Analysis of Adverse Impact for the Hogan Personality Inventory, Hogan Development Survey, and the Motives, Values, Preferences Inventory

Documentation of Psychometric and Research Evidence

HOGAN RESEARCH DIVISON

2012

THE SCIENCE OF PERSONALITY

Executive Summary

In this paper, we define Adverse Impact (AI) and provide empirical evidence for a lack of AI in personnel selection applications using the Hogan Personality Inventory (HPI), Hogan Development Survey (HDS), and the Motives, Values, Preferences Inventory (MVPI).

Defining Adverse Impact

1. The *Uniform Guidelines on Employee Selection Procedures* (UGESP, 1978) presents the four-fifths, or eighty percent rule, for examining AI based on sex or race/ethnicity. This rule has also been adopted for examining potential age discrimination (see The Age Discrimination in Employment Act of 1967 [ADEA], 1967).

2. Researchers have proposed alternative methods for examining AI, although none have been as widely adopted as the four-fifths rule.

3. A statistical significance test for mean group differences on an individual assessment scale used in a selection profile does *not* provide evidence for AI.

Neither meaningful mean group differences nor AI is evident in selection profiles created using the HPI, HDS, or MVPI.

1. Statistically significant mean group differences on HPI, HDS, or MVPI scales do not indicate AI and are not practically meaningful as indicated by effect sizes.

2. There is no evidence of AI from selection profiles using HPI scales across seven job families encompassing all U.S. occupations.

3. There is no evidence of AI from selection profiles using HPI, HDS, or MVPI scales across multiple selection systems.

To date, no operational selection profile using the HPI, HDS, or MVPI has demonstrated AI, and no claims of unfair employment discrimination have resulted from an employer's use of Hogan assessments.

Analysis of Adverse Impact for the Hogan Personality Inventory, Hogan Development Survey, and the Motives, Values, Preferences Inventory

Defining Adverse Impact

The *Uniform Guidelines on Employee Selection Procedures* (UGESP, 1978) defines AI as "a substantially different rate of selection in hiring, promotion, or other employment decisions which works to the disadvantage of members of a race, sex or ethnic group" (see section 1607.16). Furthermore, in examining the potential of AI, the UGESP outlines the four-fifths rule, stating that the "selection rate for any race, sex or ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact." (1978, see section 1607.4 D). Courts have also applied this rule to cases involving age discrimination. The Age Discrimination in Employment Act (ADEA) of 1967 prohibited discrimination in selection contexts against individuals 40 years of age or older.

An employer is not required to conduct validity studies of selection procedures where no AI results. Nevertheless, best professional practices encourage an examination of the potential for AI and the accumulation of validation evidence for each step of any selection process. Furthermore, a statistical significance test for mean group differences on individual assessment scales is often informative, but does not provide evidence of AI when a selection profile includes multiple assessment scales. For example, Hogan typically creates selection profiles using multiple HPI, HDS, and MVPI scales. In such cases, organizations must examine AI at the overall profile level, or the point at which selection decisions are made, rather than examining differences on individual scales within the profile.

Examining Adverse Impact

AI can be very costly to organizations. In examining AI cases from 1998 to 2008, Williams (2010) found that fees for out-of-court settlements averaged $590,266 for Equal Employment Opportunity Commission (EEOC) cases and $668,785 for Office of Federal Contract Compliance Program (OFCCP) cases. Similarly, in cases where individuals filed discrimination complaints outside the EEOC or OFCCP, compensation averaged $12,292,492 for cases settled out of court and $13,306,346 for court rulings favoring the plaintiff (Williams, 2010).

The implications of AI are not only financial, but may also impact the image of the organization. For example, although the *Dukes v. Wal-Mart* (2010) ruling favored the retail giant in a sex discrimination case, it still cost the organization 10 years of litigation and bad press. Conversely, the more positive an organization's image, the more attractive it is to both applicants (Chapman, Uggerslev, Carroll, Piasentin, & Jones, 2005) and customers (Keh & Xie, 2009).

In comparing methods for examining AI, the Technical Advisory Committee on Testing (TACT) of the California Fair Employment Practices Commission adopted the four-fifths rule as a “trigger rule” to avoid the complexities of statistical significance testing, which may be difficult to understand by people affected by the decision rule (as cited in Roth, Bobko, & Switzer, 2006). The use of the four-fifths rule in selection contexts is appropriate because it is (a) based on selection rates, (b) not affected by large sample sizes, and (c) often simpler to understand than many statistical tests. The four-fifths rule indicates whether the impact ratio (i.e., the difference in the selection rates of two groups) is large enough to be of practical concern (Morris & Lobsenz, 2000). That is, AI most commonly occurs when the differences between two groups’ scores in one or more parts of a selection procedure are large enough to be of practical concern, resulting in the organization selecting one group at a substantially lesser rate than another group.

Organizations may use the four-fifths rule to monitor the effects of their own selection processes. Also, the federal government might use it to determine compliance or if enforcement is necessary when civil actions occur (Roth et al., 2006). A selection procedure is in compliance when “such use has been validated in accord with these guidelines” (UGESP, 1978, see section 1607.16 C).

