Age Bias in Laboratory and Field Settings: A Meta-Analytic Investigation[1]

RANDALL A. GORDON[2]
University of Minnesota, Duluth

RICHARD D. ARVEY
University of Minnesota, Twin Cities

A meta-analytic review of age-discrimination research from laboratory and field settings spanning the years 1963 to 1998 revealed a significant, yet modest overall effect size with younger applicants and workers evaluated more positively than older applicants and workers. The present predictions and findings were compared and contrasted with those from previous meta-analytic investigations by Kite and Johnson (1988) and by Finkelstein, Burke, and Raju (1995). A number of significant moderational relationships were revealed, including negative linear relationships between the relative generalizability of the research (in terms of participants, setting, and stimulus materials) and effect size. A significant negative relationship was also found between the publication year of a study and effect size. In addition, the type of design used (between-subjects vs. within-subject) and the specific type of dependent measures used to evaluate applicants and workers (e.g., potential for development, interpersonal skills, worker reliability and stability) were found to significantly moderate effect size. These results, along with suggestions for future investigations, are discussed.A

A

Beginning with the work of Tuckman and Lorge (1952) and Kirchner and Dunnette (1954), research on attitudes toward older workers has spanned five decades. Similarly, research investigating age bias and stereotyping that operate to disadvantage relatively older employees in evaluation and decision-making contexts has been accumulating over the years. However, the investigation of age bias and discrimination toward older workers has generated a rather mixed set of findings. There is evidence supporting the existence of negative evaluations of older applicants and job incumbents (Cleveland & Landy, 1981; Crew, 1984; Gordon, Rozelle, & Baxter, 1988; Haefner, 1977; Lee & Clemmons, 1985; Rosen & Jerdee, 1976a; Singer, 1986).

However, other studies have revealed no significant differences (Erber, Caiola, & Pupo, 1994; Hitt & Barr, 1989) or, in some instances, more positive

1

evaluations of older workers than of younger workers (Gibson, Zerbe, & Franken, 1993; Liden, Stillwell, & Ferris, 1996). Such inconsistency in existing research is likely for a number of factors. Support for or against age bias may differ according to the criterion variables used, or may be a function of several contextual variables (e.g., type of job, age of subjects, amount of information available about the target workers evaluated).

Much of the research on age discrimination has been guided by congruency models. These models suggest that the relative *fit*—the match between an applicant's or worker's age and the specific knowledge, skills, abilities, and other worker characteristics viewed as necessary for a job—is responsible, in part, for the evaluations of applicants and workers (Arvey, 1979; Heilman, 1983). Such models are based on the notion that when stereotypes regarding the behaviors or personality characteristics of specific groups (e.g., female workers, older workers) do not match the perceived requirements for a job, bias is more likely to occur. Consistent with much of the gender-discrimination research that has supported Heilman's lack-of-fit model (Heilman, 1984; Heilman, Martell, & Simon, 1988), when a relative match exists between stereotyped traits, behaviors, and a job, evaluations tend to be more positive. Given the variety of traits (both positive and negative) associated with age, it is likely that the results of specific studies will be dependent on the perceived stereotypicality of an applicant or incumbent, whether or not the activated stereotype matches perceived job requirements and, lastly, whether or not stereotype application occurs.

It is well known that meta-analytic procedures can be used to summarize such diverse sets of findings across studies, as well as to help search for potential moderators of existing relationships (Glass, McGaw, & Smith, 1981; Hunter & Schmidt, 1990; Linn & Dunbar, 1986). In fact, two meta-analyses have been published previously aggregating the results of independent studies that examined age bias in work settings. The first, conducted by Kite and Johnson (1988), summarized studies dealing with attitudes as well as decision and appraisal outcomes toward older workers. While an overall effect size for studies examining such variables within work-related settings was modest (the mean weighted effect size, *d*, was .19, indicating that older workers were evaluated less positively than were younger workers), there were only 11 studies included in this analysis.[3B] Although their review included only studies conducted prior to 1981, there was some evidence that field studies produced different findings compared to laboratory studies. Laboratory-based studies did not show differential evaluations according to age, whereas field studies exhibited more negative evaluations of older workers.

B

---

[3]Kite and Johnson's (1988) results also produced evidence that there were moderators of the relationship observed; that is, while older workers were evaluated less positively, this was not a unitary phenomenon.

Finkelstein, Burke, and Raju published a second meta-analysis of age bias and discrimination in 1995. This review was restricted to data collected from simulated employment contexts. An estimate of the overall weighted effect size (d) in this review was .29.[4] However, their analysis revealed that age bias was moderated by a number of variables, including the age of the evaluators, whether positive or negative information was available about the rated workers, and whether older and younger target workers were evaluated by the same evaluator (a within-subject design) or whether only a relatively young or old worker was evaluated (a between-subjects design) by an evaluator.

There are reasons to be cautious regarding the outcomes of Finkelstein et al.'s (1995) analyses. The most important limitation is that it included only laboratory studies (9 studies, 15 independent samples). The limitations of studies conducted in laboratory settings are well recognized (Sackett & Larson, 1990). Laboratory studies often place great reliance on proxy subjects (e.g., students), as well as on artificial stimulus simulations and decision tasks. These study characteristics may produce results substantially different from those observed in more realistic settings. In addition, the moderator analyses conducted by Finkelstein et al. included estimates derived from sometimes as few as two effect sizes. Caution should be exercised because of the limited generalizability of their results to other than laboratory settings as well as the relatively small number of effect sizes used to produce category estimates in some of their moderator analyses.

Purpose of the Current Review

The purpose of the current study is to conduct a comprehensive review of research examining age bias in laboratory and field settings. Age discrimination in the workplace continues to be an important issue, as evidenced by the number of age-discrimination cases that continue to be filed with the Equal Employment Opportunity Commission (EEOC). A fact sheet distributed by the government's Administration on Aging (2002) notes that, on average, 17,000 workers brought age-discrimination complaints to the EEOC each year from 1991 to 1995. In addition, as the overall age of workers continues to increase, this issue is likely to take on increased importance and visibility. On the basis of 1995 Census data (U.S. Bureau of the Census, 1996), the government estimates there will be a

---

[4]This estimate is based on providing as comparable a data set as possible among the current effect sizes and transforming the overall weighted effect size (r) in Finkelstein et al. (1995) to a d of .29 in the present investigation. The M_r was not indicated in Finkelstein et al. However, computations based on the moderational assessments found in Table 1 in their article show the overall M_r to be approximately .17. Using DSTAT (Johnson, 1993) for the computations, the M_r in the present studies restricted to the 15 independent effect sizes from Finkelstein et al. was .16.

66.9% population increase among the 55- 64-year age group between the years of 1995 to 2010.

Given that the central goal of Kite and Johnson's (1988) review was not[C] the assessment of age bias in employment settings, coupled with the small number of relevant effect sizes in that review and in the more contemporary investigation by Finkelstein et al. (1995), the current meta-analysis would appear to be warranted and important in order to provide a more comprehensive evaluation of the research on age bias in this domain. As such, we have included studies that span four decades (1963-1998) involving field and laboratory research designs to produce a more current assessment of age bias. In addition, the present investigation provides a more comprehensive assessment of the potential moderators of age bias, compared to these earlier studies.

