EEO INSiGHT®

Volume 5 : Issue 1

The Leading Resource for the EEO/AA Community

EEO INSIGHT

EEO INSIGHT®

The Leading Resource for the EEO/AA Community

Volume 5: Issue 1

EEO Insight® is published by Biddle Consulting Group, Inc., 193 Blue Ravine Road, Suite 270, Folsom, California 95630-4760. Electronic subscriptions (PDF e-mail delivery) are free for a limited time. Sign up online at *www.EEOInsight.com/subscribe.html.* Annual hardbound editions containing a compilation of published articles will be printed and made available for purchase. For more information on submitting an article for publication in EEO Insight, please contact Heather J. Patchell, Managing Editor, at editor@eeoinsight.com or (916) 294-4250 ext. 155.

ISSN: 1947-1548

Dear **READERS**

Around the cubicle farm, we are calling this issue of the *EEO Insight* our comp issue. This should come as no surprise to most of our readers in light of some very recent movement by the OFCCP. While, we are still awaiting some proposed changes on other topics, a new compensation Directive from the OFCCP has arrived! This new Directive and the modifications found therein have many Federal contractors wondering how this will affect their efforts with compliance as well as increase their workloads. Our featured article, *OFCCP's Compensation Directive 307: Should I Freak Out or is Everything Going to be Okay?*, gives the reader a practical review of how Federal contractor's should respond to the OFCCP's new compensation Directive.

This issue's two-part article, *How to Compute Liability When Statistical Evidence of Pay Disparities Exists*, offers the reader action-based steps for approaching compliance. The first article provides the reader with steps that organizations can take to identify compensation owed when disparities are found. The second article gives employers information on calculating the liabilities and gives three plausible methods for distribution, as this is a dynamic and complex issue.

Finally, *The Difference Between Differential Validity and Differential Prediction: Time to Clear Up the Mess*, discusses two relevant concepts in validation. This article highlights some of the confusion amongst the two concepts, but describes solutions for academicians and practitioners alike.

While this issue is very applied in nature, we are open to theoretical and applied articles and we always love hearing from our readers, and encourage you to contact us if you want to submit a response to something you've read in *EEO Insight*, have an idea for an article, or would like to submit an article of your own. Please contact me at *editor@eeoinsight.com*.

Heather J. Patchell
Managing Editor, *EEO Insight*

EEO INSiGHT

TABLE *of* CONTENTS

OFCCP's Compensation Directive 307:

A PRACTICAL REVIEW OF HOW FEDERAL CONTRACTORS SHOULD PREPARE FOR THE OFCCP'S NEW COMPENSATION DIRECTIVE

Jim Higgins, Ed.D. | *Biddle Consulting Group, Inc.*
Patrick M. Nooren, Ph.D. | *Biddle Consulting Group, Inc.*

THE HEAT IS ON!

Since the February 27, 2013 release of the latest iteration of the OFCCP's stated procedures for investigating compensation discrimination (*i.e.*, Directive 307 - effective February 28, 2013), there has been no shortage of webinars, blogs, and warnings from consultants and attorneys alike, providing insights into how this new directive will impact employers, and what they should do to get ahead of the inevitable tidal wave of compensation enforcement activity. Of course, if you've been around long enough, you will recognize this "heat" as being very similar to 2006, when the original *Standards*[1] and *Guidelines*[2] were first released. When that happened, many consultants and attorneys (present company included) went on the speaking circuit pontificating about how the "compensation sky was falling" and how employers had better start preparing for "compageddon."

Fast forward several years and it became abundantly clear that this forecasted apocalypse didn't occur. "Why not?" you may ask. The answer is multi-faceted and really came down to just a few key issues: 1) systemic compensation investigations are inherently time and resource-intensive and can take years to reach a resolution (and the OFCCP needs success/performance metrics "now"), and 2) the OFCCP is/was already very good at investigating adverse impact in hiring (in fact, it has long been their bread-and-butter). So, given the historical precedent, can employers anticipate the same false start this time around?

Not a Chance!

What employers need to understand is that the OFCCP *cannot* allow itself to fail this time. It's simply not acceptable to have large increases in both budget and staff (~$84M and 575 FTEs to ~$105M and 755 FTEs since 2009, respectively[3]), and *not* show that the money was well-spent (*i.e.*, commensurate increases in enforcement metrics). Couple this with the Obama Administration's continued focus on the "pay gap," the signing of the Lilly Ledbetter Equal Pay Act, the creation of the Equal Pay Enforcement Task Force, the focus on compensation enforcement in the OFCCP's congressional budget justification documents, and anyone can start to see the "writing on the walls."