Despite their prevalence, both the Uniform Guidelines and the four-fifths rule have their critics. Recently, debate has emerged concerning the need to update the Guidelines to reflect current scientific knowledge and standards. Much of this debate has focused on the types of validity evidence considered, the statistical definition of adverse impact, and the UGESP not being updated in over 30 years (e.g., Jacobs, Deckert, & Silva, 2011; McDaniel, Kepes, & Banks, 2011). Furthermore, researchers and practitioners have suggested a number of alternatives to the four-fifths rule (e.g., moderated multiple regression, one-person rule, and the N of 1 rule). We examine several of these alternative methods below.

Moderated multiple regression (MMR) provides a means for examining AI using the slopes and intercepts at any given cutoff score in a selection procedure. MMR generates a predictive equation to examine relationships between a selection procedure and performance ratings. By including demographic variables in these equations, MMR helps analysts identify small differences in how well a selection procedure predicts performance across groups. Unfortunately, MMR results can be influenced by a number of factors associated with statistical power, unequal error variance, unequal sample sizes, or variable reliability (Hough, Oswald, & Ployhart, 2001). Also, it provides no direct means for examining the practical significance of group differences (Hough et al., 2001).

Another alternative is the one-person rule, which is based on a comparison of the number of expected and actual minority hires. When using it, one calculates the expected number of minority hires by multiplying the number of minority applicants by the overall selection ratio, and then rounds this product down to the nearest whole number. According to this rule, a difference of one or more between expected and actual minority hires provides evidence of potential AI (Roth, Bobko, & Switzer, 2006). Although one advantage of the one-person rule is that it does not require large samples, current research indicates it is unlikely to be used or to be persuasive in court (Roth et al., 2006).

Finally, the *Adoption of Questions and Answers to Clarify and Provide a Common Interpretation of the Uniform Guidelines on Employee Selection Procedures* (Adoption of Questions, 1979) presents the *N* of 1

rule as a follow-up to the four-fifths rule with small samples. This rule states that even if the four-fifths rule is violated, differences could have occurred by chance if the selection ratio is reversed after reducing the number of majority hires by one and increasing the number of minority hires by one (Roth et al., 2006). According to this document, when such a reversal occurs, “there is no requirement for additional validity evidence or expectation of enforcement actions unless the difference in selection rates continues for a period of time” (Adoption of Questions, 1979, see question 21).

Although these methods address potential limitations of the four-fifths rule, they may be best suited for self monitoring rather than an evaluation of potential legal liability. In contrast, the four-fifths rule is more common, relatively easy to use, and less ambiguous than these alternatives (Morris & Lobsenz, 2000).

Another important consideration is that factors independent of selection procedures may affect selection ratios, such as the number of qualified minorities available in the current applicant pool. Accordingly, selection procedures need to be monitored continuously and re-evaluated to ensure compliance. With regards to the HPI, HDS, and MVPI, there have been no demonstrated cases of AI and no claims of employment discrimination to date using these assessments.

Adverse Impact and the Courts

Courts most commonly rely on the four-fifths rule for assessing AI (Risavy, & Hausdorf, 2011). As such, an organization’s best course of action is to first validate a selection procedure, and then record and monitor applicant information to ensure compliance. Although the UGESP only requires a validation study when selection procedures show evidence of AI (Hough et al., 2001), using a validated assessment ensures that a selection system identifies high potential job applicants effectively.

Legal challenges are most likely to result when organizations fail to apply selection procedures consistently across applicants or properly validate procedures prior to their use (Williams, 2010). For example, unstructured interviews account for 57% of federal court cases involving to AI (Terpstra, Mohamed & Kethley, 1999). Unstructured interviews, although common, are often informal conversations that lack consistency on a variety of factors such as the questions asked, the interview duration and setting, and the method used to rate responses. Conversely, assessments such as the HPI, HDS, and MVPI are consistent in both administration and scoring and, when used appropriately, supported by documented validity evidence.

Adverse Impact and Assessment Tools

AI is dependent on several factors relating to both mean group differences in assessment scores and how organizations use those scores to make selection decisions. The most legally defensible selection procedures use assessments that are both job-related and demonstrate the smallest mean group differences. The failure to meet either of these criteria increases an organization’s legal liability.

Many organizations use cognitive ability tests as part of their selection processes. Research demonstrates the predictive validity of cognitive ability assessments across industries, organizations,

and jobs (Schmidt & Hunter, 2004). However, cognitive ability tests often result in large mean group differences. For example, Blacks typically score one standard deviation below Whites on cognitive ability assessments (Hough et al., 2001). In *EEOC v. NationsBank of Tennessee* (2001), the U.S. District Court, Eastern District of Tennessee, found an assessment of cognitive ability to be discriminatory against Hispanics, who typically score one-half of a standard deviation below Whites (Hough et al., 2001). Despite the fact that the cognitive test was predictive of job performance, it required a level of English proficiency not required for the job that resulted in adverse impact against Hispanics. One method for avoiding AI with cognitive ability assessments is to use cutoff scores allowing most applicants to pass (e.g., 90% or higher), although this tends to also reduce their overall utility because they remove so few applicants from the selection process (Pyburn, Ployhart, & Kravitz, 2008).

