Journal of Applied Psychology
2015, Vol. 100, No. 5, 1497–1510

© 2015 American Psychological Association
0021-9010/15/$12.00 http://dx.doi.org/10.1037/a0038991

Cognitive Predictors and Age-Based Adverse Impact Among Business Executives

Rachael M. Klein
University of Minnesota and Korn Ferry, Minneapolis, Minnesota

Stephan Dilchert
Baruch College, CUNY

Deniz S. Ones
University of Minnesota

Kelly D. Dages
General Dynamics Information Technology, Chicago, Illinois

Age differences on measures of general mental ability and specific cognitive abilities were examined in 2 samples of job applicants to executive positions as well as a mix of executive/nonexecutive positions to determine which predictors might lead to age-based adverse impact in making selection and advancement decisions. Generalizability of the pattern of findings was also investigated in 2 samples from the general adult population. Age was negatively related to general mental ability, with older executives scoring lower than younger executives. For specific ability components, the direction and magnitude of age differences depended on the specific ability in question. Older executives scored higher on verbal ability, a measure most often associated with crystallized intelligence. This finding generalized across samples examined in this study. Also, consistent with findings that fluid abilities decline with age, older executives scored somewhat lower on figural reasoning than younger executives, and much lower on a letter series test of inductive reasoning. Other measures of inductive reasoning, such as Raven's Advanced Progressive Matrices, also showed similar age group mean differences across settings. Implications for employee selection and adverse impact on older job candidates are discussed.

Keywords: age, adverse impact, cognitive abilities, general mental ability, executive selection

The adverse impact potential of cognitive ability measures has been frequently examined with regard to several protected classes. Race and ethnic group differences have been most heavily researched; the Black–White mean-score difference on general mental ability (GMA) typically amounts to about one standard deviation unit in workplace settings (Roth, Bevier, Bobko, Switzer, & Tyler, 2001). Sex differences have also been examined. Although they are negligible on GMA, on specific ability measures (e.g., visual–spatial ability, quantitative ability, and technical aptitudes) sex mean-score differences exist and could potentially lead to adverse impact against women (Hedges & Nowell, 1995; Ones, Dilchert, Viswesvaran, & Salgado, 2010; Schmidt, 2011). Applied psychologists have therefore examined the effectiveness of a wide range of strategies for reducing racioethnic and sex differences (cf. Ployhart & Holtz, 2008).

In contrast, age differences on cognitive measures commonly used in personnel decisions, as well as strategies to reduce age-based adverse impact, have received much less attention in the psychological and human resource management literatures (Ones et al., 2010). Given the growing representation of older individuals in the workforces of most industrialized countries (United Nations Population Fund, 2012) and the widespread use of cognitive ability tests in personnel selection (Ryan, McFarland, Baron, & Page, 1999), it is important for organizations to have a better understanding of the potential for age-based adverse impact of cognitive ability test scores. To this end, this study examines executives' performance on GMA and specific ability tests to determine the magnitude of age differences, as well as how they might contribute to or reduce adverse impact on older individuals. We conclude by

This article was published Online First March 30, 2015.

Rachael M. Klein, Department of Psychology, University of Minnesota and Korn Ferry, Minneapolis, Minnesota; Stephan Dilchert, Department of Management, Zicklin School of Business, Baruch College, CUNY; Deniz S. Ones, Department of Psychology, University of Minnesota; Kelly D. Dages, General Dynamics Information Technology, Chicago, Illinois.

A paper based on portions of this research was presented at the annual conference of the Society for Industrial and Organizational Psychology, San Diego, CA, April 2012, and was selected for the International Personnel Assessment Council's 2013 James C. Johnson Award. Data for Samples 1 and 2 were provided by General Dynamics Information Technology. We thank John W. Jones for providing access to these unique data and for giving us full control over analyses and write-up. We also thank Kevin C. Stanek for facilitating access to data for Samples 3 and 4. This work was supported by the National Science Foundation through a Graduate Research Fellowship awarded to Rachael M. Klein and was completed while she was a graduate student at the University of Minnesota.

Correspondence concerning this article should be addressed to Rachael M. Klein, Korn Ferry, 33 South 6th Street, Suite 4900, Minneapolis, MN 55402. E-mail: rachael.klein@kornferry.com

discussing relevant strategies to reduce age-based adverse impact and implications for selection decisions.

Investigations of age-based adverse impact on work-related assessment tools are becoming increasingly important as the U.S. population ages and older employees remain in the workforce longer. Projections based on census data from 2000 indicate that the proportion of the general population age 65 and older will increase from 13% in 2010 to 19% in 2030 (U.S. Department of Health and Human Services, Administration on Aging, 2010). In the working population, the median age of employees rose from 35 in 1979 to 41 in 2008, with a median age of 42 projected for 2018 (U.S. Bureau of Labor Statistics, 2009). Employed adults also increasingly delay their retirement, in part motivated by financial reasons (Maestas & Zissimopoulos, 2010). The recent economic downturn has contributed to this trend by depleting the savings of many older individuals, causing them to either delay retirement or even reenter the workforce (Leahey, 2012).

In many countries, older individuals are identified as a protected group by antidiscrimination legislation. In the U.S., the Age Discrimination in Employment Act (ADEA) of 1967 protects employees age 40 and older from age-based employment discrimination. Coinciding with the demographic trends summarized above, there has been a rise in the number of discrimination charges filed under the ADEA. In 2011, there were 23,465 charges of age discrimination, an almost 20% increase compared to 1992 (U.S. Equal Employment Opportunity Commission, 2010).[1] Age discrimination suits can come at a large cost to organizations. Aggrieved employees have a 78% success rate at state and local jury trials, and the median age discrimination verdict is $300,000 in federal district courts, the highest amount for all types of discrimination (Grossman, 2003). Many of these costly and time-consuming trials could potentially be avoided by greater awareness of how selection measures can adversely impact older employees.

Considering the rising number of older employees in the workforce, as well as the rise in age-based discrimination cases, it is increasingly important to turn attention to which predictors have the potential for adverse impact and at what selection ratios adverse impact is likely to occur. This article focuses on GMA and the specific ability components that are frequently used in selection. Although there is much research on individual cognitive ability development (and its decline) over the life span, little research has focused on how measuring these abilities has the potential to adversely impact groups of older applicants in work settings. Specific guidance for hiring and promotion in organizations with applicant pools diverse in age is limited. The present study examines how older executives may be impacted by tests of GMA and specific abilities, and discusses practical implications and interventions geared toward supporting older individuals as they continue to pursue their careers or reenter the workforce.

General Mental Ability, Specific Abilities, and Age-Based Declines

Cognitive ability and *intelligence* are both alternate labels for a construct that, with broad scientific consensus, is defined as "a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience" (Gottfredson, 1997, p. 13). Spearman (1904) developed the idea of general intelligence and how to identify the general factor underlying relationships among tests. Thurstone argued for the importance of specific mental abilities, ultimately suggesting that Spearman's general factor, *g*, operates through primary mental abilities (Thurstone, 1941). This structure has come to be known as a hierarchical model. Specific abilities, such as verbal ability, figural reasoning, and inductive reasoning, are narrower abilities that are positively correlated and give rise to a positive manifold. GMA is typically the best predictor of job performance across jobs and ability levels, with little incremental validity for specific abilities (Hunter, 1986; Ones, Viswesvaran, & Dilchert, 2005; Ree, Earles, & Teachout, 1994). However, group differences typically tend to be largest for measures that are highly *g*-loaded (load highly on the general factor found in batteries of many different cognitive tests). One solution that has been proposed for reducing group differences in selection settings that utilize cognitive ability measures is to instead measure specific abilities (Kehoe, 2002), which might be more relevant to job-specific content (Lubinski & Dawis, 1992). The present research examines age differences on such specific cognitive ability measures and compares them to differences on more general measures to determine if any specific ability measures can be used to avoid or minimize age-based adverse impact.