On the basis of prior research and the general stereotypes that exist regarding older workers, we expect an overall age effect to be manifested over the population of studies we examine (i.e., more negative evaluations of older applicants and workers), but we predict that such an aggregate effect will be minor to moderate in size. Moreover, based on prior research and theory, there are a number of research design and contextual variables that we examine as potential moderators of any observed age bias. In addition, a set of moderator variables that examine the issue of generalizing from less ecologically valid data (e.g., undergraduates being asked to make hiring recommendations based on a modicum of information) to more ecologically valid data (e.g., workers' and supervisors' personnel decisions) is assessed.

<div align="center">Moderator Variables</div>

Date of Publication

Across the entire range of studies in their review (1960-1986), Kite and Johnson (1988) found a negative relationship between publication date and age bias. As these authors suggested, the aging population and workforce have increased awareness and attention to the problems faced by the elderly. This increased awareness might play a role in reducing prejudice. In addition, as the workforce has aged, so too have many of the individuals responsible for evaluating job applicants and workers in applied settings.

Finkelstein et al. (1995) found some evidence for an in-group bias in their review that examined bias in simulated employment contexts. Their analyses revealed that younger raters provided more positive evaluations of younger workers than of older workers in the areas of overall job qualifications and potential for development. Older raters showed no difference in their evaluations of younger and older workers. On the basis of all these findings, we expect to find a negative linear relationship between the year of publication and the degree of bias against older applicants and workers.

Type of Stimulus Persons

Age-discrimination research has investigated bias in selection and performance appraisal processes. Therefore, it seemed fitting to compare the degree of bias directed toward these two different categories of people (job applicants vs. job incumbents). On the basis of the amount of information that an evaluator is likely to have regarding applicants versus workers and the likelihood that less information will be available about job applicants (especially in field settings), we expect to find a greater degree of bias in studies that evaluated job applicants than in studies that involved judgments of job incumbents.

Type of Dependent Variable

Findings from studies that have examined age stereotypes, along with some of the early research by Rosen and Jerdee (1976a, 1976b, 1979), suggest that age stereotypes of older applicants/workers are not uniformly negative and depend greatly on the particular dependent variable being examined. Based on this research and the findings from Finkelstein et al. (1995), we expect the type of dependent variable to significantly moderate effect size.

The most commonly used dependent measures in the literature have focused on overall evaluations (e.g., hiring recommendations, promotions, salary increases); however, a relatively large number of investigations have also included more specific evaluations that have been tied to age stereotypes (e.g., potential for development, interpersonal skills, and the reliability and stability of the worker). Findings from Finkelstein et al. (1995) revealed more positive evaluations of younger workers in the areas of potential for development and job qualifications. However, they also found significantly more positive evaluations for older workers on the dimension of stability.

Type of Research Design

It is likely that the type of research design (between-subjects vs. within-subject) will moderate observed effect sizes. Kite and Johnson (1988) and Finkelstein et al. (1995) suggested that age bias would be more likely observed in within-subject designs where decision-makers/evaluators review multiple targets with different age levels. These authors suggested that this would occur as a function of worker age becoming comparatively more salient, compared to between-subject studies where participants would evaluate either a younger or an older worker. Indeed, Kite and Johnson's as well as Finkelstein et al's analyses revealed significantly more bias against older workers in studies that employed a within-subject design. However, we believe that these results may be an artifact of the limited number of effect sizes (11 and 15, respectively) included in the two previous reviews.

Examination of prejudicial attitudes or discriminatory behavior has always been problematic as a function of participants wanting to appear "good" to themselves and to others (Tourangeau, Rips, & Rasinski, 2000). The proliferation of research on the development and empirical investigation of implicit measures of prejudice speaks to the general consensus within the discipline regarding respondents' desire to manage an impression of equity (Greenwald et al., 2002). We agree with the previous reviewers that asking participants to provide an overt comparison should increase the saliency of worker age. However, we believe that this increased saliency should lead to greater social desirability bias and, thus, a greater degree of age bias in between-subjects as opposed to within-subject designs. It should be mentioned that a similar moderator assessment in a recent investigation of gender bias (that was also limited to simulated employment contexts) failed to reveal type of design (between-subjects vs. within-subject) as a significant moderator (Davison & Burke, 2000).

Generalizability I: Type of Participant

In their qualitative review of the literature, Arvey and Faley (1988) noted that the age of the evaluator seemed to play a role in whether or not age bias was observed. Stereotype research has revealed that perceivers categorize and evaluate in-group and out-group members differentially, and that this often depends on specific characteristics of the evaluator (Zarate & Smith, 1990).

There is some limited evidence that the age of the evaluator/decision-maker impacts the degree of age bias observed. Offering the theory of in-group versus out-group bias, Finkelstein et al. (1995) showed that younger raters demonstrated age bias, whereas older raters did not, but these findings were again based on a limited number of studies and only included laboratory studies. In addition, a study that examined beliefs about older workers revealed that older workers held more positive beliefs about older workers than did younger employees (Hassell & Perrewé, 1995).

Making comparisons across the most common groups of participants (undergraduate students, graduate students, and supervisors) allowed for a general assessment of the in-group/out-group variable in the present data. Consistent with findings from previous research, we expect to find greater bias from undergraduate and graduate evaluators than from supervisors.

Generalizability II: Type of Research Setting

We expect that the type of research setting will moderate the effect sizes observed. As mentioned earlier, our estimate of the overall effect size in Finkelstein et al. (1995) was .29. This effect size is based on laboratory data only. Kite and Johnson (1988) found considerable differences between the effect sizes

observed under work-related field studies (mean $d = .40$) and work-related laboratory studies (mean $d = .07$). However, it should be noted that the mean d for the work-related field studies, noted above[D], was based on only three effect size estimates from the 1970s.

In addition, a recent meta-analytic review that examined gender bias in actual work settings—where a number of confounding variables including organizational level, experience, and education were controlled—found little evidence for gender bias in performance appraisals in nonconfounded field studies (Bowen, Swim, & Jacobs, 2000). Given the significantly larger group of current field studies included in the present review and the estimated overall effect size from the laboratory studies reviewed by Finkelstein et al. (1995), we predict that effect sizes will be relatively larger under laboratory conditions as a result of the relative isolation of the age variable, compared to the complexities involved in real-world settings.

Generalizability III: Type of Stimulus Person Presentation

On the basis of the previous age stereotype research, Kite and Johnson (1988) expected to find, and did indeed discover, greater age bias in studies that employed general information (e.g., "old person") as a written target label than in studies that provided specific information about the target person. Given that the studies included in the present review range from investigations that involved the use of the generic "old person" category (i.e., the stimulus person was simply noted to be of a given age) to résumé studies that provided additional individuating information to videotaped presentations of applicants/workers to the evaluation of actual applicants or workers, we believe that this assessment is also of interest in the present review. Consistent with findings by Kite and Johnson (1988), we predict a greater degree of bias in studies that employ more general than specific stimulus person information.