Other commonly used selection methods have also failed legal challenges. For example, in *EEOC v. American Airlines* (2002), the airline settled because a measure of physical strength resulted in AI against women and the airline had set the cutoff score at a higher level than the job required. In a similar case, *EEOC v. The Dial Corporation* (2006), the U.S. District Court, Southern District of Iowa, found that a physical performance test discriminated against women not only because the test resulted in AI, but because some women who passed were still rejected based on assessment administrator comments.

In contrast, personality assessments have fared better than these selection measures. One reason is that minority groups often score slightly higher than members of majority groups on a number of personality measures (Ng & Sears, 2010). Unlike most assessments, personality typically does not result in AI when using a reasonable passing rates (McDaniel, Kepes, & Banks, 2011), and mean group differences based on ethnicity and sex are often favorable to minority groups (Hough et al., 2001) Also, research has shown consistently that multiple personality scales predict job performance across industries, organizations, and jobs (Barrick & Mount, 1991; Oh, Wang, & Mount, 2011). In a review of U.S. federal court cases, Terpstra, Mohamed, and Kethley (1999) could not find a case between 1978 and 1997 where there was a legal challenge of personality assessments in a selection procedure. In a more recent work, Williams (2010) did not report a case between 1998 and 2008 of a legal challenge of personality assessments in a selection procedure.

Adverse Impact and the HPI, HDS, and MVPI

Next, we examine mean group differences based on age, sex, and race/ethnicity for the HPI, HDS, and MVPI. Then, we present AI results for sample selection profiles. We conducted HPI analyses using a normative sample of HPI data including cases from over 156,000 job incumbents and applicants representative of jobs from the U.S. workforce (Hogan & Hogan, 2007). We used a similar approach with HDS data, using a normative sample including cases from over 109,000 job incumbents and applicants representative of the HDS's intended population of managers, executives, professionals, and service and support personnel (Hogan & Hogan, 2009). Finally, our normative sample for the MVPI included cases from over 68,000 job incumbents and applicants representative of jobs from the U.S. workforce (Hogan & Hogan, 2010).

Section 1. Group Mean Differences on the HPI

Gender Differences

Tables 1.1 through 1.3 present comparisons of mean group differences across HPI scales. The mean scores for females are higher on Adjustment, Interpersonal Sensitivity, Prudence, and Learning Approach. Males have higher mean scores on Ambition, Sociability, and Inquisitive. Although Analysis of Variance (ANOVA) results indicate these differences are statistically significant, an examination of effect sizes shows this is primarily the result of large sample sizes. Based on traditional ranges for interpreting effect sizes (Cohen's *d*; Cohen, 1988), all differences are small.

Age Differences

Mean scores for those 40 years of age and younger are higher than those older than 40 on all seven HPI scales. Again, ANOVA results indicate these differences are statistically significant, although an examination of effect sizes indicates all differences are small.

Race/Ethnicity Differences

To examine differences based on race/ethnicity, we compared Whites, which most frequently serve as the majority group in the U.S., to all other groups. We used ANOVA and Dunnett post-hoc comparison results to examine statistical significance. The mean score for Whites was significantly higher in 8 of 28 comparisons (28.57%), significantly lower in 16 of 28 comparisons (57.14%), and similar enough to not reach statistical significance for the remaining 4 comparisons (14.29%). Of those that reached statistical significance, an examination of effect sizes shows all differences are small.

Table 1.1
Mean HPI Scale Differences by Gender

HPI scale	*Gender*		
	Male (n = 60,722)	Female (n = 60,730)	Effect size
Adjustment	31.16 (4.69)	31.27* (4.69)	0.02
Ambition	26.29 (3.19)	25.54* (3.48)	−0.22
Sociability	14.61 (4.72)	14.02* (4.60)	−0.13
Interpersonal Sensitivity	20.19 (1.81)	20.80* (1.38)	0.37
Prudence	22.95 (4.03)	23.83* (3.69)	0.23
Inquisitive	17.23 (4.37)	15.90* (4.56)	−0.30
Learning Approach	9.87 (3.09)	10.78* (2.74)	0.31

Note . All analyses conducted with the 2005 HPI norming sample (n = 121,452). Group means for each HPI scale are presented with the standard deviations in parentheses. Negative effect sizes indicate higher scale means for the male group. * = $p < .05$.

