying age-related variables in the workplace, illustrating the point by citing the work of Clark et al. (1996) which, after controlling for 80 covariates, documented a U-shaped relationship between age and job satisfaction (viz., satisfaction declines from initial levels into the 30s and then rises into the early 60s). All of the studies on which this research report is based conform to a common research design that controls for dozens of workplace covariates of age and experience.

### Research questions

We frame the issues investigated as addressable research questions rather than as deduced hypotheses because there is, as yet, no integrated, interdisciplinary theory from which to infer specific hypotheses about expected effects at the unit level and because of the problematic nature of making cross-level inferences about the age–performance relationship at the unit level when empirical findings of the two disciplines so often conflict. The research questions we investigate are:

1. What is the impact of a work unit’s general human capital, as measured by average age, on performance?
2. What is the impact of a work unit’s firm-specific human capital, as measured by average tenure with the current employer, on performance?
3. What is the impact of work unit leaders’ general human capital on unit performance?
4. What is the impact of work unit leaders’ firm-specific human capital on unit performance?
5. Are the effects of general (age) and firm-specific (tenure) human capital equal?
6. Does within-unit disparity in age or tenure affect unit performance?
7. Are there nonlinearities in age and tenure relationships to performance?

### Overview of the data collection paradigm

Before presenting methodological details, it is useful to describe the general paradigm in which the original data for this study were generated. All data come from a body of research-based applied consulting projects—studies—conducted in business organizations. All projects followed a standard data analytic protocol referred to as "business impact modeling" (Nalbantian et al., 2003, p. 105). This protocol involves statistical modeling and analyzes proprietary employer data, which can be supplemented by additional data, to assess the impact of workforce attributes, management practices, and contextual factors on work performance. More specifically, all studies in this analysis share three key features that are core to the analytic protocol. One, statistical models draw on longitudinal data and are constructed to assess the impact of current-period factors (e.g., average age of unit members) on next-period unit performance.[2] Two, the models involve extensive numbers of control variables. Three, the impact of age and tenure on performance is measured repeatedly over relatively long intervals (weeks, months, or years). Taken together, these features support the causal interpretation of observed age and tenure effects.

From the applied consulting perspective the studies provide for organization-specific, evidence-based recommendations for changes to improve performance. A study's results can become the basis for changing how an organization staffs or rewards units to improve performance, for example, and strong causal inference increases the confidence that recommended changes will have their desired effect. From a research perspective—the interest of this article—these studies are now sufficient in number to have yielded a sizeable database of findings worthy of formal investigation. This article capitalizes on that unique collected body of findings by reporting the results of a meta-analysis of the impact of age and tenure on unit performance.

## Methods

### Sources of data and variables

#### *Performance data*

Performance measures are situationally specific and drawn from databases created and maintained by organizations as part of their normal business processes. Three types of performance measures are in the analysis: Financial, customer, and operational. Financial measures include such things as the dollar value of sales, profit, and revenue growth. Customer measures include retention rates, local market share, frequency of customer referrals, and consumer satisfaction ratings. Operational measures have to do with efficiency, speed, and quality of work processes and include measures of wasted resources, time required to fill an order, and compliance with quality standards. A unit's performance could be measured by one, two, or all three types. We reverse-coded as required when higher values of a measure corresponded to worse outcomes (e.g., longer time to fill an order). To avoid errors and biases potentially associated with subjective measures of performance, no ratings by unit leaders or members were used as dependent variables.[3] For purposes of maximizing statistical power for some analyses, we combined the three performance measures into a single metric.

[2]Two studies in this analysis were an exception and modeled same-period performance.

[3]Individual-level performance ratings were used as covariates in some studies.

#### *Unit member data*

In all cases an organization's HRIS (human resource information system) database was the source of measures of age and tenure of work unit members and of their leaders. Unit-level measures were calculated from the individual-level HRIS employee data: mean age, mean tenure, and unit-level standard deviations (*SD*s) of age and tenure. The HRIS database also was the source of data on reporting relationships that was used to identify work unit leaders.

#### *Control variables*

HRIS databases were a source of numerous control variables, such as demographics (race/ethnicity, gender), type of work performed, compensation, work status (e.g., part-time vs. full-time), most recent individual performance rating, and work location. In many cases, control variables based on work location were introduced, such as zip-code level average household income or county-level unemployment rates. Such variables came from both publicly available and proprietary sources. Other than HRIS databases provided additional sources of employee information used as control variables in several cases, such as databases containing records of training courses completed by employees. Many important performance-relevant control variables were highly specific to a unit's work and circumstances, such as measures of the case complexity in a hospital ward, the size in square feet of a department in a retail store, and the time since an establishment was last refurbished. Participating organizations provided data on these types of control variables.

### Case inclusion criteria and sample size

A search of the consulting firm's (Mercer's) library of client cases utilizing the business impact modeling paradigm yielded 40 potentially eligible cases for inclusion in the meta-analysis. Three decision rules determined which among the 40 could be included. First, at least one dependent variable (DV) had to be a financial, customer, or operational measure. Second, unit performance had to be the focus of the analysis (vs. individual performance). Third, age or tenure had to be included in an ordinary least squares (OLS) regression model as a continuous variable. Figure 1 illustrates the decision rules. The final sample for the meta-analysis consists of 21 and 23 cases in the age and tenure meta-analyses, respectively.

Projects typically involved several statistical models that differed in their specifications. For the meta-analysis, we selected only those findings from a study's final model, that which was the basis for action recommendations. There are several reasons why multiple models with differing specifications exist per case. One is that multiple models were specified a priori as means of investigating different issues, such as distinct models for issues that pertain to only some parts of an organization. Another is that multiple models arose out of an iterative, collaborative model refinement process involving consultant-researchers and client organizations. Such iterations reflected the dynamic nature of applied research. Reasons for such iterations include the emergence of new hypotheses and questions stimulated by reviews of preliminary results, the removal of variables from prior models due to revealed deficiencies, new data becoming available during the course of a project, and the convergence on the best-fit model. A caveat to reliance on the final model concerned our interest in using as many observations as possible. Specifically,

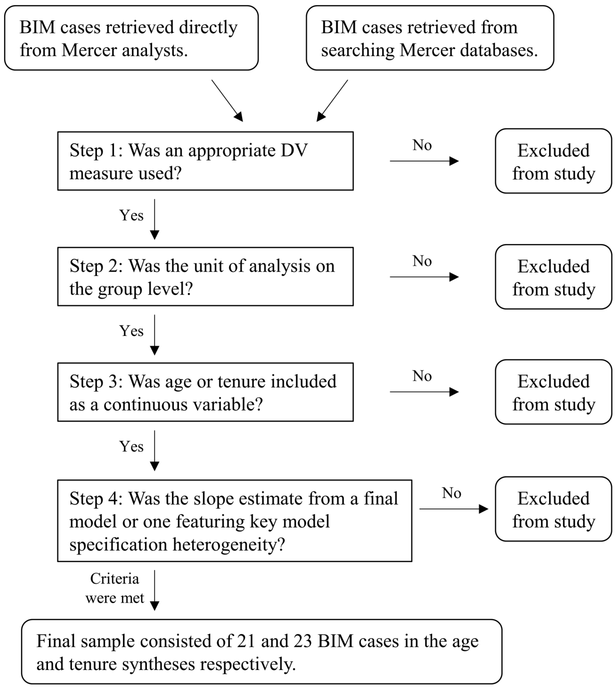


**Figure 1.** Decision rules for study inclusion.

if a case's final model did not contain the full complement of variables needed to assess relationships of interest yet the case contained the requisite data to do so, we chose to include findings from other-than final models to maximize the number of observations available. This choice resulted in the inclusion of 12 additional effect size estimates from two cases.

Because of multiple performance measures and because studies were conducted separately for distinct parts of some organizations, such as when analyses were performed separately for each of a large company's manufacturing, distribution, and service businesses, many organizations contributed more than one effect size to the analysis. This is analogous to prior meta-analyses' inclusion of results from multiple samples from the same study (e.g., 622 effect sizes from 169 studies reviewed by Gonzalez-Mulé et al., 2019).