Generalizability IV and V: Amount and Type of Stimulus Person of Information and Amount and Type of Job Information[E]

Speaking about gender bias and stereotypes, Nieva and Gutek (1980) stated, "The more task-related information about the 'evaluatee' and the greater the clarity about the criteria to be used in the evaluation situation, the less likely it is that 'actuarial prejudice' will operate" (p. 273). There is previous evidence that when information is presented about a particular stimulus target in addition to age (or other isolated protected group characteristic: gender, race, etc.)—such as relative competence, educational and skill background, work history, and so forth—age bias is diminished. In the context of gender bias, Tosi and Einbender (1985) reviewed 21 studies investigating gender bias and showed that judges faced with

limited information about candidate competence or job requirements tended to make more biased or stereotyped judgments, whereas those with more information did not.

We predict, therefore, that studies that include more information about stimulus targets will reveal less age bias, compared to studies that include less information. Similarly, we expect that studies that include relatively more information about the particular job in question (e.g., job description, desired competencies) will reflect less bias against older workers, compared to studies that contain relatively little information about the job.

Method

Studies Sampled

The procedure used to obtain the final sample of studies in our meta-analysis was based on techniques described by Cooper and Hedges (1994), Glass et al. (1981), Hedges and Olkin (1985), Hunter, Schmidt, and Jackson (1982), and Rosenthal (1984). Computer searches were conducted using a variety of sources including the PsycINFO database (Psychological Abstracts), the ERIC database, and the World Business Abstracts (WBA). Each of these databases was searched from their inception through 1999. A variety of keyword searches were conducted for title and abstract. These included age stereotype*, age bias, age discrimination, older worker*, and older applicant*.[5F] Each of the keyword searches listed above[G] was conducted searching for the term appearing contiguously and separately in the title or abstract.

F
G

Inclusion Criteria

The inclusion criteria used in our analysis were significantly broader than those employed by Finkelstein et al. (1995). For a study to be included in the present investigation, the research: (a) had to involve a comparison of subjects' judgments of at least two different age groups of job applicants or job incumbents; (b) the evaluations provided had to focus on some aspect of hiring, promotion, or performance appraisal in a work context; and (c) the article had to include the appropriate statistical information to allow for the computation of effect sizes. The search process resulted in the identification of 39 articles that met the inclusion criteria. As a function of a number of publications reporting the results of more than one study or analyzing the data from separate subgroups, 52 independent samples of subjects were found across the 39 articles. As a function of the inclusion criteria, only 6 of the 11 relevant samples from Kite and Johnson

[5]The asterisks at the end of the keywords denote the use of wildcards in the search process.

(1988) are included; all of the samples found in Finkelstein et al. are included in the present data set.

In order to capitalize on the extent to which a study included more than one type of dependent measure (e.g., studies in which evaluators/decision-makers supplied evaluations of overall job qualifications, potential for development, and applicant/incumbent suitability), separate effect sizes were computed for each of the relevant dependent variables within a single study. To address the issue of dependency across effect sizes, Cooper's (1998) shifting unit of analysis strategy was used. This approach involves an initial coding of each relevant statistical test within a study as an independent event. As in the example referred to above[H], a single group of subjects' hiring decisions and assessment of the applicant's potential for development would be coded as two separate effect sizes. These two effect sizes were subsequently used in the categorical moderator analysis for dependent variable. However, for the assessment of an overall effect size, the two effect sizes referred to in the example above[I] were integrated (combined) so that the study would contribute only one independent effect size for the overall analysis.

Variables Coded From Each Study

The following variables were coded to allow for the examination of factors that may significantly moderate age bias and discrimination: publication date of the study, type of stimulus person (applicant vs. incumbent), type of dependent variable (overall evaluation, interpersonal skills, potential for development, and stability), and type of research design (between-subjects vs. within-subject). The impact of generalizability on effect size was examined by coding each study on the following five scales: type of research participant, type of research setting, type of stimulus person presentation, the amount of information available to participants about the stimulus person, and the amount of job-relevant information provided to participants. A more specific description of each of these five variables, along with the coding scales, is presented in Table 1. Data from each study were coded by two raters. The interrater agreement across the moderator variables for the 52 independent samples was .84. An additional rater determined final coding decisions on any discrepancies.

Analysis of Effect Sizes

The effect size used in the study was d, which, in the present analysis is equivalent to the differences between the experimental (younger worker) and the control (older worker) means divided by a pooled standard deviation. Each of the 52 independent effect sizes included in the review was converted to d and subsequently analyzed with DSAT, a meta-analytic, commercially available software

Table 1

Coding of Generalizability Moderator Variables[AG]

Variable	Potential range of values
Participant	1 = undergraduate students (low work experience) 2 = graduate students (low to moderate work experience) 3 = supervisors/managers who are actively involved in selection or performance appraisal (high work experience)
Research setting	1 = laboratory study/experiment (low consequences) 2 = training (assessment center data or training exercise) 3 = actual selection or performance ratings
Stimulus person	1 = generalized perceptions of younger/older workers (stereotype presentation studies) 2 = résumé studies (with or without pictures) 3 = videotape stimulus persons 4 = actual job applicants/incumbents
Amount and type of job information	1 = no job information 2 = job title or job description 3 = job title and description 4 = specific information on job duties, tasks, knowledge, and skills required for successful job performance
Amount and type of stimulus person information	1 = age only 2 = age and other demographic work history 3 = age, demographic information, work history, and general performance information 4 = age demographic information, work history, and specific (detailed) performance information

package (Johnson, 1993). Effect sizes are positive when the evaluations/decisions directed toward the older target are negative. Each effect size was corrected for bias as a result of sample size, without which there is overestimation of the population effect size (Hedges, 1981).

Hedges (1982) indicated that while a pooled estimate of an effect size provides a summary of the results of a series of independent studies, this pooled estimate could be misleading when the effect sizes are not homogeneous. A homogeneity statistic, Q_w, was computed that compared how much variance within a particular distribution of effect sizes is explained by sampling error. The statistic has an approximate chi-square distribution with $m - 1$ degrees of

freedom, where *m* is equivalent to the number of effect sizes within each distribution examined. A significant Q_w statistic is indicative of heterogeneity among a specific set of effect sizes. In other words, there are potential moderator variables operating such that the variance observed cannot be explained by sampling error alone. Kite and Johnson (1988) also employed this method when examining effect sizes for moderator variables, [J]as well as other authors conducting meta-analyses in other psychosocial areas (e.g., Swim, Borgida, & Maruyama[K], 1989). Moreover, a recent Monte Carlo study that compared three commonly employed meta-analytic techniques for the assessment of moderators revealed that the approach based on Hedges and Olkin (1995) provided the most accurate results for moderator estimation (Steel & Kammeyer-Mueller, 2002).