Table 1.2
Mean HPI Scale Differences by Age

HPI scale	*Age*		
	40 and younger (n = 42,679)	Over 40 (n = 18,213)	Effect size
Adjustment	31.00 (4.81)	30.27* (5.22)	−0.15
Ambition	25.95 (3.41)	25.46* (3.64)	−0.14
Sociability	14.63 (4.69)	12.81* (4.84)	−0.38
Interpersonal Sensitivity	20.39 (1.68)	20.03* (2.07)	−0.20
Prudence	23.08 (4.00)	22.63* (3.95)	−0.11
Inquisitive	16.80 (4.44)	15.16* (4.50)	−0.36
Learning Approach	9.94 (2.99)	9.17* (3.23)	−0.25

Note . All analyses conducted with the 2005 HPI norming sample (n = 60,892). Group means for each HPI scale are presented with the standard deviations in parentheses. Negative effect sizes indicate higher scale means for the 40 and younger group. * = $p < .05$.

Table 1.3
Mean HPI Scale Differences by Ethnicity

HPI scale	*Ethnicity*								
	White (n = 72,975)	Black (n = 13,006)	Effect size	Hispanic (n = 15,034)	Effect size	Asian/PI (n = 5,067)	Effect size	Native American (n = 2,208)	Effect size
Adjustment	31.24 (4.75)	31.60* (4.27)	0.08	31.89* (4.04)	0.14	30.54* (4.66)	−0.15	31.12 (4.70)	−0.03
Ambition	25.85 (3.49)	26.42* (2.75)	0.17	26.07* (2.95)	0.07	25.53* (3.41)	−0.09	25.68 (3.36)	−0.05
Sociability	14.54 (4.72)	13.14* (4.61)	−0.30	14.06* (4.44)	−0.10	14.89* (4.32)	0.07	14.64 (4.44)	0.02
Interpersonal Sensitivity	20.59 (1.61)	20.42* (1.49)	−0.11	20.58 (1.42)	−0.01	20.25* (1.74)	−0.21	20.50* (1.62)	−0.06
Prudence	23.22 (3.89)	24.24* (3.64)	0.27	24.31* (3.63)	0.28	23.56* (3.80)	0.09	23.81* (3.84)	0.15
Inquisitive	16.46 (4.54)	16.08* (4.40)	−0.08	17.17* (4.45)	0.16	17.72* (4.32)	0.28	17.86* (4.34)	0.31
Learning Approach	10.19 (2.98)	10.72* (2.88)	0.18	10.88* (2.75)	0.23	10.80* (2.77)	0.20	10.90* (2.75)	0.24

Note . All analyses conducted with the 2005 HPI norming sample (n = 108,290). Group means for each HPI scale are presented with the standard deviations in parentheses. Negative effect sizes indicate higher scale means for the White group. * = $p < .05$.

Section 2. Group Mean Differences on the HDS

Gender Differences

Tables 2.1 through 2.3 present comparisons of mean group differences across HDS scales. The mean scores for females are higher on Excitable, Skeptical, Cautious, Leisurely, Bold, Imaginative, Diligent, and Dutiful. Males have higher mean scores on Reserved, Mischievous, and Colorful. Although ANOVA results indicate these differences are statistically significant, an examination of effect sizes shows all differences are small.

Age Differences

Mean scores for those above the age of 40 are higher than those 40 or younger on Excitable and Cautious. Those 40 or younger have higher mean scores on Skeptical, Leisurely, Bold, Mischievous, Colorful, Imaginative, Diligent, and Dutiful. Again, ANOVA results indicate these differences are statistically significant, although effect sizes are small.

Race/Ethnicity Differences

To examine differences based on race/ethnicity, we again compared Whites to all other groups. We used ANOVA and Dunnett post-hoc comparison results to examine statistical significance. The mean score for Whites was significantly higher in 6 of 44 comparisons (13.64%), significantly lower in 29 of 44 comparisons (65.91%), and similar enough to not reach statistical significance for the remaining 9 comparisons (20.45%). Of those that reached statistical significance, all but one of the effect sizes are small. The only exception is the moderate difference between the White and Asian/Pacific Islander groups on Skeptical.

Table 2.1
Mean HDS Scale Differences by Gender

HDS scale	*Gender*		
	Male (n = 52,607)	Female (n = 50,039)	Effect size
Excitable	2.72 (2.25)	2.87* (2.23)	0.07
Skeptical	4.14 (2.32)	4.25* (2.28)	0.05
Cautious	2.63 (2.25)	2.96* (2.42)	0.14
Reserved	4.11 (2.10)	3.87* (1.89)	−0.12
Leisurely	4.34 (2.04)	4.58* (1.89)	0.12
Bold	7.45 (2.70)	7.68* (2.60)	0.09
Mischievous	5.85 (2.55)	5.15* (2.36)	−0.28
Colorful	7.34 (2.86)	7.14* (2.59)	−0.07
Imaginative	5.20 (2.43)	5.30* (2.44)	0.04
Diligent	9.58 (2.17)	10.00* (1.98)	0.20
Dutiful	7.87 (2.08)	8.50* (2.07)	0.30

Note. All analyses conducted with the 2009 HDS norming sample (n = 102,646). Group means for each HDS scale are presented with the standard deviations in parentheses. Negative effect sizes indicate higher scale means for the male group. * = $p < .05$.