Table 1 presents basic descriptive data about the studies in this analysis, including industry type, the number of parameter estimates each case contributed to the analysis of unit age and tenure and to the analysis of unit leaders' age and tenure. In Table 1, for example, the first case contributed two effect sizes to the age–performance relationship for unit leaders and four and two effect size estimates, respectively, to the tenure–performance analyses for unit members and unit leaders. Table 1 also reports sample size as "person years." Person-years are calculated as the number of employees working in units multiplied by the number of years units were studied. Total person-years are 1,209,237 and 1,255,479 for analyses of age and tenure, respectively. Over 1.2 million years of employee behavior in the workplace is an appreciable basis for analyzing the impact of age- and tenure-based experience on work unit performance. Table 2 presents basic demographics of the samples of employees in the separate analyses for unit age and tenure. Average unit age and tenure correlate $r$ = .79, similar in magnitude to Ng and Feldman's (2013b) reported correlation of 0.70 at the individual level.

## Coding and data validation checks

Our coding protocol followed prevailing recommendations, such as creating a coding manual, unambiguous definitions for all variables in the source studies, and discussion-to-consensus as a research team to resolve ambiguities and make low-inference judgment calls (Lipsey & Wilson, 2001; Vevea et al., 2019; Wilson, 2019). Information from source studies was coded directly into Microsoft Excel using data validation rules to safeguard data integrity. A single member of the research team double-coded all source cases. The intrarater reliability coefficients calculated using a two-way mixed effects model and the absolute agreement definition (Shrout & Fleiss, 1979; Koo & Li, 2016) resulted in a minimum intraclass correlation coefficient (ICC) across all coded variables of 0.95. As an additional fail-safe, other team members spot-checked data coding.

## Regression models and encoding effect sizes

All cases use time series data and regression models with extensive control variables. All models account for time by including a set of fixed-effects indicators for months, years, or weeks as appropriate to the case. The focus of analysis is thus unit-by-period observations (e.g., a unit's monthly performance with performance measures lagging other unit measures by 1 month). In each case, the source models follow this form in predicting unit-level performance per period:

$$Y_i = \beta_0 + \gamma_1 Age_i + \gamma_2 Tenure_i + \beta_1 X_{1i} + \cdots + \beta_p X_{pi} + \varepsilon_i \quad (1)$$

where $Y_i$ refers to a performance measure for unit-period observation $i$, $\gamma_1$ is the impact on $Y$ for a 1-year increase in mean employee age within the unit, $\gamma_2$ is the impact on $Y$ for a 1-year increase in mean employee tenure within the unit, $X_{pi}$ and the corresponding $\beta_p$ represent the impacts of covariates. The average number of covariates in models assessing age is 45 and the average number in models assessing tenure is 42. Average $R^2$ for models testing age–performance relationships is .56 and .57 for models testing tenure–performance relationships.

In this meta-analysis, we seek to compare slope estimates of the age–performance and tenure–performance parameters across cases. That is, we analyze partial effect sizes. Aloe and Thompson (2013) review three different partial effect size indices suitable for synthesis through meta-analysis: the standardized slope ($b$), partial correlation ($r_p$), and the semi-partial (or part) correlation ($r_{sp}$). The partial correlation was chosen as the effect size index for this analysis. It communicates the strength of association between two variables while holding other variables constant and it requires only the $t$-statistic and degrees of freedom ($df$) to calculate error terms ($df = n - p - 1$ where $p$ refers to the number of covariates and $n$ is the regression sample size). Squaring the partial correlation produces the proportion of (unique) unexplained variance in $Y$ attributable to an independent variable beyond all other variables in the model (Cohen & Cohen, 1975). In a simulation study, Aloe (2014) demonstrated that, in meta-analysis, the partial correlation performs well with respect to bias and root mean squared error and has smaller Type I error rates for the homogeneity test as compared to the part correlation. The partial correlation for each slope relation is (Aloe & Thompson, 2013):

**Table 1.** Studies in the meta-analysis.

| | | | Number of effect size estimates | | | |
|---|---|---|---|---|---|---|
| | | | Age | | Tenure | |
| Case ID | Industry | Person-years | Members | Leaders | Members | Leaders |
| 1 | Transportation | 17,074 | 0 | 2 | 4 | 2 |
| 2 | Food | 12,017 | 3 | 0 | 0 | 10 |
| 3 | Food | 19,500 | 2 | 0 | 2 | 3 |
| 4 | Financial services | 728 | 5 | 0 | 10 | 0 |
| 5 | Retail | 2,200 | 0 | 1 | 0 | 1 |
| 6 | Health | 20,627 | 19 | 0 | 35 | 0 |
| 7 | Health | 19,500 | 2 | 2 | 2 | 2 |
| 8 | Retail | 769,231 | 2 | 4 | 0 | 4 |
| 9 | Financial services | 30,552 | 4 | 4 | 4 | 4 |
| 10 | Services | 21,288 | 3 | 0 | 3 | 0 |
| 11 | Food | 29,450 | 2 | 2 | 1 | 1 |
| 12 | Mining | 6,687 | 4 | 0 | 4 | 0 |
| 13 | Distribution | 29,733 | 1 | 1 | 0 | 1 |
| 14 | Food | 13,210 | 10 | 0 | 10 | 0 |
| 15 | Health | 69,792 | 1 | 0 | 1 | 0 |
| 16 | Health | 22,496 | 5 | 0 | 5 | 0 |
| 17 | Education | 6,897 | 0 | 1 | 0 | 2 |
| 18 | Health | 1,755 | 2 | 2 | 2 | 2 |
| 19 | Health | 27,972 | 0 | 0 | 6 | 0 |
| 20 | Food | 18,270 | 0 | 0 | 0 | 5 |
| 21 | Healthcare | 28,456 | 1 | 0 | 1 | 1 |
| 22 | Retail | 52,136 | 6 | 0 | 6 | 0 |
| 23 | Facility services | 35,908 | 2 | 0 | 2 | 0 |

**Table 2.** Sample characteristics.

| | Age analyses | | Tenure analyses | |
|---|---|---|---|---|
| Variable | Mean | *SD* | Mean | *SD* |
| Unit age | 39.46 | 5.31 | 39.18 | 4.73 |
| Person-level *SD* of age | 10.00 | 2.45 | 10.00 | 2.45 |
| Between-group *SD* of age | 3.87 | 2.57 | 3.67 | 2.44 |
| Unit tenure | 8.30 | 4.54 | 7.12 | 4.31 |
| Person-level *SD* of tenure | 5.55 | 2.48 | 5.48 | 2.51 |
| Between-unit *SD* of tenure | 3.04 | 2.03 | 3.21 | 2.01 |
| Proportion of females in unit | 0.53 | 0.25 | 0.56 | 0.22 |

*Note*. These descriptive statistics are unweighted sample statistics.

$$r_p = \frac{t}{\sqrt{t^2 + df}} \tag{2}$$

And the variance of $r_p$ is:

$$Var\left(r_p\right) = \frac{\left(1 - r_p^2\right)^2}{n - p - 1} \tag{3}$$

The $t$-statistic was imputed with three different methods when not directly available in the source data. The first and second methods produce an exact $t$-statistic by dividing the regression coefficient by its standard error (*SE*) in the first case and working backwards from a $p$ value and its degrees of freedom in the second. The third approach achieves an approximate $t$-statistic and was used when missing data rendered the first or second method unusable. Not all cases recorded exact $p$ values; some recorded only significance levels that were exceeded (.10, .05, or .01). In these cases, we followed the recommendations of Greenberg et al. (2003) and of Stanley and Doucouliagos (2012) and assumed that a rounded $p$ value lies at the midpoint of a statistical significance range. For example, rounded $p$ values of .05 and .01 were corrected to .03 and .005, respectively.

### *Addressing statistical dependency*

There are two sources of dependency in our data. One is correlated effect sizes. This can arise when one study contributes multiple effect sizes that are correlated for reasons not able to be accounted for by the observed covariates and, if it exists, can lead to underestimation of the *SE*s of effect sizes (Hedges et al., 2010; Hedges, 2019; Tanner-Smith et al., 2016). To correct for potential bias in the *SE*s of effect sizes we applied robust variance estimation (RVE) techniques. The RVE random-effects model is expressed as (Tanner-Smith et al., 2016):

$$Y_{ij} = \beta_0 + u_i + \varepsilon_{ij} \tag{4}$$

where $i = 1\ldots k$, $j = 1\ldots m$, $Y_{ij}$ is the $j$th observed effect size from the $i$th study, $\beta_0$ is the mean population effect size,

$u_i$ is the study-level random effect where $u_i \sim N(0, \tau^2)$, and $\varepsilon_{ij}$ is the observed within-study sampling error where $\varepsilon_{ij} \sim N(0, \sigma_{ij}^2)$. The within-study error $\sigma_{ij}^2$ varies across effect sizes and is computed via Equation 3 from above. $\tau^2$ quantifies the between-study variance in effect sizes and is estimated with the method-of-moments estimator (Fisher et al., 2017). The estimation procedure of $\tau^2$ depends on the value of $\rho$, a statistical parameter which governs the degree of relation between effect sizes within the same case. We initially specified $\rho = 0.8$, a conservative value indicative of relatively high within-case dependence, and we further tested values ranging from 0.1 to 0.9. The results were robust across all values of rho and, consequently, we implemented the correction along with Tipton's (2015) small sample size correction and we report the main results with $\rho = 0.8$.