Results

Overall Effect Size

The mean effect size across the 52 independent samples corrected for sample size bias was .11 with $Q_w(51) = 353.64$, $p < .0001$.[6] This positive effect size indicates more positive evaluations for younger workers or applicants. This relationship was based on a total [L]*n* of 28,344 subjects. The 95% confidence interval for this statistic ranged from .09 to .12, and the individual effect sizes ranged from -0.44 to 1.98.[7M]

Three conclusions may be made based on these data. First, while the mean effect size across all studies was significant and positive, the size of the effect was small, using Cohen's (1988) rule of thumb that *d* values of .20 represent a relatively small effect size. Second, the mean effect size observed here was less than that observed by Kite and Johnson (1988), who reported a mean *d* value of .19 for work-related studies. Similarly, while Finkelstein et al. (1995) did not report an overall mean *d* value, on the basis of restricting the analysis to the 15 independent samples included in their review, we found a mean *d* value of .29, which is significantly larger than the value we obtained here. A third element is that the homogeneity index (Q_w) was significant, indicating the probable presence of moderating variables.

---

[6]To achieve homogeneity across the effect sizes, an analysis of outliers was conducted (Hedges, 1987; Lipsey & Wilson, 2001). In the present study, 19 of the 52 independent effect sizes had to be eliminated to achieve homogeneity, $Q_w(32) = 47.673$, $p < .068$. Doing so did not produce a meaningful difference in the mean weighted effect size ($d = .09$) or accompanying confidence interval (.07 to .11). Thus, all effect sizes were included across the moderator analyses.

[7]Given that the five samples found in Avolio, Waldman, and McDaniel (1986) comprise a significant proportion of the total number of subjects included in the review, all analyses including all the categorical and continuous moderator analyses were also conducted without these five samples. The results of all analyses without the five samples led to the same inferential conclusions.

Table 2

Tests of Categorical Models for Age Discrimination Effect Sizes

AH Variable and class	Between-class effect (Q_B)	n	Weighted effect size $(d_{i+}{}^{AH})$	95% CI for d (lower/ upper)	Homo-geneity within each class $(Q_{Wi})^a$
Type of stimulus					
Persons	9.56**				
Applicant		21	.20	.14/.26	87.09**
Incumbent		31	.10	.08/.12	251.08**
Dependent variable	480.13**b				
Overall evaluations		45	.10	.08/.12	645.95**
Potential for development		17	.45	.39/.51	266.94**
Interpersonal skills		5	.11	-.01/.22	26.18**
Stability		9	-.67	-.75/-.59	84.73**

Note. Effect sizes are positive for more favorable evaluations of younger job applicants or workers. CI = confidence interval.

[a]Significance indicates rejection of the homogeneity hypothesis. [b]All contrasts across the four dependent variable categories were significantly different at $p < .0001$, with the exception of the job qualifications and interpersonal skills comparison which was non-significant ($p > .90$).

**$p < .001$.

Moderator Analyses

Date of publication. Examination of the relationship between year of publication and effect size was obtained by conducting a continuous moderator analysis using Rosenthal and Rubin's (1982) focused-comparison technique. This continuous model technique provides an estimate of the linear relationship between the predictor variable (year of publication) and effect size. As previously mentioned, the studies spanned the years of 1963 to 1998. The results indicated a significant negative relationship between effect size and year of publication (z for model = -5.48, $p < .01$, one-tailed). Thus, as predicted, less age bias was found in more recent research studies.

Type of stimulus person(s). Table 2 lists the results of examining age bias as a function of the type of person(s) being evaluated (applicants vs. incumbents).

Consistent with our predictions, job applicants received significantly more negative evaluations (mean d = .20) than did job incumbents (mean d = .10), $Q_B(1)$ = 9.56, p < .002. A similar outcome was revealed limiting the analysis to the 33 effect sizes from laboratory studies. Once again, applicants received significantly more negative evaluations (mean d = .29) than did incumbents (mean d = .12), $Q_B(1)$ = 12.76, p < .0004. Given the small number of field studies that evaluated job applicants (n = 2), a similar comparison could not be made among the field research.

Type of dependent variable. The moderator analysis for the type of dependent variable is shown in Table 2. The specific type of evaluation being made significantly moderated effect size, $Q_B(3)$ = 480.13, p < .0001. Significant positive effect sizes (ps < .05) were found on the overall evaluation (mean d = .10) and potential for development (mean d = .45) measures, where younger workers were evaluated more positively, whereas older workers were evaluated more positively than younger workers in the area of stability (mean d = -.67). While no significant difference was found between ratings of older and younger workers on the assessment of interpersonal skills (mean d = .10), this relationship was marginally significant (p < .10). A series of contrasts across the four categories revealed significant differences between each of the four dependent variable categories, all $\chi^2(1^N)$ > 25.76, all ps < .0001, with the exception of the comparison between the categories of overall evaluation and the interpersonal skills, $\chi^2(1^O)$ = .02, p > .98. Notably, even the group of five effect sizes that comprised the category of interpersonal skills revealed significant heterogeneity, as did the other three categories in this analysis.

Type of research design. To provide a thorough comparison between this assessment and a similar analysis conducted by Finkelstein et al. (1995), a set of three related moderator analyses was conducted. This involved making the between-subjects/within-subject comparison across all 52 effect sizes and the following two subsets: the 33 effect sizes from laboratory research in the present review, and the 15 effect sizes that comprised Finkelstein et al.'s (1995) review. The results are summarized in Table 3. Consistent with our predictions, older workers received more negative evaluations in between-subjects designs (mean d = .30) than in within-subject designs (mean d = .10), $Q_B(1)$ = 20.90, p < .001. Even when constraining our analyses to studies that were conducted in laboratory settings (33 independent effect sizes), the mean d for between-subjects age-manipulation studies was significantly larger (mean d = .31) than that produced from within-subject designs (mean d = .14), $Q_B(1)$ = 20.87, p < .001.

Given the possibility that the outcome for the analysis in the present investigation might be attributed to the specific meta-analytic technique used for the moderator analysis (cf. Johnson, Mullen, & Salas, 1995; Schmidt & Hunter, 1999), the same analysis was conducted using the 15 effect sizes examined in the previous meta-analytic investigation by Finkelstein et al. (1995). On the basis of

Table 3

Tests of Categorical Models for Age Discrimination Effect Sizes: Between-Subjects Versus Within-Subject Designs

Variable and class	Between-class effect (Q_B)	n	Weighted effect size (d_{i+}^{AI})	95% CI for d (lower/ upper)	Homo-geneity within each class (Q_{Wi})[a]
All studies					
Design	20.90**				
Between-subjects		18	.30	.21/.38	52.46**
Within-subject		34	.10	.08/.12	274.37**
Laboratory studies					
Design	20.87**				
Between-subjects		15	.31	.22/.39	45.27**
Within-subject		18	.14	.09/.20	157.44**
Finkelstein et al. (1995)					
Design	11.34**				
Between-subjects		6	.11	-.03/.26	11.03
Within-subject		9	.40	.31/.49	28.70**

Note. Effect sizes are positive for more favorable evaluations of younger job applicants or workers. CI = confidence interval.
[a]Significance indicates rejection of the homogeneity hypothesis.
*$p < .01$. **$p < .001$.

these 15 effect sizes, the results of the moderator analysis were consistent with those of Finkelstein et al.—a significantly larger mean effect size for within-subject age manipulations (mean $d = .40$), compared to data derived from between-subjects manipulations (mean $d = .11$), $Q_B(1) = 11.34$, $p < .001$. Therefore, it would appear that the conclusions derived from their meta-analytic investigation might have been a function of the restricted range of studies used.