Table 2.2
Mean HDS Scale Differences by Age

HDS scale	*Age*		
	40 and younger (n = 38,515)	Over 40 (n = 31,972)	Effect size
Excitable	2.73 (2.16)	2.79* (2.15)	0.03
Skeptical	4.37 (2.23)	3.67* (2.15)	−0.31
Cautious	2.49 (2.21)	2.85* (2.33)	0.16
Reserved	3.93 (1.91)	3.92 (2.01)	0.00
Leisurely	4.53 (1.88)	4.30* (1.92)	−0.12
Bold	7.84 (2.61)	7.05* (2.66)	−0.29
Mischievous	5.46 (2.45)	5.11* (2.43)	−0.14
Colorful	7.37 (2.61)	6.89* (2.78)	−0.18
Imaginative	5.30 (2.42)	4.87* (2.39)	−0.18
Diligent	10.09 (1.92)	9.69* (2.11)	−0.20
Dutiful	8.53 (2.00)	8.04* (2.04)	−0.24

Note. All analyses conducted with the 2009 HDS norming sample (n = 70,487). Group means for each HDS scale are presented with the standard deviations in parentheses. Negative effect sizes indicate higher scale means for the 40 and younger group. * = $p < .05$.

Table 2.3
Mean HDS Scale Differences by Ethnicity

HDS scale	*Ethnicity*								
	White (n = 45,677)	Black (n = 12,696)	Effect size	Hispanic (n = 13,897)	Effect size	Asian/PI (n = 3,846)	Effect size	Native American (n = 144)	Effect size
Excitable	2.77 (2.25)	2.87* (1.87)	0.05	2.88* (2.03)	0.05	3.12* (2.43)	0.16	2.74 (2.56)	−0.01
Skeptical	3.75 (2.17)	4.67* (2.16)	0.42	4.81* (2.36)	0.47	5.07* (2.50)	0.59	4.20 (2.31)	0.21
Cautious	2.86 (2.39)	2.16* (1.86)	−0.30	2.55* (2.13)	−0.13	3.27* (2.43)	0.17	3.36* (2.68)	0.21
Reserved	3.79 (1.93)	4.33* (1.94)	0.28	4.13* (1.94)	0.17	4.06* (2.09)	0.14	4.01 (2.07)	0.11
Leisurely	4.28 (1.92)	4.86* (1.71)	0.31	4.68* (1.81)	0.21	5.00* (2.10)	0.37	4.46 (2.19)	0.09
Bold	7.07 (2.64)	8.34* (2.45)	0.48	8.25* (2.63)	0.44	8.24* (2.75)	0.44	7.48 (2.64)	0.16
Mischievous	5.29 (2.47)	5.08* (2.32)	−0.08	5.26 (2.37)	−0.01	5.76* (2.57)	0.19	6.03* (2.48)	0.30
Colorful	7.13 (2.78)	7.13 (2.37)	0.00	6.95* (2.53)	−0.07	7.10 (2.85)	−0.01	7.63 (2.97)	0.18
Imaginative	4.96 (2.39)	5.22* (2.40)	0.11	5.37* (2.43)	0.17	6.05* (2.51)	0.45	5.65* (2.58)	0.29
Diligent	9.72 (2.08)	10.49* (1.64)	0.38	10.40* (1.69)	0.34	10.25* (2.05)	0.26	9.24* (2.17)	−0.23
Dutiful	8.15 (2.06)	8.81* (1.85)	0.33	8.87* (1.94)	0.35	8.55* (2.22)	0.19	7.72* (1.94)	−0.21

Note. All analyses conducted with the 2009 HDS norming sample (n = 76,260). Group means for each HDS scale are presented with the standard deviations in parentheses. Negative effect sizes indicate higher scale means for the White group. * = $p < .05$.

Section 3. Group Mean Differences on the MVPI

Gender Differences

Tables 3.1 through 3.3 present comparisons of mean group differences across MVPI scales. The mean scores for females are higher on Aesthetic, Affiliation, Altruistic, Hedonistic, and Security. Males have higher mean scores on Commercial, Power, Recognition, Scientific, and Tradition. Although ANOVA results indicate these differences are statistically significant, an examination of effect sizes shows all differences are small.

Age Differences

Mean scores for those over the age of 40 are higher than those 40 and younger on Aesthetic and Tradition. Those 40 and younger have higher mean scores on Affiliation, Altruistic, Commercial, Hedonistic, Power, Recognition, Scientific, and Security. Again, ANOVA results indicate these differences are statistically significant, although effect sizes are small.

Race/Ethnicity Differences

To examine differences based on race/ethnicity, we compared Whites to all other groups for each MVPI scale. We used ANOVA and Dunnett post-hoc comparison results to examine statistical significance. The mean score for Whites was significantly higher in 6 of 40 comparisons (15.00%), significantly lower in 31 of 40 comparisons (77.50%), and similar enough to not reach statistical significance for the remaining 3 comparisons (7.50%). Of those that reached statistical significance, all but one of the effect sizes are small. The only exception is the moderate difference between the White and Black groups on Security.