Two second-order factors, fluid and crystallized intelligence, have often emerged in the extensive and systematic factor analyses of Thurstone's (1941) primary mental abilities (Cattell, 1987; Horn, 1972; Horn & Cattell, 1967). Fluid intelligence (Gf) involves the ability to reason, and is often measured by tests of nonverbal figural and inductive reasoning, such as letter series, picture series, and figure classifications (Horn, 1975; Horn, 1976; Horn & Cattell, 1967). Crystallized intelligence (Gc) includes specific abilities that often draw on explicit knowledge, such as tests of vocabulary, verbal comprehension, and general information (Horn, 1976). Aging leads to declines in cognitive ability, but the patterns of decline are different for Gf and Gc (Horn, 1982; Salthouse, 1988, 2013; Schaie & Willis, 1993). Declines in fluid abilities begin as early as young adulthood and continue steadily (Horn, 1975; Schaie, 1994), while crystallized abilities typically

[1] The U.S. Equal Employment Opportunity Commission's (1978) *Uniform Guidelines on Employee Selection Procedures* protect members from any race, sex, or ethnic group from *adverse impact*, which is defined as a substantially different rate of selection in hiring, promotion, or other employment decisions. Selection rates for any subgroup of a protected class that are less than four fifths of the group with the highest rate are considered evidence of adverse impact. Although the *Uniform Guidelines* do not encompass age, the Supreme Court explicitly ruled in Smith v. the City of Jackson (2005) that the ADEA can similarly "authorize recovery on a disparate-impact theory" (p. 11). The ADEA had relied on a disparate impact theory of liability for two decades following the Griggs v. Duke Power Co. (1971) decision. That is, the adverse consequences of employment practices for various subgroups based on race, color, religion, sex, or national origin was interpreted as discrimination regardless of the intent of the company. However, after the Supreme Court's decision in Hazen Paper Co. v. Biggins (1993), some courts "concluded that the ADEA did not authorize disparate-impact theory of liability" (Smith v. the City of Jackson, 2005, p. 8) The issue of whether ADEA could rely on disparate impact theory for determinations about age discrimination was not directly examined in the courts until Smith v. City of Jackson (2005).

continue to increase into middle age and show slower rates of decline over time (Kanfer & Ackerman, 2004; Salthouse, 2013). For instance, with respect to vocabulary, Verhaeghen's (2003) meta-analysis of 210 published articles revealed that older individuals scored 0.80 standard deviation units higher than younger individuals on Gc (although the effect appeared to be confounded by education; Ones et al., 2010).

One factor contributing to age differences in cognitive ability, particularly with respect to fluid abilities, is the Flynn effect. The *Flynn effect* describes the steady rise in average intelligence test performance during the 20th century that has been observed in the United States and other industrialized countries. The average magnitude of the increase in GMA scores has been about 0.2 standard deviations per decade (Neisser et al., 1996), and is accounted for by increases in Gf more so than Gc (Flynn, 1998). The precise causes of the Flynn effect are unknown but likely numerous; hypothesized causes include improved nutrition, health care, and increasingly enriched cognitive environments. Increases in cognitive test scores over time may lead to age-based adverse impact on older cohorts when tests of cognitive ability are used to select or promote employees (Hough, Oswald, & Ployhart, 2001). Not only do older candidates face cognitive declines in fluid abilities, but they may also have to compete with individuals from younger cohorts advantaged in terms of Gf.

One study that examined age differences in work-related cognitive abilities was Avolio and Waldman's (1994) large-scale investigation of 25,140 employees who had taken the General Aptitude Test Battery (GATB). Age was negatively related to GMA ($r = -.15$) and all specific ability subtests. The lowest observed difference between younger and older employees was found for verbal ability. A very large difference was observed for form perception measures; older individuals between the ages of 45 and 54 scored more than one standard deviation unit lower than the 20- to 34-year-old comparison group (Ones et al., 2010). As Avolio and Waldman noted, one limitation of their study was that it did not adequately represent more complex jobs. In contrast, the present study investigates age differences on GMA and specific abilities in a sample of high-level executives.

Investigations of high job-complexity samples are an important extension of past research on age differences in cognitive ability. There is evidence that rates of decline in cognitive ability might be slower for individuals with higher levels of initial baseline ability (Deary, MacLennan, & Starr, 1998), for employees in more complex and enriched jobs (Schooler, Mulatu, & Oates, 1999), and for those engaged in intellectually stimulating activities (Arbuckle, Gold, & Andres, 1986; Hultsch, Hertzog, Small, & Dixon, 1999; Masunaga & Horn, 2001). Presumably, performing complex, intellectually engaging activities allows individuals to use their cognitive abilities to a greater extent, buffering them against cognitive decline (cf. Rizzuto, Cherry, & LeDoux, 2012). Based on these findings with respect to job complexity and engagement in intellectually stimulating activities, cognitive ability declines among business executives may be smaller than in the general working population. Executive-level positions frequently require individuals to solve complex tasks, plan, adapt, think abstractly, and learn quickly and, thus, offer a challenging work environment. The present study seeks to determine the extent of age differences in GMA and specific abilities in a sample of such highly engaged executives. It is expected that due to different patterns of age-based declines, specific ability measures assessing Gc, such as verbal ability, should minimize adverse impact against older executives, whereas measures of Gf will adversely impact older individuals.

Methods and Results: Primary Samples

Sample 1: Participants

The sample consisted of 3,375 individuals applying to executive-level jobs, all of which were vice president or general manager positions within professional, technical, line, and sales occupations. The sample was 60.7% male; 143 applicants did not indicate their age, leaving 3,232 executives on which age analyses were conducted. Age ranged from 20 to 74 years ($M = 42.87$, $SD = 9.48$) at time of application (assessment period: November 2005–July 2014). Among individuals 40 years of age and younger, 64% had previously held between one and three positions with personnel responsibility, 30% had held four or more; among the older group, these proportions were 41% and 58%.

Sample 1: Measures

Executives completed standardized psychological tests that were part of a specially developed managerial and professional test battery. This battery was based on a job analysis; measures were identified based on job relevance. The GMA and specific ability measures were used to determine individual potential and were part of a battery that included measures of experience and background, managerial and professional skills, creative potential, and management style, among others. Given the manner in which the battery is utilized, no single scale is the dominant determinant of the overall evaluation. Brief descriptions of the specific tests examined in this study are provided below.

General mental ability and specific abilities. Cognitive ability was assessed with three objective tests of specific mental abilities measuring verbal and reasoning facilities. Two of these measures assessed fluid abilities: figural and inductive reasoning. Figural reasoning was measured with a 36-item short form of Corsini's (1957) *Non-Verbal Reasoning Test*. Respondents were instructed to identify one of four pictures that best matched a target picture. High scores on figural reasoning are indicative of more advanced nonverbal deductive and analytical reasoning skills. Reliabilities estimated using the KR-20 and KR-21 formulae have ranged from .61 to .85 for various groups for the full 44-item test (Kennedy, 1965). Reliability for the 36-item version was estimated as ranging from .56 to .82 using the Spearman–Brown prophecy formula. The split-half reliability (corrected) was reported as .79, resulting in an estimate of .75 for the 36-item version.

The second fluid intelligence measure assessed inductive reasoning via a 26-item speeded letter series test (Vangent, 1993). Each of the items included a series of letters, with respondents having to correctly identify the letter that would complete the sequence ($\alpha = .79$ in the present sample). High scores indicate a good facility for inductive, or abstract, reasoning by quickly identifying patterns in a series (Horn & Cattell, 1967).

The third measure assessed verbal ability, and thus constituted a measure of crystallized cognitive ability. Verbal ability assesses

one's capacity to understand the meaning of words and language. A short form (46 items) of the *Bruce Vocabulary Inventory* (Bruce, 1974) was used for this purpose. Respondents were instructed to select one word from a row of five that had the same or nearly the same meaning as the first word in the row. The examiner's manual reports a split-half (odd–even) reliability of .92 for a 100 item test (.84 when adjusted for the 46-item test) and a test–retest reliability of .84. An overall indicator of GMA was created by combining scores on the three measures of specific abilities using unit-weights into a total score.

Sample 1: Analyses

We computed standardized effect sizes (Cohen's *d* values) to examine the magnitude of age-related group mean-score differences in GMA and specific fluid and crystallized abilities. Executives between the ages of 20 and 29 were used as a comparison group (numbers of executives in each age group are listed in Table 1). This age group was chosen as a reference group because declines in abilities such as inductive reasoning begin around the age of 30 (Schaie, 1994). The rest of the sample was grouped into 5-year increments, which were then compared with the reference group (see Table 2). Positive effect sizes indicate that younger individuals scored higher on average. In addition to the groups listed above, effect sizes were computed to compare applicants under age 40 (the relevant ADEA cutoff) to older executives. Effect sizes of 0.8 or greater are typically considered large effects, those around 0.5 moderate, and those around 0.2 small (Cohen, 1977).

When to expect adverse impact from cognitive ability and specific ability measures. Based on the effect sizes between executives under age 40 and those 40 years and older, we computed when there would be the potential for age-based adverse impact from the use of GMA and specific ability measures. To this end, we used Sackett and Ellingson's (1997) technique. First, we took the observed *d* values between older and younger workers from Table 3 and used the cumulative density function for a normal distribution to determine the lower-scoring group's selection ratio at each value of the selection ratio for the higher scoring group. The lower-scoring group's selection ratios were divided by the corresponding higher-scoring group's selection ratios to obtain adverse impact ratios. When these ratios are less than .80 (or four fifths), adverse impact is likely to occur for the lower-scoring group if the respective cognitive ability measures are used in strict top-down selection (see Table 4).