Generalizability I: Type of participant. The results of the five categorical generalizability analyses appear in Table 4. As predicted, the type of research participant significantly moderated the degree of age bias, $Q_B(2) = 480.13$, $p < .0001$. Findings were similar to those found in the comparison of younger versus older raters by Finkelstein et al. (1995). Undergraduate student (mean $d = .31$)

Table 4

Tests of Categorical Models for Generalizability Scales

AJ

Variable and class	Between-class effect (Q_B)	n	Weighted effect size (d_{i+}^{AJ})	95% CI for d (lower/ upper)	Homogeneity within each class $(Q_{Wi})^a$
Participants[b]	39.56**				
Undergraduate students		19	.31	.23/.38	97.09**
Graduate students		6	.38	.23/.53	8.47
Supervisors		26	.10	.08/.12	144.63**
Setting	20.87**				
Laboratory		33	.19	.14/.24	212.61**
Training		8	-.10	-.26/.04	69.20**
Field		11	.10	.08/.12	45.05**
Stimulus persons	25.57**				
Generic		7	.05	-.03/.13	131.83**
Résumé		25	.26	.20/.33	69.93**
Videotape		7	.21	.06/.36	20.78**
Job incumbents		13	.10	.08/.12	99.62**
Amount and type of stimulus information	20.97**				
Age only		10	.09	.02/.17	165.24**
Age and other demographic information or work history		13	.29	.21/.37	25.69*
Age, demographic information, work history, and general performance information		17	.08	.01/.14	95.15**
Age, demographic information, work history, and detailed performance information		12	.10	.08/.12	40.48**

(*table continues*)

Table 4 (Continued)

Variable and class	Between-class effect (Q_B)	n	Weighted effect size (d_{i+})	95% CI for d (lower/ upper)	Homogeneity within each class $(Q_{Wi})^a$
Amount and type of job information	39.16**				
No job information		6	.46	.33/.59	13.22*
Job title or description		16	.09	.03/.16	141.84**
Job title and description		16	.23	.14/.32	48.88**
Specific KSAOS[AK]		14	.10	.08/.11	104.14**

Note. Effect sizes are positive for more favorable evaluations of younger job applicants or workers. CI = confidence interval.

[a]Significance indicates rejection of the homogeneity hypothesis. [AL][b]The number of effect sizes totals 51 for the type of research participants because only one of the samples involved supervisee ratings.

*p < .01. **p < .001.

[AK]

[AL]

and graduate student (mean d = .38) participants made significantly more negative evaluations of older workers than did supervisors (mean d = .10), both $\chi^2(1^P) > 16.31$, $ps < .003$. Ratings from undergraduate and graduate students were not significantly different, $\chi^2(1^Q) = 2.25$, $p > .69$.

[P]

[Q]

Generalizability II: Type of research setting. Table 4 also includes the results of the analysis based on the type of research setting within which the study was conducted: laboratory, field, or training (assessment center) settings. This categorical moderator analysis revealed a significant effect, $Q_B(2) = 20.87$, $p < .0001$. The mean d for field settings was .10, whereas the mean d for laboratory studies was .19. These values were significantly different from one another, $\chi^2(1^R) = 13.13$, $p < .002$, confirming our prediction. The -.10 mean d found for the eight effect sizes from the assessment center studies was not significantly different from zero. However, this mean d was significantly different from both the laboratory, $\chi^2(1^S) = 13.60$, $p < .002$; and field research, $\chi^2(1^T) = 6.89$, $p < .032$.

[R]

[S/T]

Generalizability III: Type of stimulus person presentation. Analysis of the data as a function of stimulus person presentation showed a significant effect, $Q_B(3) = 25.57$, $p < .0001$. The use of résumé formats for stimulus materials resulted in significantly more negative evaluations of older workers (mean d = .26) than did the use of a videotape format (mean d = .21). Somewhat surprisingly, when generalized perceptions of older or younger workers were asked for

in studies, the age bias observed was nonsignificant (mean $d = .05$). Comparisons across the four categories revealed only two significant differences. The résumé stimulus person category showed significantly greater bias than did the generic category, $\chi^2(1^U) = 15.38$, $p < .002$; and, consistent with the generalizability hypothesis, use of résumé materials showed significantly greater bias than data from studies that involved the evaluation of actual workers, $\chi^2(1^V) = 22.04$, $p < .0001$.

Generalizability IV and V: Amount and type of stimulus person of information and amount and type of job information. The results concerning the amount and type of information about the stimulus target revealed a significant between-group effect, $Q_B(3) = 39.16$, $p < .0001$, but were more complex. As expected, studies that had more limited information about the target (age and other demographic information or work history) showed the most bias (mean $d = .29$), compared to studies that presented more performance information (.08 and .10 for the two performance information conditions), both $\chi^2(1^W) > 16.22$, $ps < .001$. However, contrary to our predictions, studies that provided only information regarding the age of the target worker showed significantly less bias than did the age and other demographic information or work history category, $\chi^2(1^X) = 12.13$, $p < .007$. We expected more bias to be demonstrated under this condition.

Finally, the amount and type of information about the job also moderated age bias, $Q_B(3) = 20.97$, $p < .0001$. Once again, there was mixed support for our predictions. As expected, when no information about the job was presented or known to participants, fairly pronounced age bias was observed (mean $d = .46$). This level of bias was significantly greater than in any of the three higher information categories, $\chi^2(1^Y) > 8.82$, $ps < .032$. Conversely, the high information category (when participants are provided or have information about the specific knowledge, skills, and abilities related to job performance) revealed significantly less age bias (mean $d = .10$) than did the job title and description category (mean $d = .23$), $\chi^2(1^Z) = 8.01$, $p < .045$. None of the other contrasts between categories were significant.

Inasmuch as there are undoubtedly varying degrees of overlap across the many categories of the five generalizability scales, a final assessment that involved the development of a composite measure based on summing across the various scale values listed in Table 1 was conducted. The possible range of scores on this composite measure could range from 5 (*low generalizability*) to 22 (*high generalizability*[8]). Use of Rosenthal and Rubin's (1982) focused-comparison technique revealed a significant negative linear relationship between the relative generalizability of the data and effect size (*z* for model = -8.15, *M* = 14,

---

[8]The discrepancy between this upper scale limit (22) and the number of levels within categories found in Table 4 (18) was a result of there being less than three effect sizes comprising a single level across a number of the generalizability scales. This precluded analysis of those levels.