A second type of statistical dependence occurs when one study contributes effect size estimates to more than one performance measure or contributes more than one independent variable of interest (e.g., mean unit age and age of leader) to a single measure of performance. Addressing these dependencies statistically requires knowing the covariance structure of effect sizes within each study (Becker, 2000; Tanner-Smith et al., 2016). As this was unknown, we followed the convention of performing separate subgroup analyses of effect sizes (Becker, 2000; De Vibe et al., 2012).

## Methods of analysis

As is typical in meta-analyses, initially we report the point-value pooled estimates of effect sizes for unit age and unit tenure and test whether they reliably differ from zero. Further, because prior meta-analyses of age–performance relationships report considerable heterogeneity of studies' effect sizes, and because considerable heterogeneity across studies is a characteristic of this meta-analysis, we report $I^2$ and τ. $I^2$ measures the proportion of true between-study variance in the observed effect sizes and τ expresses the magnitude of between-study *SD* (Borenstein et al., 2009). Higgins et al., (2003) suggest the following intervals for $I^2$ as indicating low, moderate and high degrees of heterogeneity in effect sizes respectively: 0–50%, 50–75%, 75%+. We complement these descriptive measures with prediction intervals, which we describe in detail below. Further, because it best tests certain of the research questions of interest, we also employ random effects meta-regression. Given our small sample size, we test one moderator at a time in separate meta-regression models (Valentine, 2019; De Vibe et al., 2012).

## Software

All analyses were conducted with version 3.6.0 of the R environment (R Core Team, 2017). The RVE method was implemented with the *robumeta* package (Fisher et al., 2017) while the forest plots we present were created with the *metafor* package (Viechtbauer, 2017).[4]

[4]Supplemental materials available on request include a meta-analysis coding manual, list of control variables appearing in source studies' models (many variables are de-identified and made generic to protect organizations' identities), and software code for effect sizes, transformations to partial correlations, and meta-analyses.

# Results

The pooled effect size estimates for age and tenure are reported in Table 3. With regard to research question 1, there is no simple main effect of average member age on unit performance. That is, for all three measures—financial, customer, and operational—the 95% confidence intervals (CIs) for age include zero. In contrast, effects exist for tenure (research question 2). For two of three performance measures, financial and operational, the lower bound of the 95% CIs for average member tenure is positive and above zero, thus answering research question 2 affirmatively. Specifically, greater firm-specific experience (proxied by tenure) leads to higher unit performance.

Table 3 also answers research questions 3 (age, negatively) and 4 (tenure, affirmatively) concerning the impact of unit leader experience on the performance of the units they lead. Leaders' age is unrelated to any performance measure while the tenure of leaders is related to operational effectiveness. The presence of an effect for leaders' firm-specific experience and the absence of an effect for general work experience resemble the unit-level findings.The findings summarized in Table 3 concern central tendency and quantify the degree of uncertainty about whether a population mean differs from zero (in this case, at the 95% level of confidence). The τ statistic in Table 3 estimates the *SD* of effects across cases (Borenstein, 2019). The *SD* is high, indicating considerable heterogeneity in effect sizes in this body of evidence, consistent with what prior meta-analyses have reported in syntheses of age–performance relationships at the individual level. Such heterogeneity is an important finding in its own right, which we now consider in greater detail.

The forest plots of Figures 2 and 3 visually communicate heterogeneity in effect sizes for studies with measures of financial performance, Figure 2 for unit age and Figure 3 for unit tenure. The lack of sameness across studies is striking. Figure 2 includes cases where the effects of unit age on performance are significantly negative and others significantly positive, although the average effect size across all studies does not differ from zero. Average unit tenure also includes cases of significantly negative effects, although there is a decided surplus on the positive side and the average effect of tenure is significantly positive (Figure 3). Further, Figure 3 shows a greater proportion of studies with comparatively high within-study variability in the effects of unit tenure.

Heterogeneity in the present findings informs the range of effect sizes that can be expected to be observed in future research on age and tenure. Prediction intervals (PIs) quantify the range of plausible values—lower and upper bounds—of such future observations and in this way complement CIs (Borenstein, 2019). PIs can be appreciably wider than CIs.[5] That is the case here, as can be seen in Table 4, which reports 95% PIs. To illustrate, the CI for the average effect of unit tenure on financial performance is .01 to .15 (Table 3) while the associated PI is −.10 to .25 (Table 4). This PI indicates that the effects of work units' firm-specific experience on financial performance could reasonably be expected to be several times larger—or possibly opposite in direction of—the average effect documented here.

Heterogeneity has important implications for theory, research, and practice. Focusing on the upper bound of effects, the forest plots (Figures 2 and 3) and PIs (Table 4) imply that the effects of experience—firm-specific especially, but also

[5]A prediction interval is computed as $r_p \pm t_{df}\sqrt{SE_{rp}^2 + \tau^2}$ (Borenstein, 2019).

**Table 3.** Summary estimates and subgroup analyses.

| Variables | $N$ | $i$ | $j$ | $r_p$ | $SE$ | $df$ | 95% CI | $\tau$ | $I^2$ |
|---|---|---|---|---|---|---|---|---|---|
| Unit member age | | | | | | | | | |
| Financial | 722,987 | 10 | 36 | 0.006 | 0.012 | 8.2 | [−0.02, 0.03] | 0.03 | 92.9 |
| Customer | 17,143 | 7 | 8 | 0.016 | 0.007 | 2.9 | [−0.01, 0.04] | 0.00 | 0.0 |
| Operational | 30,062 | 7 | 30 | −0.001 | 0.017 | 4.3 | [−0.05, 0.04] | 0.05 | 79.0 |
| Unit member tenure | | | | | | | | | |
| Financial | 189,979 | 10 | 51 | 0.079 | 0.030 | 8.7 | [0.01, 0.15] | 0.07 | 96.7 |
| Customer | 13,675 | 6 | 8 | 0.036 | 0.018 | 4.6 | [−0.01, 0.08] | 0.04 | 61.9 |
| Operational | 26,754 | 8 | 39 | 0.025 | 0.009 | 4.1 | [0.00, 0.05] | 0.04 | 64.7 |
| Unit leader age | | | | | | | | | |
| Financial | 1,141,810 | 7 | 13 | 0.006 | 0.018 | 5.9 | [−0.04, 0.05] | 0.04 | 96.3 |
| Customer | 2,389 | 2 | 3 | −0.068 | 0.079 | 1.0 | [−1.00, 0.94] | 0.11 | 89.4 |
| Operational | 6,866 | 2 | 3 | 0.001 | 0.037 | 1.0 | [−0.47, 0.47] | 0.11 | 57.0 |
| Unit leader tenure | | | | | | | | | |
| Financial | 1,194,344 | 10 | 19 | 0.010 | 0.012 | 8.9 | [−0.02, 0.04] | 0.04 | 95.7 |
| Customer | 22,284 | 5 | 7 | 0.015 | 0.023 | 3.6 | [−0.05, 0.08] | 0.04 | 80.9 |
| Operational | 108,934 | 5 | 12 | 0.028 | 0.004 | 3.0 | [0.02, 0.04] | 0.02 | 69.6 |

general work experience—can become quite favorable in the right contexts. Indeed, these predicted upper bounds are high relative to observed average effects. Recall that we are analyzing partial effect sizes in this meta-analysis—that is, the effects of two types of experience after accounting for dozens of other variables that also can affect the outcomes of interest.

Research question 5 asks whether the effects of age- and tenure-based experience are equal. The most direct answer to this research question comes from analyzing the impact of one while controlling for the influence of the other. This test was performed through a meta-regression analysis which recasts age– and tenure–performance effects from point estimates based on subgroup analyses into functional relationships with control variables. To maximize the statistical power and overcome sample size limitations, we pooled (averaged) observations from all three performance measures into a single dependent variable. Our sample sizes for leader age and tenure are too small to include in meta-regression, thus question 5 is addressed at only the unit level of analysis.