Sample 1: Results

Table 1 presents means and standard deviations on GMA, verbal ability, figural reasoning, and inductive reasoning separately by age group. The *d* values comparing each age group to the reference group of 20- to 29-year-olds are presented in Table 2.

There were small age differences in GMA among executives in this sample. All comparison groups scored lower on GMA than the reference group of 20- to 29-year-olds. Standardized mean differences between the reference group and each of the three comparison groups in the 30- to 44-year-old age range were minimal (ds $\leq$ 0.13). Effect sizes for the three age groups in the 45- to 59-year-old age range were in the 0.20s (ds = 0.22 to 0.26). The largest differences in GMA were observed in executives over 59 years of age. Scores for executives between 60 to 64 years were 0.47 standard deviation units lower than the 20- to 29-year-old comparison group. When examined by the ADEA's definition of the protected age group, executives age 40 and older were at a small disadvantage compared with younger executives (d = 0.16). Adverse impact would be likely to occur when organizations are fairly selective (i.e., selecting less than 22% of younger candidates) using GMA scores obtained from the test battery.

Focusing on crystallized abilities, older individuals had higher scores on average compared with younger individuals. Mean scores on verbal ability tended to increase across most age groups, and all age groups scored higher on average than the comparison group of 20- to 29-year-olds. The *d* value for 30- to 34-year-olds indicated that this group scored 0.14 standard deviation units higher than the comparison group on the vocabulary measure. The largest difference with respect to the comparison group was observed in the 55- to 59-year-old group (d = −0.81). Differences for 60- to 64-year-olds were −0.63. When examining executives younger than 40 compared with those 40 and older, the older group has an advantage on the measure of verbal ability (d = −0.33).

Table 1
Summary of Descriptive Statistics by Age (Sample 1)

		General mental ability		Verbal ability (Gc)		Figural reasoning (Gf)		Inductive reasoning (Gf)	
Age group	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
< 25	48	152.56	21.91	50.75	11.33	47.45	9.47	54.06	9.89
25–29	210	156.81	21.56	53.01	9.78	50.68	10.10	53.12	10.27
30–34	404	154.38	21.11	53.94	9.81	50.01	9.27	50.43	10.23
35–39	587	153.69	19.33	55.49	8.91	49.42	9.66	48.79	9.41
40–44	580	153.43	19.78	56.39	8.89	49.06	9.60	47.99	9.38
45–49	551	151.37	20.40	56.72	9.18	48.34	9.94	46.30	9.18
50–54	457	150.70	20.57	57.82	9.22	47.86	9.71	45.02	9.06
55–59	265	150.71	19.00	59.75	7.63	47.01	9.75	43.95	8.89
60–64	106	146.22	18.71	58.60	8.10	45.42	9.74	42.20	7.77
≥ 65	24	132.88	17.01	56.08	9.45	41.58	8.87	35.21	6.97

Note. Gc = crystallized intelligence; Gf = fluid intelligence.

Table 2
Standardized Age Group Mean-Score Differences in Cognitive Ability Among Applicants to Executive Positions (Sample 1)

Age group comparison (reference group 20–29 years)	*n*	General mental ability		Verbal ability (Gc)		Figural reasoning (Gf)		Inductive reasoning (Gf)	
		d	90% CI	*d*	90% CI	*d*	90% CI	*d*	90% CI
30–34	404	0.08	[−0.05, 0.20]	−0.14	[−0.26, −0.01]	0.01	[−0.11, 0.14]	0.28	[0.15, 0.41]
35–39	587	0.12	[0.00, 0.23]	−0.31	[−0.42, −0.20]	0.07	[−0.04, 0.19]	0.47	[0.35, 0.58]
40–44	580	0.13	[0.01, 0.24]	−0.41	[−0.52, −0.30]	0.11	[0.00, 0.22]	0.55	[0.44, 0.67]
45–49	551	0.22	[0.11, 0.34]	−0.44	[−0.55, −0.32]	0.18	[0.07, 0.29]	0.74	[0.62, 0.85]
50–54	547	0.25	[0.13, 0.38]	−0.55	[−0.67, −0.42]	0.23	[0.11, 0.35]	0.87	[0.75, 1.00]
55–59	265	0.26	[0.12, 0.40]	−0.81	[−0.96, −0.66]	0.32	[0.17, 0.46]	0.98	[0.83, 1.13]
60–64	106	0.47	[0.30, 0.64]	−0.63	[−0.80, −0.46]	0.47	[0.30, 0.64]	1.17	[0.99, 1.35]
≥ 65	24	1.09	[0.89, 1.29]	−0.35	[−0.54, −0.16]	0.86	[0.66, 1.06]	1.82	[1.60, 2.05]

Note. $n = 258$ for the 20- to 29-year-old reference group. Gc = crystallized intelligence; Gf = fluid intelligence; d = standardized group mean-score differences (Cohen's d) compared with the reference group (positive effect sizes indicate that younger individuals scored higher on average); CI = confidence interval (two-tailed).

Fluid abilities showed the largest declines with age.[2] However, effects appeared to be less steep for the figural reasoning test compared with the inductive reasoning measure. The 55- to 59-year-old and 60- to 64-year-old age groups scored 0.32 and 0.47 standard deviation units below the comparison group of younger executives, respectively. The protected class of employees 40 years and older scored $d = 0.18$ lower than younger individuals and would be at risk of adverse impact when less than 29% of the younger group were being selected.

Substantial age differences and steep declines were observed in inductive reasoning. Executives age 35 to 39 scored 0.47 standard deviation units lower than the comparison group of 20- to 29-year-olds. This difference continued to increase continuously across age groups, exceeding one standard deviation unit for the two oldest age groups. Younger executives under the age of 40 scored $d = 0.46$ higher than those 40 and older. Given this difference, adverse impact against older employees would be likely to occur when the selection ratio for younger executives is less than 79%. Using inductive reasoning measures alone for personnel decisions in a top-down manner has a high potential for adverse impact, even under most lenient selection scenarios (i.e., high pass rates).[3]

Sample 2: Participants

A second sample was utilized to examine the generalizability of findings from the executive sample to a more heterogeneous working population. Sample 2 consisted of 513 individuals applying to individual contributor and managerial positions between 2009 and July 2010. The age data available were primarily categorical, coded as either under the age of 40 or age 40 and older. In the 119 cases for which specific age data were available, the average age of applicants ranged from 17 to 59, and the average age was 37.94 ($SD = 11.41$). The sample was 73.4% male.

Table 3
Standardized Age Group Mean-Score Differences Among Applicants to Executive Positions—Individuals Younger Than 40 Compared With Those Age 40 and Older (Sample 1)

Measure	n_1	n_2	*d*	90% CI
General mental ability	1,249	1,983	0.16	[0.10, 0.21]
Verbal ability (Gc)	1,249	1,983	−0.33	[−0.39, −0.27]
Figural reasoning (Gf)	1,249	1,983	0.18	[0.12, 0.24]
Inductive reasoning (Gf)	1,249	1,983	0.46	[0.40, 0.52]

Note. Gc = crystallized intelligence; Gf = fluid intelligence; n_1 = sample size for younger individuals (younger than 40 years); n_2 = sample size for older individuals (age 40 and older); d = standardized group mean-score differences (positive effect sizes indicate that younger individuals scored higher on average); CI = confidence interval (two-tailed).

Sample 2: Measures

Applicants completed the same measures of cognitive ability for hiring purposes. Because of differences in jobs, other tools in the assessment battery, as well as sequence of administration, varied.

[2] Although the present study is not longitudinal, when older age groups score lower than younger groups in cognitive ability, differences are typically referred to as declines, given past findings on developmental trajectories obtained in longitudinal investigations (Schaie, 1994). Although cohort effects (such as the Flynn effect, discussed in the introduction) might contribute to observed group differences, these differences are generally accepted to be caused mostly by maturational effects.

[3] One anonymous reviewer of this article raised the issue of range restriction in cognitive ability among executive populations, and its potential impact on the observed age group mean-score differences. Indeed, range restriction on ability measures would most likely reduce the magnitude of effects observed in a given sample, indicating that results from Sample 1 might underestimate group differences in more heterogeneous samples. However, it has been shown that restriction of range in cognitive ability test scores among executives (and especially applicants to such positions) is far smaller than typically believed. Ones and Dilchert (2009), using a sample of more than 1,000 executives and top executives, have shown that reductions in variance (compared with the general working population) range from 0% to a maximum of only 20%, and are consistent across general mental and specific abilities. By comparing results from the present sample of applicants to executive level jobs to results obtained in general population samples (Samples 3 and 4 presented below), we conclude that any potential range restriction effects on cognitive ability appear to have little impact on the overall *pattern* of results, and likely attenuate *absolute magnitudes* of effects sizes only to a small degree.