Table 3.1
Mean MVPI Scale Differences by Gender

MVPI scale	*Gender*		
	Male (n = 36,295)	Female (n = 27,391)	Effect size
Aesthetic	34.34 (7.70)	36.53* (8.01)	0.28
Affiliation	50.17 (5.07)	50.61* (4.62)	0.09
Altruistic	49.75 (6.31)	51.71* (5.57)	0.32
Commercial	47.24 (5.76)	45.11* (6.14)	−0.35
Hedonistic	38.14 (6.74)	38.92* (6.81)	0.12
Power	49.58 (5.98)	47.93* (6.43)	−0.26
Recognition	41.94 (8.16)	41.78* (7.78)	−0.02
Scientific	42.57 (7.80)	39.82* (7.89)	−0.35
Security	41.46 (7.15)	42.40* (6.90)	0.13
Tradition	48.09 (6.31)	47.82* (5.97)	−0.04

Note . All analyses conducted with the 2010 MVPI norming sample (n = 63,686). Group means for each MVPI scale are presented with the standard deviations in parentheses. Negative effect sizes indicate higher scale means for the male group. * = $p < .05$.

Table 3.2
Mean MVPI Scale Differences by Age

MVPI scale	*Age*		
	40 and younger (n = 20,086)	Over 40 (n = 16,575)	Effect size
Aesthetic	34.38 (7.64)	35.26* (7.83)	0.11
Affiliation	50.72 (4.93)	48.88* (5.34)	−0.35
Altruistic	50.13 (6.28)	49.49* (6.31)	−0.10
Commercial	46.31 (6.06)	44.67* (6.15)	−0.27
Hedonistic	39.81 (6.89)	37.30* (6.62)	−0.36
Power	49.15 (6.18)	47.11* (6.65)	−0.32
Recognition	42.45 (7.97)	39.51* (7.82)	−0.37
Scientific	41.93 (8.18)	40.16* (8.04)	−0.22
Security	41.30 (7.09)	40.13* (7.25)	−0.16
Tradition	47.58 (6.32)	48.46* (6.29)	0.14

Note . All analyses conducted with the 2010 MVPI norming sample (n = 36,661). Group means for each MVPI scale are presented with the standard deviations in parentheses. Negative effect sizes indicate higher scale means for the 40 and younger group. * = $p < .05$.

Table 3.3
Mean MVPI Scale Differences by Ethnicity

MVPI scale	*Ethnicity*								
	White (n = 35,794)	Black (n = 4,718)	Effect size	Hispanic (n = 5,205)	Effect size	Asian/PI (n = 3,033)	Effect size	Native American (n = 541)	Effect size
Aesthetic	35.00 (7.93)	35.39* (7.76)	0.05	35.54* (8.05)	0.07	36.43* (8.19)	0.18	37.00* (8.36)	0.25
Affiliation	50.33 (4.96)	49.98* (4.29)	−0.07	51.28* (3.97)	0.20	50.63* (5.12)	0.06	50.63 (4.52)	0.06
Altruistic	50.41 (6.07)	52.35* (5.13)	0.32	51.60* (5.56)	0.20	51.28* (6.13)	0.14	51.67* (5.61)	0.21
Commercial	46.06 (6.00)	48.42* (5.53)	0.39	47.85* (5.68)	0.30	47.47* (6.14)	0.23	47.72* (5.71)	0.28
Hedonistic	38.30 (6.81)	37.37* (6.62)	−0.14	38.01* (6.63)	−0.04	40.49* (6.72)	0.32	37.69 (6.61)	−0.09
Power	48.72 (6.26)	49.56* (5.90)	0.14	50.13* (5.69)	0.23	50.02* (6.14)	0.21	49.84* (5.72)	0.18
Recognition	41.53 (7.96)	42.64* (8.02)	0.14	43.29* (8.01)	0.22	44.98* (8.13)	0.43	42.29 (8.16)	0.10
Scientific	41.35 (7.98)	40.89* (7.63)	−0.06	42.36* (7.56)	0.13	43.94* (7.99)	0.32	43.30* (7.55)	0.24
Security	41.46 (6.98)	46.02* (5.93)	0.65	44.61* (6.16)	0.45	42.41* (6.78)	0.14	43.10* (6.62)	0.22
Tradition	48.08 (6.29)	49.82* (5.48)	0.28	47.46* (5.73)	−0.10	46.71* (5.82)	−0.22	48.84* (6.05)	0.12

Note . All analyses conducted with the 2010 MVPI norming sample (n = 49,291). Group means for each MVPI scale are presented with the standard deviations in parentheses. Negative effect sizes indicate higher scale means for the White group. * = $p < .05$.

Section 4. Adverse Impact Analysis in the Operational Use of the HPI, HDS, and MVPI

An examination of mean group differences shows all differences tend to be small to moderate in magnitude and often benefit minority groups. However, as previously stated, mean differences are not indicative of AI. When creating selection profiles, Hogan uses cutoff scores on multiple scales that predict performance for a specific job. As such, we examine AI for each profile. Below, we provide two examples of selection profiles using Hogan scales, one that uses the HPI and one that uses a combination of scales across the three assessments. These profiles are drawn from the Hogan Express Report (Hogan Assessment Systems, 2009a) and High Potential Report (Hogan Assessment Systems, 2009b) technical manuals.