$p < .0001$). Consistent with findings from the related categorical analyses, as the overall generalizability of the data increased, age bias decreased.

Discussion

The overall effect size in the study demonstrates a small, yet statistically significant mean effect for age bias in the research literature. The inclusion of additional relevant research and a wider range of moderator variables in the present investigation allowed for a more comprehensive examination of the age bias and discrimination research literature than in previous reviews by Finkelstein et al. (1995) and Kite and Johnson (1988). Evidence for this can be seen in the results of the analysis that examined the impact of type of design used (between-subjects vs. within-subject) on age bias. The differential outcomes found in Finkelstein et al. and the present study suggest that the previous relationship documented in Finkelstein et al. may well have been a function of the limited number of independent effect sizes included in their analysis of this moderator variable (15 vs. 52 in the present investigation).

The relatively small effect size found for age bias in the present review (especially in the more current data) suggests that age bias may actually be less of a problem today than it was in previous decades. However, a large number of variables were shown to moderate this effect. For example, the moderator assessment conducted on the type of evaluation or judgments (dependent variable) being made suggests that age stereotypes are likely to have a variety of components that lead to positive evaluations for older employees (older applicants/workers are more stable and reliable), negative evaluations for older employees (older applicants/workers are less likely to adapt to new tasks, situations; they will be resistant to change), and neutral evaluations of older employees (older applicants/workers have similar interpersonal skills to younger applicants/workers). These findings suggest that the fit between various aspects of age stereotypes and specific job demands might affect the degree of age bias (e.g., greater bias may occur when older workers are perceived as less adaptive and they are being evaluated for a job that demands the ability to adapt to changing situations).

Results from the moderator analyses also help substantiate an ongoing concern among applied psychologists regarding the relative generalizability of data from proxy subjects who are asked to make evaluations in the absence of appropriate information on stimulus persons, job-relevant tasks and behaviors, and with little or no experience in making such judgments within organizational settings. It should be reiterated, however, that it was the magnitude of the effect that did not generalize from lab to field; not the effect itself. Much of the current social psychological research on stereotypes suggests that the activation of stereotypes occurs automatically (Devine, 1989; Gilbert & Hixon, 1991). The results from this meta-analysis suggest that the degree of application and reliance

on such stereotypes is likely to be dependent on the amount of applicant/incumbent and job-relevant information that a perceiver possesses and subsequently utilizes in forming judgments and evaluations.

For the most part, the results from analyses that examined the issue of generalizability show that greater and more relevant information and greater and more relevant experience among raters, judges, or supervisors leads to less age bias. The one exception to this general trend involved one category within the moderator analysis that examined amount and type of stimulus person information. This analysis showed less age bias among subjects who were asked to make judgments based on the age of the applicant/worker alone. It is possible that the non-significant mean *d* based on data from studies that asked for generalized impressions of workers (i.e., in these studies participants received only the applicant's or job incumbent's age) might be a function of such perceivers becoming more cautious in their ratings inasmuch as they have no other demographic information and no performance information to utilize in forming judgments.

Given the fact that the category that involved stimulus person age and one other piece of information (the demographic information or work history category) produced the greatest amount of bias, one might be tempted to say "a little knowledge may be a dangerous thing." A somewhat related finding has been seen in the impression-formation literature where subjects form impressions and judgments readily and, in a more biased fashion, when a stereotype has been activated and subjects are provided a small amount of additional information, even if such information is redundant (Tversky & Kahneman, 1983). This one result notwithstanding, the overall trend for the type and amount of applicant/worker information moderator assessments suggests that as relevant information increases, age bias is reduced.

Finally, the results of the continuous moderator analysis that demonstrated a negative linear relationship between publication date and effect size are consistent with data from Kite and Johnson (1988). The observed reduction of age bias seen in the more recent research literature is likely to be a result of a variety of factors. As suggested by Kite and Johnson, this change may be a result of changes in the application of age stereotypes in the workplace. The impact of EEOC guidelines in the areas of selection and performance appraisal may be a key factor in the observed decrease in age bias found in the present data.

It is also possible that the reduced degree of bias found in more current data might reflect improved performance among more current cohorts of older workers. An examination of age and performance differences on cognitive, perceptual, and psychomotor skills suggests that age and gender accounted for a relatively small proportion of variability in ability test scores when experience and education were controlled (Avolio & Waldman, 1994). In addition, the aging of the workforce in the United States and the documented negative relationship between age of evaluator and degree of bias found in both the current review and

the previous meta-analyses might have contributed to smaller effect sizes. It is also possible that a more concerted focus on using more relevant and realistic methods to investigate age bias (especially when this occurs in laboratory research), has led to smaller effect sizes found in the more current research.

Taken as a whole, the results of the present investigation have demonstrated that the amount of age bias found in a given study is likely to be a function of a multitude of factors, including the research design used; the setting; the types of evaluations and judgments that are made; and the relative generalizability of stimulus persons, stimulus materials, and perceivers. On the basis of the current findings, future attempts at examining age bias in laboratory and field research will need to give such factors more serious consideration when designing studies and interpreting data from such investigations.

References

Note. [AA]References marked with an asterisk (*) indicate studies included in the meta-analysis.

Administration on Aging. (2002, January). *Age discrimination: A pervasive and damaging influence*. Retrieved May 15, 2002, from http://www.aoa.gov/factsheets/ageism.html

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

Arvey, R. D., & Faley, R. H. (1988). *Fairness in selecting employees*. Reading, MA: Addison-Wesley.

*Avolio, B. J., & Barrett, G. V. (1987). Effects of age stereotyping in a simulated interview. *Psychology and Aging*, *2*, 56-63.

Avolio, B. J., & Waldman, D. A. (1994). Variations in cognitive, perceptual, and psychomotor abilities across the working life span: Examining the effects of race, sex, experience, education, and occupational type. *Journal of Applied Psychology*, *9*, 430-442.

*Avolio, B. J., Waldman, D. A., & McDaniel, M. A. (1990[AB]). Age and work performance in nonmanagerial jobs: The effects of experience and occupational type. *Academy of Management Journal*, *33*, 407-422.

Bowen, C., Swim, J. K., & Jacobs, R. R. (2000). Evaluating gender biases on actual job performance of real people: A meta-analysis. *Journal of Applied Social Psychology*, *30*, 2194-2215.

*Britton, J. O., & Thomas, K. R. (1973). Age and sex as employment variables: Views of employment service interviewers. *Journal of Employment Counseling*, *10*, 180-186.

*Cleveland, J. N., & Hollmann, G. (1990). The effects of the age-type of tasks and incumbent age composition on job perceptions. *Journal of Vocational Behavior*, *36*, 181-194.

*Cleveland, J. N., & Landy, F. J. (1981). The influence of rater and ratee age on two performance judgments. *Personnel Psychology*, *34*, 19-29.