Results show that controlling for age does not change the tenure–performance slope ($\beta = 0.013$, $SE = 0.013$, $df = 8.11$, $p = .325$); the main effect of tenure remains statistically significant when controlling for age ($\beta = 0.027$, $SE = 0.007$, $df = 15.23$, $p = .002$). However, when controlling for tenure the age–performance relationship is not statistically significant ($\beta = -0.003$, $SE = 0.006$, $df = 15.19$, $p = .627$). Furthermore, in the absence of a control for tenure, there is a statistically significant increase in the relationship between age and performance ($\beta = 0.034$, $SE = 0.008$, $df = 2.64$, $p = .030$). That is, it is the change in the age–performance slope due to removing the tenure control that is significant at the .05 level, not the effect of age itself. Figure 4 visualizes the findings. Note how the distribution of the age–performance relationship shifts to the left (i.e., to 0) when controlling for tenure. These findings confirm that firm-specific experience is the more important type of experience affecting unit performance, thus answering research question 5.

Research question 6 asks if disparities of age and tenure in work groups affect their performance and research question 7 asks if nonlinearities exist in experience–performance relationships. No such effects were detected. That is, the age–performance relationship does not depend on age dispersion as measured by the within-unit $SD$ of member age nor on the value of age-squared (a test of nonlinearity). Likewise for tenure, neither tenure dispersion within work units nor tenure-squared influenced the effects of tenure on performance. Table 5 summarizes these tests. Small sample sizes make for low statistical power in these analyses and so the results should be treated with caution. That said, this lack of significant findings suggests that the search for moderators of experience-performance relationships may be better directed at contextual factors rather than at work units' internal attributes.

## Discussion

This section begins by addressing principal findings and selected methodological issues, followed by considerations regarding future research and theory, including a suggested way of creating theory that integrates psychological and economic scholarship. Implications of the findings for the workplace are then addressed, with an emphasis on the risks to employers who choose to replace tenured employees with gig and contract workers. The final section discusses the changing nature of organizationally based research and problems and prospects for such research making contributions accessible to the scientific community.

### Principal findings

This research provides strong evidence that higher levels of firm-specific experience increase work unit performance. Specifically, units' financial and operational performance rise with increased tenure of unit members with the current employer. Age, and the general work experience it brings, did not affect unit performance after accounting for tenure. It was also found that the firm-specific experience of unit leaders positively affects the operational effectiveness of the units

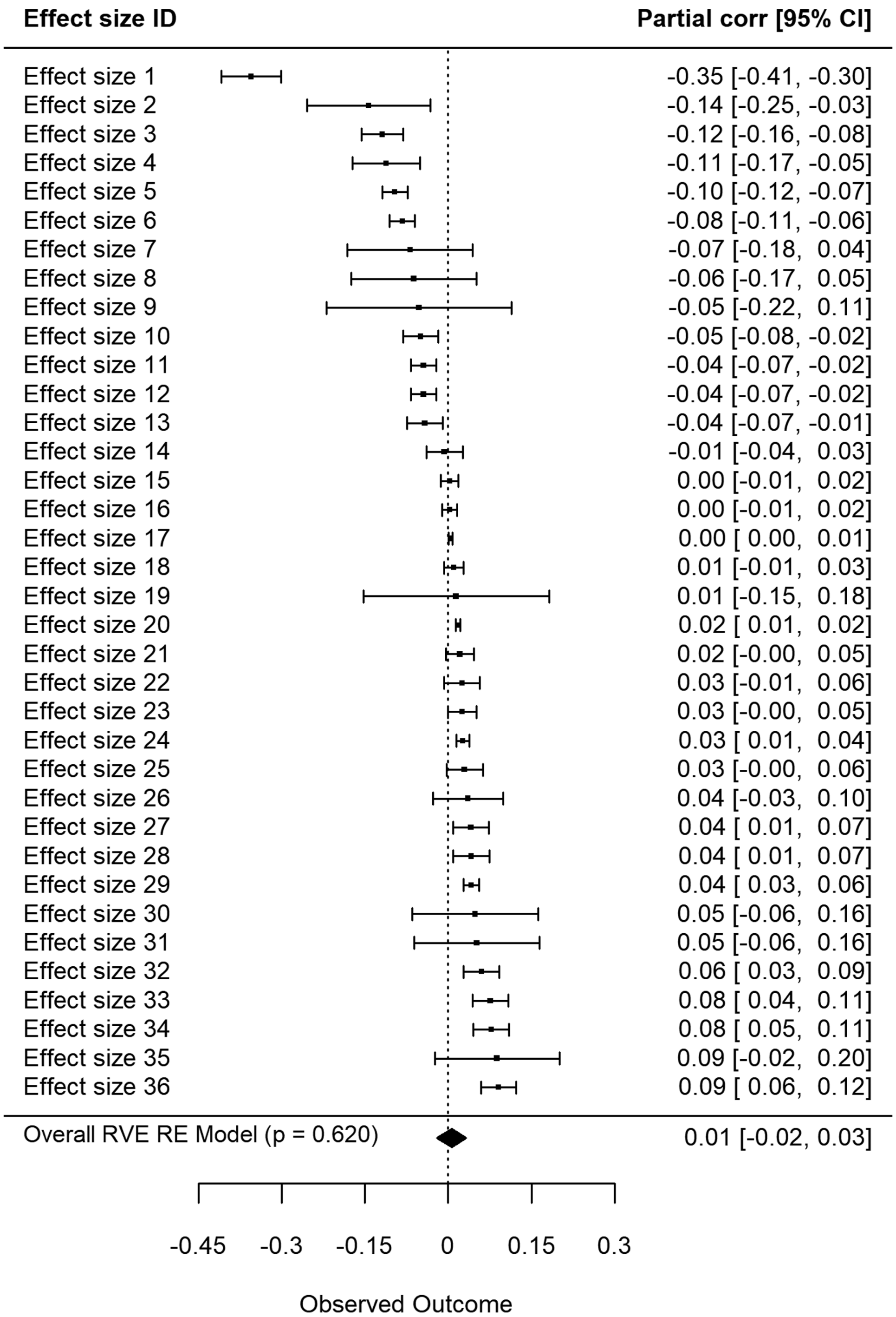


**Figure 2.** Forest plot of average age in unit on measures of financial performance.

they lead while leaders' general work experience does not. Within-unit variability in members' experience was unrelated to performance and no nonlinearities were detected in relationships of experience to performance.[6]

[6]What to make of the absence of findings regarding nonlinearities and within-unit dispersion is an open issue. Methodological influences could be at play, such as range restriction in measured variables that inhibit the detection of significant relationships and, while occasional evidence of nonlinear and dispersion relationships appears in the literature, it is difficult to discern its replicability because of publication biases that often keep nonsignificant results out of the literature.

It is not clear why the customer-based performance measures showed no effects of either unit or leader tenure. One possible explanation is that customers respond to factors not captured here, such as unit members' dispositional attributes or their styles of interaction. Further, leaders may have less direct contact with customers and thus less opportunity for influence. The one measure that registered a positive effect of leader tenure was operational effectiveness. This may indicate that, with increased tenure, leaders become more apt at bringing to the units they lead the