Another anonymous reviewer raised the issue of range *enhancement*, in this case with regard to the age range of Sample 1. The sample includes about 10% of applicants to executive positions who were younger than 30 years of age. To investigate whether inclusion of this relatively young group skewed results and adverse impact statistics, we recomputed the d values in Table 3 after excluding those individuals—the average decrease in d was 0.04, or 13% across scales.

Table 4
Scenarios Under Which to Expect Adverse Impact From Use of General Mental and Specific Ability Measures Among Applicants to Executive Positions (Sample 1)

Measure	Adverse impact scenario
General mental ability	Younger SR < .22
Verbal ability (Gc)	Older SR < .65
Figural reasoning (Gf)	Younger SR < .29
Inductive reasoning (Gf)	Younger SR < .79

Note. Values indicate the selection ratios (SRs) of the higher scoring group that may lead to adverse impact against the lower scoring group. For example, for inductive reasoning, younger SR < .79 indicates that when the SR for the higher scoring younger executives is less than 79%, adverse impact on older individuals is to be expected. Gc = crystallized intelligence; Gf = fluid intelligence.

General mental ability and specific abilities. Cognitive ability was assessed with the Thurstone Test of Mental Alertness (TMA; Thurstone & Thurstone, 1996), a 126-item test with a 20-min time limit. Fluid abilities (inductive reasoning) are assessed via 36 number series items on this test. Measures of crystallized abilities include 18 arithmetic reasoning items assessing individuals' ability to solve quantitative word problems, as well as a verbal ability scale with two item types: 36 synonym/antonym and 36 definition items. The average test–retest reliability of the two quantitative scales is reported as .94, the average test–retest reliability of the verbal scale as .89 across samples (Thurstone & Thurstone, 1996). An overall indicator of GMA was created by unit-weighting the scores on the three specific abilities into a total score.

Sample 2: Analyses

Analyses were parallel to those for Sample 1. Standardized effect sizes (Cohen's *d* values) were computed to examine the magnitude of age-related group mean-score differences in GMA and specific fluid and crystallized abilities comparing applicants under 40 to those age 40 and older. Given the nature of the age data available in this sample, comparisons across other age groups could not be conducted.

Sample 2: Results

The means and standard deviations on GMA and the specific mental abilities for applicants under the age of 40 and 40 years and older as measured by the TMA are presented in Table 5. Age differences on the measure of GMA were negligible ($d = -0.01$; Table 6). Adverse impact would be unlikely to occur when GMA is measured with the TMA.

In terms of verbal measures of crystallized abilities, older applicants had higher scores on average compared with younger individuals with a standardized mean difference of −0.13, reducing the threat of adverse impact. In contrast, older applicants scored lower on measures of crystallized quantitative ability and fluid ability than younger individuals ($d = 0.10$ and $d = 0.17$, respectively). Relying on these measures, the protected group of applicants 40 years and older would only be at risk of adverse impact if less than 4% or less than 26% of the younger group were being selected (see Table 7).

Methods and Results: Supplemental Samples

The primary goal of this study was to investigate age differences on measures of GMA and specific cognitive abilities among job applicants to executive positions to determine which predictors might lead to age-based adverse impact in making selection and advancement decisions at the executive level. The degree to which these findings might generalize to other job applicant groups and selection scenarios—particularly nonexecutive jobs and samples more heterogeneous in age and/or cognitive ability—is an important consideration. Although the results from the mixed executive/nonexecutive Sample 2 provide a first indication, we sought additional samples to investigate the generalizability of the pattern of effects.

Samples 3 and 4

Sample 3 was composed of a representative sample of employed and retired U.S. adults from the *University of Michigan Health and Retirement Study* (HRS; http://hrsonline.isr.umich.edu). Sample 4 was composed of a general population sample from the *Midlife in the United States* Study (MIDUS; http://midus.wisc.edu). Both studies focus on middle-age and older adults comparable in age range to our original executive samples, making them ideally suited to investigate generalizability of effects.

Measures and Analyses: Samples 3 and 4

Both the HRS and MIDUS are longitudinal studies that include a variety of cognitive ability measures administered to different waves of participants over time. For the purpose of this study, we identified those cognitive tests and scales available for the respective samples that most paralleled the constructs measured in Samples 1 and 2. In the HRS data, measures of fluid and crystallized abilities were available for a significant number of participants in

Table 5
Summary of Descriptive Statistics by Age (Sample 2)

		General mental ability		Linguistic ability (Gc: verbal)		Arithmetic reasoning (Gc: quantitative)		Number series (Gf: inductive)	
Age group	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
< 40	297	59.36	16.15	33.06	9.39	8.77	2.86	17.52	4.96
≥ 40	216	59.44	16.10	34.25	9.61	8.49	2.63	16.70	4.88

Note. Gc = crystallized intelligence; Gf = fluid intelligence.

Table 6
Standardized Age Group Mean-Score Differences for Applicants Younger Than 40 Compared With Applicants Age 40 and Older (Sample 2)

Measure	n_1	n_2	d	90% CI
General mental ability	297	216	−0.01	[−0.15, 0.14]
Linguistic ability (Gc: verbal)	297	216	−0.13	[−0.27, 0.02]
Arithmetic reasoning (Gc: quantitative)	297	216	0.10	[−0.05, 0.25]
Number Series (Gf: inductive)	297	216	0.17	[0.02, 0.31]

Note. n_1 = sample size for younger individuals (younger than 40 years); n_2 = sample size for older individuals (age 40 and older); d = standardized group mean-score differences (positive effect sizes indicate that younger individuals scored higher on average); CI = confidence interval (two-tailed); Gc = crystallized intelligence; Gf = fluid intelligence.

the form of number series as well as a vocabulary/verbal comprehension scale. In addition, scales assessing short- and long-term memory were available (the measures are described in detail in the official codebooks, available at http://hrsonline.isr.umich.edu/index.php?p=showcbk). In the MIDUS data, measures of fluid and crystallized abilities were available in the form of subscales from three popular ability tests: The Wechsler Adult Intelligence Scale (Vocabulary scale [Gc]), Raven's Advanced Progressive Matrices (APM), and Schaie–Thurston Letter Series (both Gf). A measure of perceptual processing speed was also available (Wechsler Digit Symbol Scale). The analyses carried out mirrored those for Samples 1 and 2. Where necessary, data from several younger age groups were combined to create reference groups large enough so that stable estimates of group mean-score differences could be computed.

Results: Samples 3 and 4

Results for the supplemental Samples 3 and 4 are presented in Appendixes A and B, respectively. Tables A1 and B1 present descriptive statistics for both samples. In terms of average standing on the different cognitive ability tests, the relative trends observed in Samples 1 and 2 are mirrored in these two general population samples: with increasing age, there is a consistent decrease in group mean scores on the fluid ability measures in Samples 3 (Number Series) and 4 (Letter Series and Raven's APM). On the crystallized abilities measures, mean scores are stable or increase among older groups. Tables A2 and B2 express these results in terms of standardized group mean-score differences in relation to the youngest available reference group. For Sample 3 (employed and retired adults), on the crystallized ability measure, all older age groups scored higher than the reference group, with effects ranging from −0.11 to −0.66 standard deviation units. On the fluid ability measure, older age groups scored progressively lower on average. However, the most notable decrement again occurs in the group age 65 and older (d = 0.68). Parallel and somewhat more pronounced age differences are observed on the two measures of short- and long-term memory (ds increasing monotonously from 0.07 to 0.87, favoring younger individuals). For Sample 4, on the crystallized ability measure, all older age groups again scored higher than the reference group (ds from −0.18 to −0.37). On the fluid ability measures (Raven's APM and Letter Series), older groups scored progressively worse, with small mean-score differences for the 35- to 44-year-old group, and large differences (ds from 0.50 to 0.82) for older individuals, with the largest effect again observed in the very oldest group (d = 1.55). Results for the measure of perceptual processing speed also mirror these findings. Despite the relatively small group sample sizes (in comparison to Samples 1 – 3), the effects for these older age groups are established with statistical confidence, as indicated by the confidence intervals reported in Table B2.

In sum, from examining the patterns of cognitive ability scores among these more general, nonapplicant samples, we conclude that the effects first observed in our two applicant samples are likely to reflect generalizable age group mean-score differences that also exist in the general population. The restriction of range in the age composition of our applicant samples (as well as indirect restriction of range via correlated variables such as managerial experience, for which candidates might be prescreened) does not appear to lead to a notable reduction in age differences observed. The pattern of effects (modest to large differences in fluid abilities, favoring younger individuals, and a consistent pattern of group differences favoring older individuals on crystallized ability measures) is observed to similar degrees in the general population. The increase in magnitude in the most extreme age group can be attributed to the inclusion of older adults in these samples, who typically do not appear even in executive job applicant samples (individuals 75 years and older).