*Cleveland, J. N., & Landy, F. J. (1983). The effects of person and job stereotypes on two personnel decisions. *Journal of Applied Psychology*, *68*, 609-619.

*Cleveland, J. N., & Shore, L. M. (1992). Self- and supervisory perspectives on age and work attitudes and performance. *Journal of Applied Psychology*, *77*, 469-484.

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

*Connor, C. L., Walsh, P., Litzelman, D. K., & Alvarez, M. G. (1978). Evaluation of job applicants: The effects of age vs. success. *Journal of Gerontology*, *33*, 246-252.

Cooper, H. (1998). *Integrating research: A guide for literature reviews*. Thousand Oaks, CA: Sage.

Cooper, H., & Hedges, L. V. (1994). *The handbook of research synthesis*. New York, NY: Russell Sage.

*Craft, J. A., Doctors, S. I., Shkop, Y. M., & Benecki, T. J. (1979). Simulated management perceptions, hiring decisions, and age. *Aging and Work*, *2*, 95-102.

*Crew, J. C. (1984). Age stereotypes as a function of race. *Academy of Management Journal*, *27*, 431-435.

Davison, H. K., & Burke, M. J. (2000). Sex discrimination in simulated employment contexts: A meta-analytic investigation. *Journal of Vocational Behavior*, *56*, 225-248.

Devine, P. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5-18.

*Drehmer, D. E., Carlucci, C. A., Bordieri, J. E., & Pincus, L. B. (1992). Effects of age on ranking for reduction in work force. *Psychological Reports*, *70*, 1203-1209.

*Erber, J. T., Caiola, M. A., & Pupo, F. A. (1994). Age and forgetfulness: Managing perceivers' impressions of targets' capability. *Psychology and Aging*, *9*, 554-561.

*Ferris, G. R., Yates, V. L., Gilmore, D. C., & Rowland, K. M. (1985). The influence of subordinate age on performance ratings and causal attributions. *Personnel Psychology*, *38*, 545-557.

*Finkelstein, L. M., & Burke, M. J. (1998). Age stereotyping at work: The role of rater and contextual factors on evaluations of job applicants. *Journal of General Psychology*, *125*, 317-345.

Finkelstein, L. M., Burke, M. J., & Raju, N. S. (1995). Age discrimination in simulated employment contexts: An integrative analysis. *Journal of Applied Psychology*, *80*, 652-663.

*Fusilier, M. R., & Hitt, M. A. (1983). Effects of age, race, sex, and employment experience on students' perceptions of job applications. *Perceptual and Motor Skills*, *57*, 1127-1134.

*Gibson, K. J., Zerbe, W. J., & Franken, R. E. (1993). The influence of rater and ratee age on judgments of work-related attributes. *The Journal of Psychology*, *127*, 271-280.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypical beliefs. *Journal of Personality and Social Psychology*, *60*, 509-517.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

*Gordon, R. A., Rozelle, R. M., & Baxter, J. C. (1988). The effect of applicant age, job level, and accountability on the evaluation of job applicants. *Organizational Behavior and Human Decision Processes*, *41*, 20-33.

*Gordon, R. A., Rozelle, R. M., & Baxter, J. C. (1989). The effect of applicant age, job level, and accountability on the perceptions of female job applicants. *The Journal of Psychology*, *123*, 59-68.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*, 3-25.

*Griffeth, R. W., & Bedeian, A. G. (1989). Employee performance evaluations: Effects of ratee age, rater age, and ratee gender. *Journal of Organizational Behavior*, *10*, 83-90.

*Haefner, J. E. (1977). Race, age, sex, and competence as factors in employer selection of the disadvantaged. *Journal of Applied Psychology*, *62*, 199-202.

Hassell, B. L., & Perrewé, P. L. (1995). An examination of beliefs about older workers: Do stereotypes still exist? *Journal of Organizational Behavior*, *16*, 457-468.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*, 107-128.

Hedges, L. V. (1982). Fitting categorical models to effect sizes from a series of experiments. *Journal of Educational Statistics*, *7*, 119-137.

Hedges, L. V. (1987). How hard is hard science, how soft is soft science? The empirical cumulativeness of research. *American Psychologist*, *42*, 443-455.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York, NY: Academic Press.

Heilman, M. E. (1983). Sex bias in the work setting: The lack of fit model. *Research in Organizational Behavior*, *5*, 269-298.

Heilman, M. E. (1984). Information as a deterrent against sex discrimination: The effects of applicant sex and information type on preliminary employment decisions. *Organizational Behavior and Human Decision Processes*, *37*, 174-186.

Heilman, M. E., Martell, R. F., & Simon, M. C. (1988). The vagaries of sex bias: Conditions regulating the undervaluation, equivaluation, and overvaluation of female job applicants. *Organizational Behavior and Human Decision Processes*, *41*, 98-110.

*Hitt, M. A., & Barr, S. H. (1989). Managerial selection decision models: Examination of configural cue processing. *Journal of Applied Psychology*, *74*, 53-61.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: Cumulating findings across research*. Beverly Hills, CA: Sage.

Johnson, B. T. (1993). *DSTAT 1.10: Software for the meta-analytic review of research literatures*. Hillsdale, NJ: Lawrence Erlbaum.

Johnson, B. T., Mullen, B., & Salas, E. (1995). Comparison of three major meta-analytic approaches. *Journal of Applied Psychology, 80*, 94-106.

Kirchner, W. K., & Dunnette, M. D. (1954). Attitudes toward older workers. *Personnel Psychology, 7*, 257-265.

Kite, M. E., & Johnson, B. T. (1988). Attitudes toward older and younger adults: A meta-analysis. *Psychology and Aging, 3*, 233-244.

*Lee, J. A., & Clemmons, T. (1985). Factors affecting employment decisions about older workers. *Journal of Applied Psychology, 70*, 785-788.

*Liden, R. C., Stilwell, D., & Ferris, G. R. (1996). The effects of supervisor and subordinate age on objective performance and subjective performance ratings. *Human Relations, 49*, 327-347.

*Lin, T., Dobbins, G. H., & Farh, J. (1992). A field study of race and age similarity effects on interviewer ratings in conventional and situational interviews. *Journal of Applied Psychology, 77*, 363-371.

Linn, R. L., & Dunbar, S. B. (1986). Validity generalization and predictive bias. In R. A. Berk (Ed.), *Performance assessment: Methods and applications* (pp. 203-236). Baltimore, MD: Johns Hopkins University Press.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

*Morrow, P. C., McElroy, J. C., Stamper, B. G., & Wilson, M. A. (1990). The effects of physical attractiveness and other demographic characteristics on promotion decisions. *Journal of Management, 16*, 723-737.

Nieva, V. F., & Gutek, B. A. (1980). Sex effects on evaluation. *Academy of Management Review, 5*, 267-276.

*Perry, E. L., & Bourhis, A. C. (1998). A closer look at the role of applicant age in selection decisions. *Journal of Applied Social Psychology, 28*, 1670-1697.