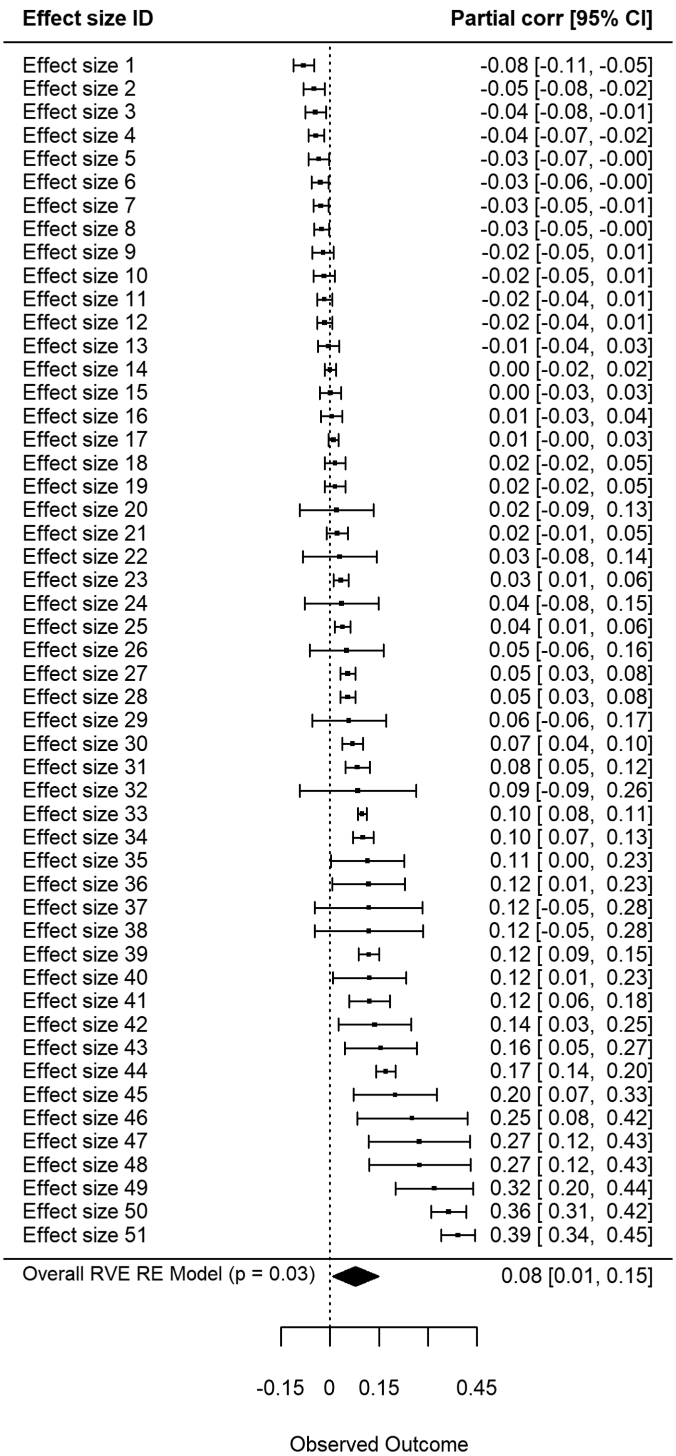


**Figure 3.** Forest plot of average tenure in unit on measures of financial performance.

organizational resources that enable them to operate more effectively. Also, with tenure and the visibility into multiple units it brings, leaders may become increasingly knowledgeable about which ways of working are best suited to a given unit and coach units accordingly. In the absence of direct measures of tenure-based differences in how leaders influence the units they lead these interpretations are plausible but speculative.

**Table 4.** Prediction intervals of partial correlation coefficients for age and tenure effects.

| | Work units | | Unit leaders | |
|---|---|---|---|---|
| | Age | Tenure | Age | Tenure |
| Financial | −.07 to .08 | −.10 to .25 | −.11 to .12 | −.09 to .11 |
| Customer service | −.01 to .05 | −.09 to .15 | – | −.13 to .16 |
| Operational | −.15 to .15 | −.09 to .15 | – | −.04 to .09 |

*Note*. Small sample sizes preclude the calculation of prediction intervals involving unit leader age for two of three dependent variables.

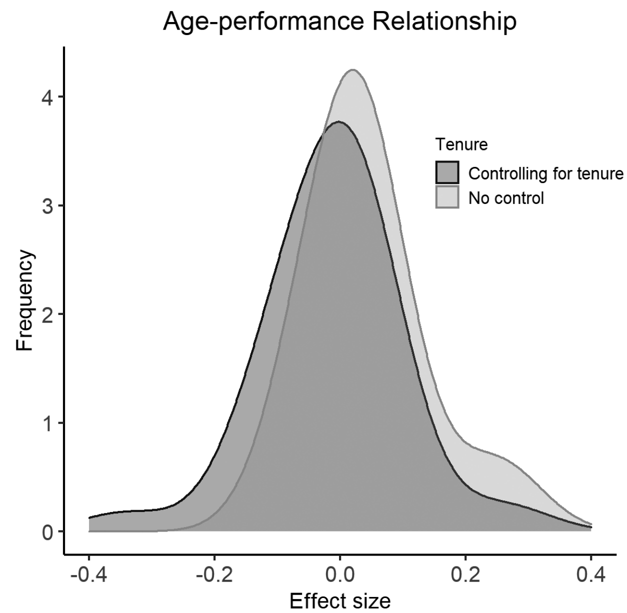


**Figure 4.** A comparison of the age–performance slope, with and without a tenure control.

## Methodological considerations

### *Model specification differences*

The estimates of age and tenure effects in this meta-analysis came from coefficients in ordinary least squares regression models, the exact specifications of which differed study-to-study. Consequently, some variance in coefficients may reflect model specification differences (Aloe et al., 2016; Stanley & Jarrell, 1989).

Control variables are a primary source of differences in model specifications, for good reason. Controls that are important to the analysis of the performance of health care teams in a hospital ward necessarily differ from those important to crews in fast food shops, building maintenance teams, operators on the factory floor, or associates in retail bank branches. The tailoring of control variables to each situation—and thus creating models with differing specifications—is a strength of the research, given the importance of accounting for relevant covariates when studying age-related variables and work (Barnes-Farrell et al., 2019). By contrast, forcing identical model specifications in all studies in the pursuit of methodological rigor and "pure" apples-to-apples comparisons of coefficients across studies would have resulted in age and tenure coefficients more rather than less

**Table 5.** Moderator analyses for age- and tenure-performance in work units.

| Variables | *i* | *j* | Intercept (*SE*) | | Slope (*SE*) | | *df* | *p* |
|---|---|---|---|---|---|---|---|---|
| Age meta-analyses | | | | | | | | |
| Tenure | 21 | 93 | −0.003 | (0.006) | 0.034 | (0.008) | 2.64 | .030 |
| Age squared | 21 | 93 | 0.008 | (0.007) | −0.007 | (0.008) | 19.00 | .388 |
| Age *SD* | 21 | 93 | −0.006 | (0.004) | 0.008 | (0.008) | 1.90 | .382 |
| Tenure meta-analyses | | | | | | | | |
| Age | 23 | 136 | 0.027 | (0.007) | 0.013 | (0.013) | 8.11 | .325 |
| Tenure squared | 23 | 136 | 0.039 | (0.015) | −0.012 | (0.018) | 21.00 | .509 |
| Tenure *SD* | 23 | 136 | 0.061 | (0.024) | −0.033 | (0.025) | 2.28 | .300 |

confounded due to the failure to account for situationally specific influences in naturally occurring settings. For purposes of making the strongest inferences about age and tenure effects, we believe that the best models are those that as much as possible account for local workplace realities, not just in this research but in all future research. Also, an inherent strength of meta-analysis is its techniques for accounting for potential method-based variability in findings, such as RVE as was applied in this meta-analysis.

Another source of differences between studies was that final model specifications almost always reflected inputs from the organizations in which the studies were conducted. Because they intended to act on the results, participating organizations were motivated to rely on best possible models and their representatives often had very astute insights and suggestions about potential variables for inclusion and interpretation of results. Sometimes these inputs arose once analyses were underway, such as when additional variables and new data became available due to their emerging relevance. The research process thus was not one of complete control by researchers over model specifications. Indeed, that the research process was a partnering between "outsider" researchers and "insider" organizational members was much to the benefit of simultaneously maximizing rigor and relevance.

### *Accounting for commonly unobserved factors*

The data and analytical approach underlying our study addresses limitations in existing economic research on firm-level productivity and performance. Econometric studies of age–performance relationships at the firm level include a mix of cross-sectional and time series analyses that rely on very limited or highly aggregated sets of explanatory variables and organization-specific statistical controls. Nonlabor factors influencing output and value added measures, such as facility, customer, and market characteristics are often ignored, as are critical workforce management practices—e.g., staffing ratios, supervisory spans of control, reward structure—that also can influence performance or condition the impact of age and tenure. Such omissions are natural limitations arising from the typical practice of relying on matched employee–employer panel data sets created from periodic surveys. The limitations of such typical datasets are not fully overcome, in our view, by reliance on fixed-effects regression modeling, a common methodology used by economists analyzing panel data sets to capture the effects of unobservable latent characteristics unique to each firm. Fixed effects models produce estimates of the effects of specific factors like age and tenure on performance but they do so at the considerable cost of discarding information emanating from the variance of performance across firms. The studies in this meta-analysis, which are based on downloads of extensive firm-level workforce and organizational data, capture intra-firm variation and account for a wide range of control variables within each firm. Formally comparing the estimates of partial correlations from these highly controlled models, as done in this meta-analysis, supports conclusions that are based on the most inclusive set of observables available. In the age of big data, factors once treated as unobservable—and thus ignorable—can now be explicitly measured and investigated. In addition to controlling for confounding influences, using data to make observable what has previously been treated as unobservable is essential to gaining insights into how contextual factors within organizations may moderate experience–performance relationships, something that fixed-effects models simply cannot achieve, and serves as a reminder that methodological choices do indeed influence theory development.