The Hogan Express Report predicts performance for managers and executives using cutoff scores on four HPI scales. High performing managers and executives tend to be calm and stable (Adjustment), competitive and achievement oriented (Ambition), conscientious and rule-following (Prudence), and friendly and agreeable (Interpersonal Sensitivity). Table 4.1 displays these cutoff scores and Table 4.2 presents selection ratios based on this profile. As shown in these tables, this profile helps identify high potential managers and executives without resulting in AI.

Table 4.1
Recommended Cutoff Scores for Managers and Executives Jobs

Scale	*Below Requirements*	*Meets Requirement*
Adjustment	Fails to Satisfy **Meets Requirements** Scores	≥ 12%
Ambition		≥ 11%
Prudence		≥ 12%
Interpersonal Sensitivity		≥ 20%
Expected Pass Rates		73.90 %

Note. Cutoff scores were drawn from Express technical manual (Hogan Assessment Systems, 2009a) for managers and executives (n = 4,523).

Table 4.2
Selection Ratios and Adverse Impact for Managers and Executives Jobs Using Recommended Cutoff Scores

Demographic Groups		*AI Ratio*	*AI Result*
Sex	Men		
	Women	0.98	No AI
Age	≤ 40		
	>40	1.04	No AI
Race	Black	1.00	No AI
	Hispanic	1.00	No AI
	Asian/PI	0.94	No AI
	Native American	1.10	No AI
	White		

Note. Analyses were drawn from Express technical manual (Hogan Assessment Systems, 2009a) for managers and executives (n =4,487). Numbers greater than 1.00 indicate that the results for minority group applicants fall within acceptable ranges more frequently than the majority group. Ratios below 1.00 indicate that the minority group fall within the acceptable ranges less frequently than the majority group. AI would exist when the ratio is less than 0.80.

Hogan's High Potential Report includes a number of competency-based profiles for predicting job performance for high level professionals and managers/executives. These profiles include cutoff scores on HPI, HDS, and MVPI scales (Hogan Assessment Systems, 2009b). One competency included in the report is Strategic Self-Awareness, or recognizing strengths and weaknesses and using that information to guide personal growth and development. Individuals who receive high scores on Strategic Self-Awareness tend to stay calm under pressure and adjust well to fast-paced environments (HPI Adjustment), accept feedback and avoid excessive self-promotion (HDS Bold), and are comfortable with ambiguity (MVPI Security). Table 4.3 displays cutoff scores for Strategic Self-Awareness and Table 4.4 presents selection ratios based on this profile. As shown in these tables, this profile helps identify high potential managers and executives without resulting in AI.

Table 4.3
Recommended Cutoff Scores for Strategic Self-Awareness

Scale	*Below Requirements*	*Meets Requirement*
HPI Adjustment		≥ 26% & ≤ 98%
HDS Bold	Fails to Satisfy **Meets Requirements** Scores	≥ 15% & ≤ 95%
MVPI Security		≤ 95%
Expected Pass Rates		50.80%

Note . Cutoff scores were drawn from High Potential technical manual (Hogan Assessment Systems, 2009b) for managers and executives (n = 246).

Table 4.4
Selection Ratios and Adverse Impact for Strategic Self-Awareness Using Recommended Cutoff Scores

Demographic Group		*AI Ratio*	*AI Result*
Sex	Men		
	Women	1.06	No AI
Age	≤ 40		
	>40	1.00	No AI
Race	Black	0.98	No AI
	Hispanic	0.96	No AI
	Asian/PI	1.04	No AI
	Native American		
	White		

Note . Analyses were drawn from High Potential technical manual (Hogan Assessment Systems, 2009b) for managers and executives (n = 3,974). Analyses were not conducted for Native Americans due to small sample sizes. Numbers greater than 1.00 indicate that the results for minority group applicants fall within acceptable ranges more frequently than the majority group. Ratios below 1.00 indicate that the minority group fall within the acceptable ranges less frequently than the majority group. AI would exist when the ratio is less than 0.80.

Section 5. Conclusions and Considerations

The use of assessments such as the HPI, HDS, and MVPI helps to identify high potential job applicants while reducing the probability of AI. Even when using properly validated assessments, organizations should monitor their selection procedures continually. For example, when using a top-down selection procedure, the four-fifths rule may be violated even when no subgroup differences exist (Roth et al., 2006). Also, selection procedures need to be consistent across applicants. Structured assessments are less likely to be successfully challenged in court as compared to inconsistent selection procedures (Williams, 2010).

Next, to monitor AI, it is necessary to maintain current job descriptions and accurate applicant records (Williams, 2010). In *Dennis v. Columbia Medical Center* (2002), the court found in favor of the plaintiff because the organization's reasons for not hiring him did not match requirements outlined in the job description.