*Perry, E. L., Kulik, C. T., & Bourhis, A. C. (1996). Moderating effects of personal and contextual factors in age discrimination. *Journal of Applied Psychology, 81*, 628-647.

*Perry, J. S., & Varney, T. L. (1978). College students' attitudes toward workers' competence and age. *Psychological Reports*, *42*, 1319-1322.

*Peterson, C. C. (1980). Are young people biased against older teachers? *Journal of Genetic Psychology, 136*, 309-310.

*Querishi, M. Y., & Kay, J. P. (1986). Physical attractiveness, age, and sex as determinants of reactions to résumés. *Social Behavior and Personality, 14*, 103-112.

*Rosen, B., & Jerdee, T. H. (1976a). The influence of age stereotypes on managerial decisions. *Journal of Applied Psychology, 61*, 428-432.

*Rosen, B., & Jerdee, T. H. (1976b). The nature of job-related age stereotypes. *Journal of Applied Psychology, 61*, 180-183.

*Rosen, B., & Jerdee, T. H. (1979). Influence of employee age, sex, and job status on managerial recommendations for retirement. *Academy of Management Journal, 22*, 169-173.

*Rosen, B., Jerdee, T. H., & Lunn, R. O. (1981). Effects of performance appraisal format, age, and performance level on retirement decisions. *Journal of Applied Psychology, 66*, 515-519.

Rosenthal, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage.

Rosenthal, R., & Rubin, D. B. (1982). Comparing effect sizes of independent studies. *Psychological Bulletin, 92*, 500-504.

Sackett, P. R., & Larson, J. R. (1990). Research strategies and tactics in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 419-489). Palo Alto, CA: Consulting Psychologists Press.

Schmidt, F. L., & Hunter, J. E. (1999). Comparison of three meta-analysis methods revisited: An analysis of Johnson, Mullen, and Salas (1995). *Journal of Applied Psychology, 84*, 144-148.

*Schwab, D. P., & Heneman, H. G. (1978). Age stereotyping in performance appraisal. *Journal of Applied Psychology, 63*, 573-578.

*Shore, L. M., & Bleicken, L. M. (1991). Effects of supervisor age and subordinate age on rating congruence. *Human Relations, 44*, 1093-1105.

*Singer, M. S. (1986). Age stereotypes as a function of profession. *Journal of Social Psychology, 126*, 691-692.

*Singer, M. S., & Sewell, C. (1989). Applicant age and selection interview decisions: Effect of information exposure on age discrimination in personnel selection. *Personnel Psychology, 42*, 135-154.

Steel, P. D., & Kammeyer-Mueller, J. D. (2002). Comparing meta-analytic moderator estimation techniques under realistic conditions. *Journal of Applied Psychology, 87*, 96-111.

AC [AC]Swim, J., Borgida, E., Maruyama, G., & Myers, D. G. (1989). Joan McKay versus John McKay: Do gender stereotypes bias evaluations? *Psychological Bulletin, 105*, 409-429.

Tosi, H. L., & Einbender, S. W. (1985). The effects of the type and amount of information in sex discrimination research: A meta-analysis. *Academy of Management Journal*, *28*, 712-723.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge, UK: Cambridge University Press

*Triandis, H. C. (1963). Factors affecting employee selection in two cultures. *Journal of Applied Psychology*, *47*, 89-96.

Tuckman, J., & Lorge, I. (1952). Experts' biases about the older worker. *Science*, *115*, 685-687.

Tversky, A., & Kahneman, D. T. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293-315.

AD Uniform guidelines on employee selection procedures. (1978). *Federal Register*, *43*, 38290-38309.

AE U.S. Bureau of the Census. (1996). *Table 1—Projections of the population by age and sex: 1995-2050.*

Zarate, M. A., & Smith, E. R. (1990). Person categorization and stereotyping. *Social Cognition*, *2*, 161-185.

AF6 The number effect sizes total 51 for the type of research participants because only one of the samples involved supervisee ratings.

AUTHOR QUERIES

A   The abstract is too long ("Abstracts should not exceed 120 words," APA *Publication Manual*, 2001, p. 13). Please delete some material (do not rewrite the abstract, since it has already been copy edited) to make the length more manageable.

B   See Footnote 3: I adjusted this footnote somewhat. Please make sure that it reads as you intended.

C   Shouldn't you indicate what the central goal *was*, and not what the central goal *was not*? Please reword to clarify your meaning and to avoid this awkwardness.

D   You should not use "above" to indicate location in your text. Please correct.

E   This section heading is a bit wordy. Please correct, if possible.

F   See Footnote 5: Will the term "wildcards" be widely understood? Please clarify your meaning, if necessary.

G   You should not use "above" to indicate location in your text. Please correct.

H   You should not use "above" to indicate location in your text. Please correct.

I   You should not use "above" to indicate location in your text. Please correct.

J   This sentence does not make sense. Shouldn't this read "as did other authors . . ."? Please make the necessary corrections to clarify your meaning.

K   The apparently corresponding entry in the References has an additional author (Myers). Please make the necessary corrections or provide the missing reference.

L   You should not use a statistical symbol as a word in your text.

M   See Footnote 7: 1986 or 1990? See References and provide correct date or missing reference.

N   You should provide N for this chi-square statistic. Please correct.

O   You should provide N for this chi-square statistic. Please correct.

P   You should provide N for this chi-square statistic. Please correct.

Q   You should provide N for this chi-square statistic. Please correct.

R   You should provide N for this chi-square statistic. Please correct.

S   You should provide N for this chi-square statistic. Please correct.

T   You should provide N for this chi-square statistic. Please correct.

U   You should provide N for this chi-square statistic. Please correct.

V   You should provide N for this chi-square statistic. Please correct.

W   You should provide N for this chi-square statistic. Please correct.

X   You should provide N for this chi-square statistic. Please correct.

Y   You should provide *N* for this chi-square statistic. Please correct.

Z   You should provide *N* for this chi-square statistic. Please correct.

AA   Shouldn't the studies that were in the meta-analysis also be cited somewhere in your text (28 of your reference entries are not cited in your text)? Please make any necessary corrections.

AB   1990 or 1986? See query M and provide correct date.

AC   See query K and make the necessary corrections.

AD   Not cited in your text. Please cite or delete from the References.

AE   This reference entry is incomplete. Provide city and state of publication and publisher.

AF   I moved this note to Table 4 (see query AL). Please make any necessary corrections, including deleting this note from this place.

AG   I reworded the table title. Please make sure that it reads as you intended.

AH   What does this statistical symbol mean? Please clarify your meaning.

AI   What does this statistical symbol mean? Please clarify your meaning.

AJ   What does this statistical symbol mean? Please clarify your meaning.

AK   You must spell out this abbreviation in a table note. Please correct.

AL   Is this note supposed to be here? I moved it from another place in the paper (see query AF), and it appears that it fits here. Please make the necessary corrections.

Summary of Comments on 01P0183.fm

Page: 4

Sequence number: 1
Author: Cathy
Date: 2/10/2004 12:28:21 PM
Type: Note

Insert the word "to" after "55-"