### *Generalizability*

Several features of the research support the generalizability of findings to other workplaces: All units were naturally occurring entities in employing organizations; the organizations, work units, and types of work performed were diverse; the number of employees in the studies were large and the durations of study long; and multiple dimensions of performance were measured. However, all work units existed in organizations willing to hire outsiders to conduct applied research, which may suggest a limitation to generalizability as not all organizations are prone to do so (e.g., larger organizations prevail in this research).

## Implications for research and theory

One important implication of this research is that, to deliver unambiguous insights, future research and theory on employee age and work performance need to account for tenure. Employee age and tenure are necessarily correlated yet their effects are distinguishable, as this study's findings demonstrate. Specifically, the finding that the impact of firm-specific human capital (tenure) is larger than that of general human capital (age) indicates that firm-specific experience is the more proximal determinant of performance and the further finding that the impact of age goes to zero after accounting for tenure indicates that tenure with current employer is a mediator of the impact of general work experience. That is, the value of the education, skills, and capabilities that are acquired elsewhere and brought to a current employer is expressed through the buildup of firm-specific experience, experience that activates general human capital and adapts its application to the current work context. While the necessity of accounting for employee tenure is demonstrated at the unit level of analysis in this research, we believe that the point also holds for employee age–performance research and theory at other levels of analysis, especially individual but also very likely at the organizational level of analysis. An exception to the need to account for tenure applies to the self-employed, which in the United States constitute 10% of the workforce (US Bureau of Labor Statistics, 2021).

A second important implication of this study relates to the observed heterogeneity in the effects of age and tenure, especially tenure. Heterogeneity is a signal that much remains to be discovered about factors that moderate the impact of age and tenure on performance. There is great need for research that focuses on such moderators, the results of which will provide evidence that will drive the formation of more situationally specific theoretical accounts of the impact of age and age-related experience on performance. Such advances will lead not only to more accurate but also more useful theories.

The set of potential moderators seems rather large. Occupational differences may moderate. For example, the impact of a physician's general human capital on performance seems less likely to be as strongly mediated by organizational tenure than that of a design engineer. Tenure with an employer cultivates in that engineer knowledge of details of such things as production capabilities and limitations and of proprietary libraries of designs that can be critical to channeling their stock of general human capital into valuable contributions in their current organization. Organizations' workforce management practices are likely moderators, such as the extent to which work is organized around units, differences in the amount of autonomy given to those units, effects of different leadership styles, and the extent to which incentives are provided for unit-level performance. Industry differences, too, are likely moderators. In the aviation industry, standardization of instrumentation and highly detailed role responsibilities leave little opportunity, it would seem, for average tenure with a current employer to influence the in-flight performance of cockpit crews. Industries with comparatively low barriers to worker entry and high rates of turnover, such as some sectors of the construction, entertainment, and hospitality industries, may be less likely to have work unit performance tuned to tenure. In contrast, tenure may matter greatly, for example, in the commercial brokerage sector of the insurance industry where tenure brings stable client relationships and mastery of complex products as they relate to nuanced business needs and in the pharmaceutical industry where tenure may be essential to enabling R&D units to make the most effective use of proprietary scientific knowledge and processes when creating new drugs and therapies.

One other implication from this research is the need to find the way toward more integrative interdisciplinary theory. As this article's literature review illustrated, psychology and economics share an interest in age and performance yet there is little integration of the two disciplines' abundant bodies of evidence or theory. They stand apart. Only sporadically are the two referenced in the same research on age–tenure–performance relationships and sometimes to create a competition between derived hypotheses (e.g., Ng & Feldman, 2013a). We believe that the fruitful path to integrative theory starts with the assumption that each discipline, and thus each discipline's theories and bodies of research evidence, have substantial validity ("truth") to them. Further, rather than constructing empirical tests designed to reject one versus the other, we believe a better way forward is to construct empirical tests for the purpose of identifying circumstances in which predictions derived from each perspective hold most true or are most informative. So, for example, economics as a discipline has more to say than psychology about industry-level differences and thus its literature will be the more valuable source of testable hypotheses about how industry differences influence the value of firm-specific human capital for work unit performance. Both literatures offer insights into the impact of rewards and incentives on collective performance and thus both are relevant to understanding how differences in organizational reward practices may influence the impact of unit tenure on unit performance. Psychology's literature, given its extensive research on leadership, will be the more valuable source of insight into moderating influences of leadership style. Future research that emphasizes complementarity rather than competition will best advance the identification of sources heterogeneity in the effects of age and tenure on work performance. Fortunately, such an empirical approach is made realistic by the availability of "big data" in organizations which enables the efficient testing of multiple moderators in a single investigation. Also essential to productive interdisciplinarity is an increased willingness in each discipline to learn from the other and adopt the other's strengths with regard to data collection, measurement, analytic methods, and interpretations.

## Implications for practice

### *Defeating ageism*

Ageism in the workplace is an unfortunate reality with consequences. It reduces the willingness to hire older workers and to endorse practices that support age diversity in organizations, according to a recent meta-analysis (Jones et al., 2017), with the magnitudes of effect sizes for ageism just shy of those for racism. Our study's finding of no relationship between average age of employees and their work units' performance is not alone in contributing to an evidentiary basis for defeating ageism. We believe it is time to ratchet up the prominence of research results showing that age has no consistently detrimental impact on business performance. Whether through business journalism or other means, this research-based fact demands greater appreciation. Ageism also can sometimes find its expression by reference to generational differences. The concept of generational differences may have intuitive appeal but the consensus report of the National Academies of Sciences, Engineering, and Medicine (2020) is that there is little reason to believe that multigenerational workplaces are dysfunctional or unproductive and our finding that greater age dispersion in work units is unrelated to performance is consistent with that conclusion. The National Academies report also highlights conceptual and methodological limitations to existing research on generations as a useful concept in the workplace. These limitations include the failure to separate the effects of age from generation (cohort) membership and the failure to specify how it is that age-based cohorts come to acquire systematic differences in attitudes and behaviors relevant to the workplace.

### *The value of traditional forms of employment*

Two key findings in this research, that tenure with one's current employer leads to higher unit performance and that general work experience has no effect after accounting for tenure, have a very significant implication for employers: Organizations opting for traditional terms of employment that build tenure are advantaged over competitors opting for nonemployee workforces, such as gig, contingent, and contract workers, even when those nonemployees may have many years of work experience. This should give pause to employers seeking to create or convert to nonemployee workforces.

Organizations' opting for nonemployee workforces do so for assorted stated rationales, such as a need for increased organizational agility, flexibility, and the pursuit of change and reinvigoration. Inescapably, these reasons coincide with an interest in reducing costs. Wages typically rise with increasing employee tenure and the wages of older, tenured employees are likely to exceed those paid to a replacement contract or gig worker of any age. Employers also realize cost reductions by not offering benefits that employees would receive. A shortcoming of the cost-driven approach is that it fails to account for the value created through employee tenure. That value is testable and demonstrable, as this research shows. In fact, value relative to costs was quite directly assessed in two ways in this research. One, the majority of studies controlled for wages paid when assessing tenure's effects and, two, many of the studies used a performance measure, profit, which explicitly accounts for the financial gains created by a work unit relative to the costs (including wages) of creating those gains. Thus, while the costs of a traditional employee workforce may exceed the costs of gig or contract workers who replace them, the value created by tenured employees often exceeds those costs *irrespective of the age of employees*. This is not to say that workforce change is a bad thing or that there is no role for alternatives to employees.[7] Rather, our caution to employers is that replacing high value assets—tenured employees—with lower-cost alternatives can result in poorer financial and operational performance.

## Organizationally based research, big data, and scientific contributions

The research reported here in many ways heralds change to come in the nature of research and research reports. Specifically, we believe that the partnering of external researchers and host organizations is destined to become an increasingly common research process. Today's organizations—large ones, especially—control a vast amount of ever-expanding data and are populated by research-savvy individuals. Rather than simply giving over data to outsiders (consultants, academics) for analysis, organizations progressively will demand active partnering in the research process. This, we believe, will serve scientific objectives very well. The insider–outsider research partnership will bring about a more optimal use of available data resulting in better measurement of constructs, stronger tests of relationships, and more rapid development of theory that recognizes contextual influences. Ideally, such a partnering will yield a new stream of published research studies on a wide variety of work-related topics, although no such stream has yet materialized.