Finally, recruitment procedures should be appropriate for the job (Williams, 2010). That is, they should result in an applicant pool that matches the demographics of the area if possible. In *Allen v. Tobacco Superstore* (2007), the court found in favor of the plaintiff because the demographics of the current incumbent pool did not match the demographics of the area. The company relied on word-of-mouth recruitment, which resulted in a primarily White applicant pool in a predominantly Black community.

In conclusion, AI may result from a number of factors beyond the assessments an organization uses for selection. The continual monitoring and evaluating of selection procedures is necessary to ensure against AI. Also, proper validation in accordance with the UGESP not only helps ensure compliance, but ensures that a selection procedure identifies high potential job applicants. When creating selection profiles, Hogan not only adheres to all UGESP requirements by examining and reporting validity evidence, but examines the potential for AI using large, representative normative datasets.

References

Age Discrimination in Employment Act of 1976 [ADEA], Pub. L. No. 90-202, et seq.

Allen v. Tobacco Superstore (8th Cir. 2007) 475 F.3d 931.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.

Chapman, D. S., Uggerslev, K. L., Carroll, S. A., Piasentin, K. A., & Jones, D. A. (2005) Applicant attraction to organizations and job choice: A meta-analytic review of the correlates of recruiting outcomes. *Journal of Applied Psychology, 90*, 928-944.

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

Dennis v. Columbia Colleton Medical Center (4th Cir. 2002) 290 F.3d 639.

Dukes v. Wal-mart Stores, Inc. (CA9 2010) 603 F.3d 571.

EEOC v. American Airlines, 2002, Settled

EEOC v. NationsBank Tennessee, 2001, Settled

EEOC v. The Dial Corporation (8th Cir. 2006) F.3d 735.

Equal Opportunity Employment Commission, Civil Service Commission, U.S. Department of Labor, & U.S. Department of Justice. (1978). Uniform guidelines on employee selection procedures. *Federal Register, 43*, 38290-38309.

Equal Opportunity Employment Commission, Civil Service Commission, U.S. Department of Labor, & U.S. Department of Justice. (1979). Adoption of questions and answers to clarify and provide a common interpretation of the uniform guidelines on employee selection procedures. *Federal Register, 44*, 6570-6602.

Hogan Assessment Systems (2009a*). Hogan Express Report technical manual*. Tulsa, OK: Author.

Hogan Assessment Systems (2009b). *High Potential Candidate Assessment Report technical manual*. Tulsa, OK: Author.

Hogan, R., & Hogan, J. (2007). *Hogan Personality Inventory manual* (3rd ed.). Tulsa, OK: Hogan Assessment Systems.

Hogan, R., & Hogan, J. (2009). *Hogan Development Survey manual* (2nd ed.). Tulsa, OK: Hogan Assessment Systems.

Hogan, R., & Hogan, J. (2010). *Motives, Values, Preferences Inventory manual: 2010 administrative and norming updates*. Tulsa, OK: Hogan Assessment Systems

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*, 153-194.

Jacobs, R., Deckert, P. J., & Silva, J. (2011). Adverse impact is far more complicated than the uniform guidelines indicate. *Industrial and Organizational Psychology, Perspectives on Science and Practice, 4*, 558-561.

Keh, H. T., & Xie, Y. (2009). Corporate reputation and customer behavioral intentions: The roles of trust, identification and commitment. *Industrial Marketing Management, 38*, 732-742.

McDaniel, M. A., Kepes, S., & Banks, G. C. (2011). The uniform guidelines are a determent to the field of personnel selection. *Industrial and Organizational Psychology, Perspectives on Science and Practice, 4*, 494-514.

Morris, S. B., & Lobsenz, R. E. (2000). Significance tests and confidence intervals for the adverse impact ratio. *Personnel Psychology, 53*, 89-111.

Ng, E. S. W., & Sears, G. J. (2010). The effect of adverse impact in selection practices on organizational diversity: A field study. *The International Journal of Human Resource Management, 21*, 1454-1471.

Oh, I., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology, 96*, 762-773.

Outtz, J. L. (2011). Abolishing the uniform guidelines: Be careful what you wish for. Industrial and *Organizational Psychology, Perspectives on Science and Practice, 4*, 526-533.

Pyburn, K. M., Ployhart, R. E., Kravizt, D. A. (2008). The diversity-validity dilemma: Overview and legal context. *Personnel Psychology, 61*, 143-151.

Risavy, S. D., & Hausdorf, P. A. (2011). Personality testing in personnel selection: Adverse impact and differential hiring rates. *International Journal of Selection and Assessment, 13*, 18-30.

Roth P. L., Bobko, P., & Switzer, F. S. (2006). Modeling the behavior of the 4/5ths rule for determining adverse impact: Reasons for caution. *Journal of Applied Psychology, 91*, 507-522.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86*, 162-173.

Terpstra, D. E., Mohamed, A. A., & Kethley, R. B. (1999). An analysis of federal court cases involving nine selection devices. *International Journal of Selection and Assessment, 7*, 27-34.

Williams, K. (2010, April). Legal risk in selection: An analysis of processes and tools. Paper presented at the 25th annual conference for the Society for Industrial and Organizational Psychology, Atlanta, GA.