Discussion

This study found a modest relationship between age and GMA in a large sample of executives. In Sample 1, the average score gradually decreased across age groups, with largest differences observed for individuals older than 60 years. Cognitive ability is one of the best predictors of job performance, but the existing age differences can lead to adverse impact when using GMA test scores to select executive candidates, at least when overall selection ratios are very low (i.e., competition for positions is strong).

Larger age group mean-scores differences were observed on some specific measures of fluid ability. In Sample 1, a measure of fluid intelligence (letter series test of inductive reasoning) showed substantial age group differences. Older executives scored much lower on average than younger executives. Average scores decreased steadily and across all age groups. This is consistent with

Table 7
Scenarios Under Which to Expect Adverse Impact From Use of General Mental and Specific Ability Measures Among Applicants (Sample 2)

Measure	Adverse impact scenario
General mental ability	Adverse impact unlikely
Linguistic ability (Gc: verbal)	Older SR < .11
Arithmetic reasoning (Gf: quantitative)	Younger SR < .04
Number Series (Gf: inductive)	Younger SR < .26

Note. Values indicate the selection ratios of the higher scoring group that may lead to adverse impact against the lower scoring group. For example, for Number Series (Gf: inductive), younger selection ratio (SR) < .26 indicates that when the SR for the higher scoring younger executives is less than 26%, adverse impact on older individuals is to be expected. Gc = crystallized intelligence; Gf = fluid intelligence.

previous findings that specific abilities that measure fluid intelligence decrease over time, and findings that inductive reasoning may start to decline as early as age 30 (Schaie, 1994). Executives age 55 to 59 scored 1.06 standard deviation units lower than 20- to 29-year-olds on inductive reasoning, with even larger effect sizes observed for older employees. In a replication (Sample 4), older adults from the general population also scored markedly lower on inductive reasoning measures. On the Raven's APM, adults between age 55 and 64 and those older than 65 scored 0.82 and 1.34 standard deviation units higher than 25- to 35-year olds, with similar declines on the Schaie–Thurston Letter Series test. Age group mean-score differences on these inductive reasoning measures reach magnitudes typically observed for more frequently investigated race and ethnic group differences. However, the potential for adverse impact against older individuals as well as ways to minimize this impact have received little attention in the literature. Organizations should be cautious when using certain tests of inductive reasoning, such as letter series tests, given the magnitude of age differences established in this research. Selecting on such measures alone will likely lead to younger individuals being selected at much greater rates than older candidates.

The test of figural reasoning included in Sample 1 also represented a measure of fluid intelligence but displayed smaller age group differences. Including such a measure of reasoning to assess applicants might reduce the potential for adverse impact compared with inductive reasoning tests. However, adverse impact potential still exists in cases in which organizations are very selective using a figural reasoning test (i.e., selection ratio $\leq$ .29 for younger individuals).

In contrast to fluid abilities, older employees and adults appear to have an advantage on tests of vocabulary and verbal ability that measure crystallized ability. Across all samples in this study, each age subgroup scored higher than the comparison group of younger individuals on these verbal and linguistic measures. In Sample 1, the magnitude of age differences at first steadily increased and then leveled off after the 55- to 59-year-old age group that performed the best on this specific ability ($d = -0.81$). Similarly, in Sample 3, 50- to 59-year-olds scored the highest compared with the reference group (*d*s ranging from -0.61 to -0.66), with individuals 65 and older still scoring higher, but to a lesser extent (-0.28). In Sample 4, the magnitude of these age differences was in the 0.30s for a majority of the age groups. This is consistent with past findings of crystallized intelligence showing the slowest rates of decline over time (Avolio & Waldman, 1994; Owens, 1966; Schaie, 1994). However, although Avolio and Waldman (1994) found that younger workers performed better than older workers on verbal ability, all four samples in this study demonstrated that older workers had an advantage on the specific tests of verbal/linguistic ability. Our finding is more consistent with the broader literature on aging that established patterns of increases in crystallized abilities over most of the life span (see, e.g., Schwartzman, Gold, Andres, Arbuckle, & Chaikelson, 1987). The samples in this study exhibit notable consistency, and one potential explanation for Avolio and Waldman's (1994) different results may be that patterns of development (including decline) might be idiosyncratic to even more specific aspects of verbal ability, and thus findings more study specific. Second, observed age differences might be more sample dependent than previously acknowledged, as suggested by prior theory linking patterns of decline to typical cognitive engagement as well as initial ability level (see above). Older individuals in Sample 1 may have had the advantage of holding complex executive jobs for which responsibilities and tasks provide a buffer against typically experienced cognitive declines. More research is also needed comparing other crystallized measures used in selection settings such as tests of job knowledge, as this study highlights the importance of analyzing adverse impact of the specific measures an organization is using with a sample that is representative of the types of applicants or candidates an organization expects to test.

In selection settings, there are legal implications in terms of adverse impact connected to the use of cognitive ability measures if there is not demonstrated job relatedness of the assessment measure or measures. However, it is also important to note that declines in cognitive ability may not always correspond to large age group differences in job performance. In these cases, fairness issues may arise if the performance of older employees is differentially predicted compared with younger employees on the basis of cognitive ability measures. The scientific literature on differential prediction using cognitive ability tests on the basis of age has yet to develop. Although conceptually distinct, Sackett and Bobko (2010) point out that from the legal perspective "a finding of group differences often constitutes the point of entry to regulatory scrutiny; it is in the context of a finding of adverse impact that an investigation of differential validity comes into play (cf. *Uniform Guidelines on Employee Selection Procedures*, U.S. Equal Employment Opportunity Commission, 1978, Section 1607.14B8[b])" (p. 214). We see the present research as a starting point in this necessary line of scholarly work (see below).

One factor potentially affecting job relatedness, as well as declines in job performance, is the extent to which a position requires Gf and Gc. Jobs that place higher demands on fluid rather than crystallized ability typically become more difficult for older workers, requiring greater effort on the part of older employees to maintain high levels of job performance (Kanfer & Ackerman, 2004). In select cases where the demands on Gf are high (and cannot be compensated for by Gc), older workers are often reassigned or offered early retirement programs. For example, air traffic controllers in the U.S. have a mandatory retirement age of 56 given the high demand on attentional effort (Gf).

However, on the whole, findings from a recent meta-analysis indicate that age is largely unrelated to core task performance across jobs (Ng & Feldman, 2008). The authors also reported that the relationship may be curvilinear, with age being positively related to performance for age groups 40 and below and negatively for individuals over age 40 ($r = -.05$). The magnitude of this effect is small, particularly in comparison to some of the observed differences on general ability measures and specific fluid abilities such as inductive reasoning. One potential explanation for the small magnitude of the age–job performance relationship despite larger age differences in cognitive abilities is that older workers are able to compensate relatively well for age-based declines in these abilities (Baltes & Baltes, 1990). Another potential explanation is that changes in relevant job knowledge (the more direct, causal determinant of performance) are not commensurate with declines in underlying abilities. In these cases, it would be important to examine differential validity and differential prediction for age subgroups. In terms of differential prediction, a test is considered biased, or unfair, if the criterion score that is predicted by the

regression line is consistently too high or too low for members of a subgroup. Differential prediction of cognitive ability measures has been frequently examined with respect to race and sex (Society for Industrial and Organizational Psychology, 2002), but studies comparing age subgroups are lacking. Although job performance data were not available to examine differential prediction in this study, examining differential prediction for older workers is an important area for further research. Based on findings from this study, tests of specific abilities such as inductive reasoning may differentially predict performance across age subgroups for jobs for which there are small differences in job performance across age groups, although this must be tested empirically with appropriate criterion data.

Future Research

Our findings of age differences that vary across crystallized and fluid cognitive abilities have interesting implications for workplace research on age and performance. Ng and Feldman's (2008) meta-analysis, as well as prior reviews (McEvoy & Cascio, 1989; Sturman, 2003), found mostly small relationships between age and job performance. However, such findings might be contingent on employees working in similar jobs throughout their career. One interesting question is how older individuals perform compared with younger individuals when the former start out in a new job or career. Career transitions for older workers are not uncommon. A study tracking working adults indicated that 27% of workers employed full time at age 51 to 55 changed occupations by age 65 to 69 (Johnson, Kawachi, & Lewis, 2009). Such is an increasingly common situation, in part due to the recent economic downturn (and associated depletion of retirement funds), but also more generally due to technological advances that require acquisition of new skills, retraining, or entire career changes as some jobs change and selected occupations become obsolete (cf. Beier, Teachout, & Cox, 2012). When older individuals make career changes without having had the opportunity to accumulate relevant job knowledge and skills that otherwise help compensate for declining fluid abilities, larger discrepancies in job performance may be observed across age groups. Future research should empirically evaluate this hypothesis.