In addition to the emergence of new research reports from such collaborations, there are bodies of existing, unpublished studies on a number of topics like that used in this meta-analysis. Such bodies of evidence, when sufficient in number and merit, can be subject to meta-analyses and other forms of integrative review and can become another potential new stream of research reports. A distinctive feature of such evidence is its freedom from publication bias, the favoring of reports with statistically significant findings over those with null results. It is estimated that more than 90% of published studies in psychology report significant results, a rate that is several times higher than in physical sciences (Fanelli & Ioannidis, 2013; Kühlberger et al., 2014). Consequently, any synthesis, meta-analysis or otherwise, of published studies may itself be biased. It will be instructive to see the extent to which syntheses of such unpublished studies may yield conclusions differing from those of published studies on the same topic.

The studies that went into this article's meta-analysis embody characteristics of big data such as high volume and variety of data and large numbers of observations. King and Persily (2019, p. 2) assert that "big data collected by firms about individuals and human societies is more informative than ever, which means it has increasing scientific value." Making the most of that scientific value requires getting research reports into the literature, which itself may become a challenging process. For example, open science practices that some journals require as ways to ensure research integrity, practices such as full disclosure and data sharing, often will be impossible for organizationally based, big-data research to comply with because full disclosure and data sharing can "violate individual privacy or help a firm's competitors" (King & Persily, 2019, p. 2; also see Guzzo, Schneider, & Nalbantian, in press). In such cases, adapted journal review processes may be needed to safeguard integrity while mitigating risks and removing barriers to publication. An example of an adapted peer review process is entrusting a third party—neither a article's authors nor journal reviewers—that is bound by nondisclosure agreements to authenticate a study's underlying data and certify data integrity. Also, the availability of big data creates opportunities for efficiently testing numerous relationships simultaneously, the number of which can far exceed the number able to be hypothesized from the stock of current theories. Consequently, we believe that greater acceptance of big data-driven abductive and inductive methods of analysis and theory development will be needed to counterbalance the heavy weight currently given to classical hypothesis testing and hypothetico-deductive reasoning in the research and publication process (Guzzo et al., in press; also and Prosperi et al., 2019). Other changes, too, surely will be needed to remove barriers that would otherwise keep good research out of the view of the scientific community.

[7]Older workers are proportionately over-represented in alternative work arrangements (Kosanovich, 2018), suggesting a fit between such arrangements and many members of this segment of the workforce.

## References

Aloe, A. M. (2014). An empirical investigation of partial effect sizes in meta-analysis of correlational data. *The Journal of General Psychology,* 141(1), 47–64. doi:10.1080/00221309.2013.853021

Aloe, A. M., Tanner-Smith, E. E., Becker, B. J., & Wilson, D. B. (2016). *Synthesizing bivariate and partial effect sizes*. The Campbell Collaboration. doi:10.4073/cmpn.2016.2

Aloe, A. M., & Thompson, C. G. (2013). The synthesis of partial effect sizes. *Journal of the Society for Social Work and Research,* 4(4), 390–405. doi:10.5243/jsswr.2013.24

Backes-Gellner, U., & Veen, S. (2009). The impact of aging and age diversity on company performance. *SSRN Electronic Journal*. doi:10.2139/ssrn.1346895

Baker, G., Gibbs, M., & Holmstrom, B. (1994). The wage policy of a firm. *Quarterly Journal of Economics,* 109(4), 921–955. doi:10.2307/2118352

Bal, P. M., & Boehm, S. A. (2019). How do i-deals influence client satisfaction? The role of exhaustion, collective commitment, and age diversity. *Journal of Management,* 45(4), 1461–1487. doi:10.1177/0149206317710722

Barnes-Farrell, J. L., Petery, G. A., Cleveland, J. N., & Matthews, R. A. (2019). Age(ing) and work attitudes. In K. S. Schults & G. A. Adams (Eds.), *Aging and work in the 21st century* (2nd ed., pp. 146–170). Taylor & Francis.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis with special reference to education*. The University of Chicago Press.

Becker, B. J. (2000). Multivariate meta-analysis. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 499–525). CAL Academic Press. doi:10.1016/B978-012691360-6/50018-5

Bezrukova, K., Jehn, K. A., Zanutto, E. L., & Thatcher, S. M. B. (2009). Do workgroup faultlines help or hurt? A moderated model of faultlines, team identification, and group performance. *Organization Science,* 20(1): 35–50. doi:10.1287/orsc.1080.0379

von Bonsdorff, M. E., Zhou, L., & Wang, M. (2016). Employee age and company performance: An integrated model of aging and human resource management practices. *Journal of Management,* 44(8), 3124–3150. doi:10.1177/0149206316662314

Borenstein, M. (2019). Heterogeneity in meta-analysis. In H. M. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (3rd ed., pp. 453–468). Russell Sage Foundation.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons. doi:10.1002/9780470743386

Børing, P. (2019). The relationship between firm productivity, firm size and CSR objectives for innovations. *Eurasian Business Review,* 9, 269–297. doi:10.1007/s40821-019-00123-y

Carstensen, L. L. (2006). The influence of a sense of time on human development. *Science,* 312, 1913–1915. doi:10.1126/science.1127488

Casanova, M. (2013). *Revisiting the Hump-shaped Wage Profile*. UCLA Department of Economics Working Paper. http://www.econ.ucla.edu/casanova/files/casanova_wage_older_workers.pdf

Charles, S., & Carstensen, L. L. (2010). Social and emotional aging. *Annual Review of Psychology,* 61, 383–409. doi:10.1146/annurev.psych.093008.100448

Choudry, P., & Haas, M. R. (2018). Scope versus speed: Team diversity, leader experience, and patenting outcomes for firms. *Strategic Management Journal,* 42(1), 30–57. doi:10.1002/smj.3215

Clark, A., Oswald, A., & Warr, P. (1996). Is job satisfaction U-shaped in age? *Journal of Occupational and Organizational Psychology,* 68, 57–81. doi:10.1111/j.2044-8325.1996.tb00600.x

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences*. Lawrence Erlbaum Associates. doi:10.4324/9780203774441

De Vibe, M., Bjørndal, A., Tipton, E., Hammerstrøm, K., & Kowalski, K. (2012). Mindfulness based stress reduction (MBSR) for improving health, quality of life, and social functioning in adults. *Campbell Systematic Reviews,* 8(1), 1–127. doi:10.4073/csr.2012.3

Fanelli, D., & Ioannidis, J. P. (2013). US studies may overestimate effect sizes in softer research. *Proceedings of the National Academy of Sciences,* 110, 15031–15036. doi:10.1073/pnas.1302997110

Fisher, Z., Tipton, E., Zhipeng, H., & Fisher, M. Z. (2017). Package 'robumeta'. https://cran.r-project.org/web/packages/robumeta/robumeta.pdf

Frederiksen, A., Lange F., & Kriechel, B. (2017). Subjective performance evaluations and employee careers. *Journal of Economic Behavior & Organization,* 134(C), 408–429. doi:10.1016/j.jebo.2016.12.016

Gonzalez-Mulé, E., Cockburn, B. S., McCormick, B. W., & Zhao, P. (2019). Team tenure and team performance: A meta-analysis and process model. *Personnel Psychology,* 73, 151–198. doi:10.1111/peps.12319

Greenberg, D. H., Michalopoulos, C., & Robins, P. K. (2003). A meta-analysis of government-sponsored training programs. *Industrial and Labor Relations Review,* 57(1): 31–53. doi:10.1177/001979390305700102

Grund, C., & Westergard-Neilsen, N. (2008). Age structure of the workforce and firm performance. *International Journal of Manpower,* 29(5), 410–422. https://www.emerald.com/insight/content/doi/10.1108/01437720810888553/full/html

Guzzo, R. A., Schneider, B., & Nalbantian, H. R. (in press). Open science, closed doors: The perils and potential of open science for research-in-practice. *Industrial and Organizational Psychology*.