Future research should focus on how to best help older workers transition to new positions or careers to support the gainful employment of older individuals and maximize their potential within the workplace. There has been some research in nonworkplace settings examining the effectiveness of training on specific fluid ability measures such as inductive (Ball et al., 2002; Blieszner, Willis, & Baltes, 1981) and figural reasoning (Willis, Blieszner, & Baltes, 1981). This work is promising in terms of its potential to help individuals regain some of these skills. Programs designed to help older individuals reenter the workforce could utilize interventions to retrain specific abilities, as well as familiarize older adults with tests similar to those used in selection settings, to enable more accurate measurement of their cognitive potential. Ng and Feldman's (2008) meta-analysis found a correlation of –.04 between age and training performance (confidence interval = –.07 to –.02), suggesting that older individuals do not do as well in training as their younger counterparts. Although this is a small difference, differences may be more substantive when older employees have to transfer acquired knowledge (Ng & Feldman, 2008) or when learning novel information, as would be the case for older individuals moving into a new career. Learning novel or difficult tasks places a high demand on attentional resources, which are determined by working memory capacity, another ability that typically declines with age (see Table A2, as well as Kanfer & Ackerman, 1989; Salthouse & Babcock, 1991). Future research is needed on training outcomes for older employees who are learning new job skills unrelated to their current position or career. This research would help guide training for older individuals learning novel skills to keep up with changes related to their current position or for transitioning to a new position or career.

Practical Implications

This study also has practical implications for designing selection systems with the goal of reducing age-based adverse impact. First, our results highlight that age is an important variable for organizations to consider with regard to adverse impact, particularly when using measures of fluid abilities. In this research, we examined GMA and a range of specific ability components, including verbal ability, inductive reasoning, figural reasoning, memory, and processing speed. As a whole, although age differences on the composite score in Sample 1 are modest, there is still potential for age-based adverse impact when fewer than 22% of younger workers are selected. However, organizations vary in terms of which cognitive ability measures they use as well as how they combine scores in making selection and promotion decisions. Among the executives investigated, two fluid ability measures assessing reasoning (inductive and figural) displayed different patterns of age-group differences and thus different potentials for adverse impact. Although many organizations do use a composite of ability measures, many rely solely on fluid reasoning measures to assess cognitive ability, partly in an effort to reduce group differences among subgroups of other protected classes (e.g., race). Samples 3 and 4 provide evidence that there can be large age group mean-score differences on measures of specific fluid abilities. In these cases it is especially important for organizations to be cognizant of such differences and consider how the respective tools fit into the selection system as a whole. This includes careful consideration of the method used to combine scores across cognitive ability tests or components (cf. Kuncel, Klieger, Connelly, & Ones, 2013), as quasiclinical approaches (or inconsistent weights in general) will result in imprecise estimates of adverse impact.

On the whole, little research in applied psychology has examined age differences in cognitive abilities and their legal and general fairness implications. Future work should model age differences and implications of other specific ability measures to determine when adverse impact would be likely to occur. Such investigations should also include other complex jobs to examine whether findings from this study generalize to jobs of similar complexity. In addition, low-complexity jobs should also be considered. The psychological literature on aging posits that age-related cognitive decline is more pronounced in individuals who start out at lower regions of the ability spectrum (see above, also Deary, 2000, Chapter 8). Based on such findings, one might predict that age group mean-score differences among applicants to lower-complexity jobs (who display lower cognitive ability mean scores in general, see Sackett & Ostgaard, 1994), might be even larger, especially on fluid ability measures. Finally, in addition to

measuring other specific abilities, researchers should explore age differences on other personnel selection measures (i.e., job knowledge tests, physical ability tests) so that the potential impact of age on mean-score levels will be better understood in the future (Hedge, Borman, & Lammlein, 2006).

Our study shows that organizations that are interested in reducing the potential for adverse impact of cognitive ability measures on older employees could consider including tests of verbal ability in assessment batteries where appropriate. When included in cognitive ability composites, tests of verbal ability can help balance out differences on other specific ability measures that show larger age-based differences favoring younger applicants. This study specifically examined verbal ability, but other measures of Gc should also be examined with respect to their potential for minimizing adverse impact in selection and promotion decisions.

Finally, in terms of the selection system as a whole, noncognitive predictors that display only small age differences might also be considered. This strategy has been found to be generally effective in reducing racioethnic and sex-based adverse impact, and should similarly be effective for age, although the magnitude of reduction will depend upon predictor validities and intercorrelations (cf. Ployhart & Holtz, 2008). For instance, one measure included in the executive assessment battery examined in Sample 1 was a measure of business ethics. Measures of integrity offer both high validity and small age differences favoring older applicants. (A meta-analysis by Ones & Viswesvaran, 1998, showed that applicants 40 years or older [$N = 9{,}743$] scored 0.08 standard deviation units higher on overt integrity tests than those younger than 40 [$N = 68{,}477$].) Measures of biodata or managerial competencies may also allow older employees to highlight their strengths and managerial experience relative to the position. Where appropriate, the inclusion of noncognitive predictors with small age-based differences can help organizations minimize adverse impact against older applicants. This is particularly important in cases in which scores on the predictor measures are not congruent with older individuals' actual performance on the job, because crystallized abilities may help them compensate for lower levels of fluid abilities.

Conclusions

Legal and fairness concerns necessitate that organizations consider group mean-score differences on assessment tools used at various stages of the human resources process. Too often, research and applied investigations fail to comprehensively examine potential group differences on our most important predictor tools. Cognitive ability tests are widely used in making personnel decisions, and thus it is important for human resource professionals to be aware of age differences on these measures. As employees remain in the workforce longer, it is important to be cognizant of how selection systems may impact older workers. This study found that older executives performed slightly worse on tests of GMA and figural reasoning, and generally much worse on tests of inductive reasoning, which assess fluid intelligence. Older executives do seem to have an advantage, however, when it comes to some tests of verbal ability, a type of crystallized intelligence. This is in contrast to Avolio and Waldman's (1994) finding with respect to measures of verbal ability as measured by the GATB, a finding that does not appear to generalize to managerial and executive assessment. Including a measure of verbal ability in cognitive ability composites should help organizations reduce the risk of creating adverse impact against older individuals, and is particularly relevant for high-complexity jobs and positions for which such applicants are more common. Even when overall scores on cognitive ability test batteries are used in making personnel decisions, awareness of the intricate patterns of age differences on the specific tests constituting the composite is crucial to responsibly estimate adverse impact potential with the goal of avoiding age discrimination.

References

Age Discrimination in Employment Act of 1967, 29 U.S.C. §§ 621–634.

Arbuckle, T. Y., Gold, D., & Andres, D. (1986). Cognitive functioning of older people in relation to social and personality variables. *Psychology and Aging, 1,* 55–62. http://dx.doi.org/10.1037/0882-7974.1.1.55

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. *Psychology and Aging, 9,* 430–442. http://dx.doi.org/10.1037/0882-7974.9.3.430

Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., . . . Willis, S. L. (2002). Effects of cognitive training interventions with older adults: A randomized controlled trial. *JAMA: Journal of the American Medical Association, 288,* 2271–2281. http://dx.doi.org/10.1001/jama.288.18.2271

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1–34). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511665684.003

Beier, M. E., Teachout, M. S., & Cox, C. B. (2012). The training and development of an aging workforce. In W. C. Borman & J. W. Hedge (Eds.), *The Oxford handbook of work and aging* (pp. 436–453). New York: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780199562930.001.0001

Blieszner, R., Willis, S. L., & Baltes, P. B. (1981). Training research in aging on the fluid ability of inductive reasoning. *Journal of Applied Developmental Psychology, 2,* 247–265. http://dx.doi.org/10.1016/0193-3973(81)90005-8

Bruce, M. M. (1974). *Bruce Vocabulary Inventory*. New Rochelle, NY: Author.

Cattell, R. B. (1987). *Intelligence: Its structure, growth, and action*. Amsterdam, the Netherlands: North-Holland.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. San Diego, CA: Academic Press.

Corsini, R. J. (1957). *Non-verbal Reasoning*. Park Ridge, IL: London House.

Deary, I. J. (2000). *Looking down on human intelligence*. New York: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198524175.001.0001

Deary, I. J., MacLellan, W. J., & Starr, J. M. (1998). Is age kinder to the initially more able?: Differential ageing of verbal ability in the healthy old people in Edinburgh study. *Intelligence, 26,* 357–375. http://dx.doi.org/10.1016/S0160-2896(99)00005-7

Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, & Department of Justice. (1978). Uniform guidelines on employee selection procedures. *Federal Register, 43,* 38290–39315.

Flynn, J. R. (1998). IQ gains over time: Toward finding the causes. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related*

measures (pp. 25–66). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10270-001

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography. *Intelligence, 24,* 13–23. http://dx.doi.org/10.1016/S0160-2896(97)90011-8

Griggs v. Duke Power Co., 401 U.S. 424 (1971).

Grossman, R. J. (2003, August). Are you overlooking older workers? *HRMagazine, 48,* 44–45.

Hazen Paper Co. v. Biggins, 507 U.S. 604 (1993).

Hedge, J. W., Borman, W. C., & Lammlein, S. E. (2006). *The aging workforce: Realities, myths, and implications for organizations*. Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/11325-000

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 769,* 41–45. http://dx.doi.org/10.1126/science.7604277

Horn, J. L. (1972). The structure of intellect; primary abilities. In R. H. Dreger (Ed.), *Multivariate Personality Research* (pp. 451–455). Baton Rouge, LA: Claitor.

Horn, J. L. (1975). Psychometric studies of aging and intelligence. In S. Gershon & A. Raskin (Eds.), *Genesis and treatment of psychological disorders in the elderly* (Vol. 2, p. 19–34). New York: Raven.

Horn, J. L. (1976). Human abilities: A review of research and theory in the early 1970s. *Annual Review of Psychology, 27,* 437–485. http://dx.doi.org/10.1146/annurev.ps.27.020176.002253

Horn, J. L. (1982). The gaining of human abilities. In B. B. Wolman (Ed.), *Handbook of developmental psychology: Research and theory* (pp. 847–870). Englewood Cliffs, NJ: Prentice Hall.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26,* 107–129. http://dx.doi.org/10.1016/0001-6918(67)90011-X

Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment, 9,* 152–194. http://dx.doi.org/10.1111/1468-2389.00171

Hultsch, D. F., Hertzog, C., Small, B. J., & Dixon, R. A. (1999). Use it or lose it: Engaged lifestyle as a buffer of cognitive decline in aging? *Psychology and Aging, 14,* 245–263. http://dx.doi.org/10.1037/0882-7974.14.2.245

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29,* 340–362. http://dx.doi.org/10.1016/0001-8791(86)90013-8

Johnson, R. W., Kawachi, J., & Lewis, E. K. (2009). *Older workers on the move: Recareering in later life*. Washington, DC: American Association of Retired Persons.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74,* 657–690. http://dx.doi.org/10.1037/0021-9010.74.4.657

Kanfer, R., & Ackerman, P. L. (2004). Aging, adult development, and work motivation. *Academy of Management Review, 29,* 440–458.

Kehoe, J. F. (2002). General mental ability and selection in private sector organizations: A commentary. *Human Performance, 15,* 97–106. http://dx.doi.org/10.1080/08959285.2002.9668085

Kennedy, J. E. (1965). *The sixth mental measurements yearbook*. Lincoln, IL: University of Nebraska Press.

Kuncel, N. R., Klieger, D. M., Connelly, B. S., & Ones, D. S. (2013). Mechanical versus clinical data combination in selection and admissions decisions: A meta-analysis. *Journal of Applied Psychology, 98,* 1060–1072. http://dx.doi.org/10.1037/a0034156

Leahey, C. (2012, January 9). The workforce of the future: Older and healthier. *CNN Money*. Retrieved from http://management.fortune.cnn.com/2012/01/09/workforce-future-older/

Lubinski, D., & Dawis, R. V. (1992). Aptitudes, skills and proficiencies. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 3, pp. 1–59). Palo Alto, CA: Consulting Psychologists Press.

Maestas, N., & Zissimopoulos, J. (2010). How longer work lives ease the crunch of population aging. *The Journal of Economic Perspectives, 24,* 139–160. http://dx.doi.org/10.1257/jep.24.1.139

Masunaga, H., & Horn, J. (2001). Expertise and age-related changes in components of intelligence. *Psychology and Aging, 16,* 293–311. http://dx.doi.org/10.1037/0882-7974.16.2.293

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

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S., . . . Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51,* 77–101. http://dx.doi.org/10.1037/0003-066X.51.2.77

Ng, T. W. H., & Feldman, D. C. (2008). The relationship of age to ten dimensions of job performance. *Journal of Applied Psychology, 93,* 392–423. http://dx.doi.org/10.1037/0021-9010.93.2.392

Ones, D. S., & Dilchert, S. (2009). How special are executives? How special should executive selection be? Observations and recommendations. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 2,* 163–170. http://dx.doi.org/10.1111/j.1754-9434.2009.01127.x

Ones, D. S., Dilchert, S., Viswesvaran, C., & Salgado, J. F. (2010). Cognitive abilities. In J. L. Farr & N. T. Tippins (Eds.), *Handbook of employee selection* (pp. 255–275). New York: Routledge.

Ones, D. S., & Viswesvaran, C. (1998). Gender, age, and race differences on overt integrity tests: Results across four large-scale job applicant data sets. *Journal of Applied Psychology, 83,* 35–42. http://dx.doi.org/10.1037/0021-9010.83.1.35

Ones, D. S., Viswesvaran, C., & Dilchert, S. (2005). Cognitive ability in personnel selection decisions. In A. Evers, O. Voskuijl, & N. Anderson (Eds.), *Handbook of personnel selection* (pp. 143–173). Oxford, United Kingdom: Blackwell.

Owens, W. A. (1966). Age and mental abilities: A second adult follow-up. *Journal of Educational Psychology, 57,* 311–325. http://dx.doi.org/10.1037/h0023962

Ployhart, R. E., & Holtz, B. C. (2008). The diversity-validity dilemma: Strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology, 61,* 153–172. http://dx.doi.org/10.1111/j.1744-6570.2008.00109.x

Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: Not much more than g. *Journal of Applied Psychology, 79,* 518–524. http://dx.doi.org/10.1037/0021-9010.79.4.518

Rizzuto, T. E., Cherry, K. E., & LeDoux, J. A. (2012). The aging process and cognitive abilities. In W. C. Borman & J. W. Hedge (Eds.), *The Oxford handbook of work and aging* (pp. 236–255). New York: Oxford University Press.

Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology, 54,* 297–330. http://dx.doi.org/10.1111/j.1744-6570.2001.tb00094.x

Ryan, A. M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology, 52,* 359–392. http://dx.doi.org/10.1111/j.1744-6570.1999.tb00165.x

Sackett, P. R., & Bobko, P. (2010). Conceptual and technical issues in conducting and interpreting differential prediction analyses. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 3,* 213–217. http://dx.doi.org/10.1111/j.1754-9434.2010.01226.x

Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multipredictor composites on group differences and adverse impact. *Person-*

nel Psychology, 50, 707–721. http://dx.doi.org/10.1111/j.1744-6570.1997.tb00711.x

Sackett, P. R., & Ostgaard, D. J. (1994). Job-specific applicant pools and national norms for cognitive ability tests: Implications for range restriction corrections in validation research. *Journal of Applied Psychology, 79,* 680–684. http://dx.doi.org/10.1037/0021-9010.79.5.680

Salthouse, T. A. (1988). Initiating the formalization of theories of cognitive aging. *Psychology and Aging, 3,* 3–16. http://dx.doi.org/10.1037/0882-7974.3.1.3

Salthouse, T. A. (2013). Within-cohort age-related differences in cognitive functioning. *Psychological Science, 24,* 123–130. http://dx.doi.org/10.1177/0956797612450893

Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology, 27,* 763–776. http://dx.doi.org/10.1037/0012-1649.27.5.763

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist, 49,* 304–313. http://dx.doi.org/10.1037/0003-066X.49.4.304

Schaie, K. W., & Willis, S. L. (1993). Age difference patterns of psychometric intelligence in adulthood: Generalizability within and across ability domains. *Psychology and Aging, 8,* 44–55. http://dx.doi.org/10.1037/0882-7974.8.1.44

Schmidt, F. L. (2011). A theory of sex differences in technical aptitude and some supporting evidence. *Perspectives on Psychological Science, 6,* 560–573. http://dx.doi.org/10.1177/1745691611419670

Schooler, C., Mulatu, M. S., & Oates, G. (1999). The continuing effects of substantively complex work on the intellectual functioning of older workers. *Psychology and Aging, 14,* 483–506. http://dx.doi.org/10.1037/0882-7974.14.3.483

Schwartzman, A. E., Gold, D., Andres, D., Arbuckle, T. Y., & Chaikelson, J. (1987). Stability of intelligence: A 40-year follow-up. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 41,* 244–256. http://dx.doi.org/10.1037/h0084155

Smith v. City of Jackson, 544 U.S. 228 (2005).