Hackman, J. R., & Oldham, G. R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance,* 16, 250–279. doi:10.1016/0030-5073(76)90016-7

Heckman, J. J., Lochner, L. J., & Todd, P. E. (2008). Earnings functions and rates of return. *Journal of Human Capital,* 2(1), 1–31. doi:10.1086/587037

Hedges, L. V. (2019). Stochastically dependent effect sizes. In H. M. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (3rd ed., pp. 281–298). Russell Sage Foundation.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods,* 1(1), 39–65. doi:10.1002/jrsm.5

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557–560. doi:10.1136/bmj.327.7414.557

Jones, K. P., Sabat, I. E., King, E. B., Ahmad, A., McCausland, T. C., & Chen, T. (2017). Isms and schisms: A meta-analysis of the prejudice-discrimination relationship across racism, sexism, and ageism. *Journal of Organizational Behavior,* 38, 1076–1110. doi:10.1002/job.2187

King, G., & Persily, N. (2019). A new mode for industry-academic partnership. *PS: Political Science and Politics*. doi:10.1017/S1049096519000l021

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. doi:10.1016/j.jcm.2016.02.012

Kosanovich, K. (2018). *Workers in alternative employment relationships*. Spotlight on Statistics, U.S. Bureau of Labor Statistics. https://www.bls.gov/spotlight/2018/workers-in-alternative-employment-arrangements/pdf/workers-in-alternative-employment-arrangements.pdf

Kühlberger, A., Fritz, A., & Scherndl, T. (2014). Publication bias in psychology: A diagnosis based on the correlation between effect size and sample size. *PLoS One,* 9(9), e105825. doi:10.1371/journal.pone.0105825

Lallemand, T., & Rycx, F. (2009). Are old workers harmful for firm productivity? *De Economist,* 157, 273–292. doi:10.1007/s10645-009-9126-5

Lazear, E. P. (1979). Why is there mandatory retirement? *Journal of Political Economy,* 87(6), 1261–1284. doi:10.1086/260835

Lazear, E. (1995). *Personnel economics*. MIT Press.

Lazear, E., & Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy,* 89, 841–864. doi:10.3386/w0401

Lemieux, T. (2006). The "Mincer equation" thirty years after schooling, experience, and earnings. In S. Grossbard (Ed.), *Jacob Mincer a pioneer of modern labor economics* (pp. 127–145). Springer.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications, Inc. doi:10.1016/j.jhsa.2006.09.002

Maestas, N., Mullen, K. J., & Powell, D. (2016). *The effect of population aging on economic growth, the labor force and productivity*

(NBER Working Paper No. w22452). National Bureau of Economic Research. http://www.nber.org/papers/w22452

Mahlberg, B., Freund, I., Cuaresma, J. C., & Prskawetz, A. (2013). Ageing, productivity and wages in Austria. *Labour Economics,* 22, 5–15. doi:10.1016/j.labeco.2012.09.005

McCauley, C. D., DeRue, D. S., Yost, P. R., & Taylor, S. (Eds.). (2014). *Experience-driven leader development*. Wiley.

McEvoy, G. M., & Cascio, W. F. (1989). Cumulative evidence of the relationship between employee age and job performance. *Journal of Applied Psychology,* 74(1), 11–17. doi:10.1037/0021-9010.74.1.11

Medoff, J. L., & Abraham, K. G. (1980). Experience, performance, and earnings. *The Quarterly Journal of Economics,* 95(4), 703–736. doi:10.3386/w0278

Medoff, J. L., & Abraham, K. G. (1981). Are those paid more really more productive? The case of experience. *Journal of Human Resources,* 16(2), 186–216. doi:10.2307/145508

Mincer, J. (1958). Investment in human capital and the personal income distribution. *Journal of Political Economy,* 66, 281–302. doi:10.1086/258055

Mincer, J. (1974). *Schooling, experience, and earnings*. Columbia University Press. doi:10.1086/260336

Murphy, K. M., & Welch, F. (1990). Empirical age-earnings profiles. *Journal of Labor Economics,* 8(2), 202–229. doi:10.1086/298220

Nalbantian, H. R. (2014). Gauging the productivity of older workers: Issues and challenges. Conference Proceedings: Adapting to an Ageing Workforce. Stanford University Center on Longevity. 162.144.124.243/~longevl0/wp-content/uploads/2016/07/Proceedings-FINAL-3.18.pdf

Nalbantian, H. R., Guzzo, R. A., Doherty, J., & Kieffer, D. (2003). *Play to your strengths: Managing your company's internal labor markets for lasting competitive advantage*. McGraw Hill.

Nalbantian, H., & K. Marciniak (in press). *Assessing the productive contribution of older workers: The role of spillover*. Mercer White Paper.

National Academies of Sciences, Engineering, and Medicine. (2020). *Are generational categories meaningful distinctions for workforce management?* The National Academies Press. doi:10.17226/25796

Ng, T. W., & Feldman, D. C. (2008). The relationship of age to ten dimensions of job performance. *Journal of Applied Psychology,* 93(2), 392. doi:10.1037/0021-9010.93.2.392

Ng, T. W., & Feldman, D. C. (2013a). Does longer job tenure help or hinder job performance? *Journal of Vocational Behavior,* 83(3), 305–314. doi:10.1016/j.jvb.2013.06.012

Ng, T. W., & Feldman, D. C. (2013b). A meta-analysis of the relationships between age and tenure with innovation-related behavior. *Journal of Occupational and Organizational Psychology,* 86(4), 585–616. doi:10.1111/joop.12031

Prosperi, M., Bian, J., Buchan, I. E., Koopman, J. S., Sperrin, M., & Wang, M. (2019). Raiders of the lost HARK: A reproducible inference framework for big data science. *Palgrave Communications,* 5(125). doi:10.1057/s41599-019-0340-8

R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rhodes, S. R. (1983). Age-related differences in work attitudes and behavior: A review and conceptual analysis. *Psychological Bulletin*, 93(2), 328–367. doi:10.1037/0033-2909.93.2.328

Salthouse, T. (2012). Consequences of age-related cognitive declines. *Annual Review of Psychology,* 63, 201–226. doi:10.1146/annurev-psych-120710-100328

Schneider, B. (2018). Being competitive in the talent management space. *Industrial and Organizational Psychology,* 11(2), 231–236. doi:10.1017/iop.2018.10

Schneider, B., & Pulakos, E. D. (in press). *Expanding the IO mindset to organizational success*. Industrial and Organizational Psychology.

Sharpe, A. (2011). Is ageing a drag on productivity growth? A review article on ageing, health, and productivity: The economics of increased life expectancy. *International Productivity Monitor,* (21), 82–94.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin,* 86(2), 420.

Skirbekk, V. (2004). Age and individual productivity: A literature survey. *Vienna Yearbook of Population Research,* 2(1), 133–154. doi:10.1553/populationyearbook2004s133

Stanley, T. D., & Doucouliagos, H. (2012). *Meta-regression analysis in economics and business*. Routledge. doi:10.4324/9780203111710

Stanley, T. D., & Jarrell, S. B. (1989). Meta-regression analysis: A quantitative method of literature surveys. *Journal of Economic Surveys*, 3(19), 54–67. doi:10.1111/j.1467-6419.1989.tb00064.x

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life-Course Criminology,* 2(1), 85–112. doi:10.1007/s40865-016-0026-5

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods,* 20(3), 375. doi:10.3102/1076998615606099

U.S. Bureau of Labor Statistics. (2021). Labor force statistics from the Current Population Survey: Labor force characteristics. https://www.bls.gov/cps/lfcharacteristics.htm#self

Valentine, C. (2019). Incorporating judgements about study quality into research synthesis. In H. M. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (3rd ed., pp. 129–140). Russell Sage Foundation.

Verissimo, J., Verhaeghen, P., Goldman, N., Weinstein, M., & Ullman, M. T. (2021). Evidence that ageing yields improvements as well as declines across attention and executive functions. *Nature and Human Behaviour,* August. doi:10.1038/s41562-021-01169-7

Vevea, J. L., Zelinsky, N. A. M., & Orwin, R. G. (2019). Evaluating coding decisions. In H. M. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (3rd ed., pp. 173–206). Russell Sage Foundation.

Viechtbauer, W. (2017). Forest plot with subgroups. http://www.metafor-project.org/doku.php

Waldman, D. A., & Avolio, B. J. (1986). A meta-analysis of age differences in job performance. *Journal of Applied Psychology*, 71, 33–38. doi:10.1037/0021-9010.71.1.33

Wang, G., Holmes, R. M., Jr, Oh, I. S., & Zhu, W. (2016). Do CEOs matter to firm strategic actions and firm performance? A meta-analytic investigation based on upper echelons theory. *Personnel Psychology,* 69(4), 775–862. doi:10.1111/peps.12140

Westelius, N. J., & Liu, Y. (2016). *The impact of demographics on productivity and inflation in Japan* (IMF Working Paper, No. 16/237). International Monetary Fund. doi:10.1142/S1793993317500089

Wilson, D. B. (2019). Systematic coding for research synthesis. In H. M. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (3rd ed., pp. 153–172). Russell Sage Foundation.

Zacher, H., & Frese, M. (2009). Remaining time and opportunities at work: Relationships between age, work characteristics, and occupational future time perspective. *Psychology and Aging,* 24(2), 487–493. doi:10.1037/a0015425