Society for Industrial and Organizational Psychology. (2002). *Principles for the validation and use of personnel selection procedures* (4th ed.). Bowling Green, OH: Author. Retrieved from http://www.siop.org/_principles/principles.pdf

Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology, 15,* 201–293. http://dx.doi.org/10.2307/1412107

Sturman, M. C. (2003). Searching for the inverted U-shaped relationship between time and performance: Meta-analyses of the experience/performance, tenure/performance, and age/performance relationships. *Journal of Management, 29,* 609–640. http://dx.doi.org/10.1016/S0149-2063(03)00028-X

Thurstone, L. L. (1941). *Factorial studies of intelligence*. Chicago, IL: University of Chicago Press.

Thurstone, L. L., & Thurstone, T. G. (1996). *Information guide for the Thurstone Test of Mental Alertness*. Arlington, VA: Vangent.

United Nations Population Fund. (2012). *Aging in the 21st century: A celebration and a challenge.* Retrieved from http://unfpa.org/ageingreport/

U.S. Bureau of Labor Statistics. (2009, November). *Labor force projections to 2018: Older workers staying more active*. Retrieved from http://www.bls.gov/opub/mlr/2009/11/art3full.pdf

U.S. Department of Health and Human Services, Administration on Aging. (2010, May). *The next four decades: The older population in the United States: 2010 to 2050*. Retrieved from http://www.aoa.gov/AoARoot/Aging_Statistics/future_growth/DOCS/p25-1138.pdf

U.S. Equal Employment Opportunity Commission. (2010). *Age discrimination in employment act*. Retrieved from http://www.eeoc.gov/eeoc/statistics/enforcement/adea.cfm

Vangent. (1993). *Computer Programmer Aptitude Battery examiner's manual*. Chicago, IL: Author.

Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging, 18,* 332–339. http://dx.doi.org/10.1037/0882-7974.18.2.332

Willis, S. L., Blieszner, R., & Baltes, P. B. (1981). Intellectual training research in aging: Modification of performance on the fluid ability of figural relations. *Journal of Educational Psychology, 73,* 41–50. http://dx.doi.org/10.1037/0022-0663.73.1.41

Appendix A

Generalizability—Results For Sample 3

Table A1
Summary of Descriptive Statistics by Age (Sample 3)

	Immediate recall (memory: short term, working)			Delayed recall (memory: long term, storage and retrieval)			Vocabulary (Gc: verbal)			Number Series (Gf: inductive)		
Age group	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
25–29	8	7.00	0.93	8	6.13	1.64	5	4.80	1.10			
30–34	15	6.47	1.88	15	6.07	1.91	10	4.50	1.90			
35–39	60	6.72	1.61	59	5.81	1.98	34	4.09	1.99	18	2.28	0.89
40–44	165	6.68	1.54	165	5.69	1.87	88	5.23	2.13	23	2.48	0.73
45–49	452	6.49	1.63	448	5.63	1.78	265	5.10	2.18	89	2.22	0.95
50–54	1,991	6.42	1.49	1,953	5.54	1.79	1,434	6.14	1.90	414	2.23	0.85
55–59	3,320	6.30	1.60	3,242	5.46	1.83	609	6.05	1.91	922	2.19	0.92
60–64	3,419	6.10	1.66	3,328	5.26	1.91	65	5.28	2.12	1,142	2.12	0.93
≥ 65	9,593	5.15	1.78	8,639	4.32	1.88	9,780	5.48	2.14	7,877	1.61	0.99

Note. Immediate Recall, Delayed Recall, and Vocabulary measures were administered during the 1998 data collection wave; Number Series was administered as part of the 2012 data collection. Gc = crystallized intelligence; Gf = fluid intelligence.

Table A2
Standardized Age Group Mean-Score Differences in Cognitive Ability (Sample 3)

	Immediate Recall (memory: short term, working)			Delayed Recall (memory: long term, storage and retrieval)			Vocabulary (Gc: verbal)			Number Series (Gf: inductive)		
	25–44 (n = 248)			25–44 (n = 247)			25–44 (n = 137)			35–49 (n = 130)		
Age group	n_{comp}	*d*	90% CI	n_{comp}	*d*	90% CI	n_{comp}	*d*	90% CI	n_{comp}	*d*	90% CI
45–49	452	0.12	[−0.01, 0.25]	448	0.07	[−0.06, 0.20]	265	−0.11	[−0.28, 0.07]			
50–54	1,991	0.18	[0.07, 0.29]	1,953	0.12	[0.01, 0.23]	1,434	−0.66	[−0.81, −0.51]	414	0.05	[−0.12, 0.21]
55–59	3,320	0.25	[0.14, 0.35]	3,242	0.16	[0.06, 0.27]	609	−0.61	[−0.76, −0.45]	922	0.09	[−0.06, 0.25]
60–64	3,419	0.35	[0.25, 0.46]	3,328	0.26	[0.15, 0.37]	65	−0.19	[−0.44, 0.05]	1,142	0.17	[0.02, 0.32]
≥ 65	9,593	0.87	[0.76, 0.97]	8,639	0.76	[0.66, 0.87]	9,780	−0.28	[−0.43, −0.14]	7,877	0.68	[0.53, 0.82]

Note. Age ranges across the top of the table are the reference group; age ranges in the first column are for the comparison group. Immediate Recall, Delayed Recall, and Vocabulary measures were administered during the 1998 data collection wave; Number Series was administered as part of the 2012 data collection. n_{comp} = comparison group sample size; *d* = standardized group mean-score differences (Cohen's *d*) compared with the reference group (positive effect sizes indicate that younger individuals scored higher on average); CI = confidence interval (two-tailed); Gc = crystallized intelligence; Gf = fluid intelligence.

(*Appendices continue*)

Appendix B

Generalizability—Results for Sample 4

Table B1
Summary of Descriptive Statistics by Age (Sample 4)

	Wechsler—Vocabulary (Gc: verbal)			Raven's APM (Gf: inductive)			Schaie–Thurston Letter Series (Gf: inductive)			Wechsler—Digit Symbol (processing speed: perceptual)		
Age group	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
25–34	57	58.40	9.76	57	8.63	3.81	57	18.86	5.92	57	63.17	11.15
35–44	48	60.50	13.30	47	8.40	3.94	47	18.00	6.07	47	60.36	11.23
45–54	64	62.58	13.21	62	6.84	3.28	63	15.70	6.74	63	56.13	11.47
55–64	53	62.11	14.79	49	5.76	3.17	52	14.77	6.75	53	51.63	11.35
≥ 65	41	62.54	12.65	37	4.16	2.48	37	9.97	5.45	39	43.71	12.12

Note. All data stem from the 1996 data collection wave. APM = Advanced Progressive Matrices; Gc = crystallized intelligence; Gf = fluid intelligence.

Table B2
Standardized Age Group Mean-Score Differences in Cognitive Ability (Sample 4)

	Wechsler—Vocabulary (Gc: verbal)			Raven's APM (Gf: inductive)			Schaie−Thurston Letter Series (Gf: inductive)			Wechsler—Digit Symbol (processing speed: perceptual)		
	25–34 (n = 57)			25–34 (n = 57)			25–34 (n = 57)			25–34 (n = 57)		
Age group	n_{comp}	*d*	90% CI	n_{comp}	*d*	90% CI	n_{comp}	*d*	90% CI	n_{comp}	*d*	90% CI
35–44	48	−0.18	[−0.50, 0.14]	47	0.06	[−0.26, 0.38]	47	0.14	[−0.18, 0.47]	47	0.25	[−0.07, 0.58]
45–54	64	−0.36	[−0.66, −0.06]	62	0.51	[0.20, 0.81]	63	0.50	[0.19, 0.80]	63	0.62	[0.31, 0.93]
55–64	53	−0.30	[−0.61, 0.02]	49	0.82	[0.48, 1.15]	52	0.65	[0.32, 0.97]	53	1.03	[0.69, 1.36]
≥ 65	41	−0.37	[−0.71, −0.03]	37	1.34	[0.95, 1.72]	37	1.55	[1.16, 1.94]	39	1.68	[1.29, 2.08]

Note. Age ranges across the top of the table are the reference group; age ranges in the first column are for the comparison group. n_{comp} = comparison group sample size; *d* = standardized group mean-score differences (Cohen's *d*) compared with the reference group (positive effect sizes indicate that younger individuals scored higher on average); CI = confidence interval (two-tailed); APM = Advanced Progressive Matrices; Gc = crystallized intelligence; Gf = fluid intelligence.

Received July 31, 2013
Revision received September 8, 2014
Accepted November 18, 2014 ■