So now that it's clear that things have to be different this time around, what exactly are the primary components of the new Directive and how are they different from what contractors have been subject to in audits over the past few years? Well, to anyone who has been "lucky" enough to have been involved in an audit recently, the Directive is going to appear awfully familiar. The reason is that the Agency has already been following the investigative strategy outlined in the Directive (with only a few notable, and substantive, exceptions, which we will discuss in detail later). The Directive itself simply served to codify the investigative approach as well as "put the final nail in the coffin" of the *Standards* and *Guidelines*.

PRIMARY COMPONENTS OF THE NEW COMPENSATION DIRECTIVE

According to the Directive, when investigating compensation, the Compliance Officer (CO) will:

- Conduct a preliminary analysis of summary data
- Conduct an analysis of individual employee-level data
- Determine the approach from a range of investigations and analytical tools
- Consider all employment practices that may lead to compensation disparities
- Develop pay analysis groups
- Investigate systemic, small group, and individual discrimination
- Review and test factors before accepting factors for analysis
- Conduct onsite investigation, offsite analysis, and refinement of the model

To the casual reader, the Directive, as written, appears to be a well-thought-out, step-by-step manual on how to properly investigate compensation disparities. Of course compensation discrimination isn't simply a systemic issue. Of course the OFCCP should be able to investigate individual differences. Of course the Agency should have a number of different analytical tools at their disposal. With only a few exceptions (most notably being the *extremely* controversial latitude of the OFCCP to create their own Pay Analysis Groups) the Directive is very straightforward. That being said, the problem is not in how the Directive is written. The problem is how it will be enforced[4] and the dramatically-increased burden associated with investigations.[5]

It is important to note that by-and-large, the Directive simply codifies what the OFCCP is already doing in audits. The Agency is already investigating compensation disparities systemically (if possible), in small groups, and on an individual cohort-level basis. If this is, in fact, the case, then what's the big deal? The big deal is the pressure they are under this time to "get it right," and, as a result, the contractor community should expect a significant number of audits to focus on compensation. This is necessary to not only demonstrate the Agency's commitment to the Administration's priorities, but also to achieve *measurable results* with respect to their enforcement responsibilities (*i.e.*, obtaining conciliation agreements with contractors). At the end of the day, this all leads to the same question, "What should conscientious and law-abiding organizations do to be proactive and to identify/rectify problem areas within their own organizations?"

RECOMMENDATIONS

1. Document! Document! Document!

Given that it will be extremely difficult for organizations to anticipate where the OFCCP may find "potential issues," (*i.e.*, at the systemic level, small group level, or on an individual basis) the only strategy that has the ability to protect an organization across all levels of investigation is to better document, in a job-related fashion, those reasons why employees are paid differently. Further, given that the vast majority of current compensation disparities are due to differences in starting salary, the advice is to start with any/all factors that impact starting salary! All too often, when hiring managers are interviewed by the Agency, they will indicate that factors such as education or prior-related experience (both quality and/or quantity) fundamentally impact a person's starting salary. However, it is the rare contractor that will actually collect this information during the hiring process. As a result, in the event of an audit or other type of investigation, organizations are left scrambling to collect this vital information as part of their defense.

Beyond recording those factors that impact an employee's initial starting salary, employers should better document decisions impacting an employee's compensation throughout their career. Performance appraisal scores, promotional decisions, allocation of training opportunities, shift differentials, distribution of sales routes, assignment to a "high potential" list (to name a few), are all decisions that can impact an employee's salary, and are all clearly identified as being on the OFCCP's new "hit list." It is going to be absolutely vital that employers be able to justify, in a job-related fashion, why they made the decisions they made and it all comes back to documentation.

2. Review Your Compensation System(s)

Directive 307 has made it abundantly clear that "pay grade theory" is back on the table. In short, pay grade theory is based on the belief that organizations, themselves, equate jobs by grouping them into pay bands and/or grades, and thus, make them combinable for analysis purposes. As a result, knowing that the Agency is going to be conducting analyses in this fashion should prompt organizations to review their compensation system and the structure of their grades/bands. Does your organization have wide grades that combine employees with extremely dissimilar skills, qualifications, and responsibility levels that perform largely different duties? If so, then it may be necessary to restructure the grades with the goal of developing more/narrower grades that, perhaps, do a better job of aggregating more similar employees. Of course, this all comes with a price. Do organizations create many grades to solely reduce exposure and better "guide" the OFCCP toward more legitimate analyses, even if it means losing the utility of the grade system in the first place? This is where organizations will need to carefully balance the needs of the organization with the desire to avoid unnecessary exposure.

Another issue occurs when employees with dissimilar skills, qualifications, and responsibility levels performing largely different duties share the same job title. One subset of the job may work on a project or in a department requiring specialized skills and, as a result, is compensated differently from the other group (for example: *Manager* in HR v. *Manager* in Engineering). If, in this circumstance, employees in the more highly compensated departments happen to be primarily non-minority, then it would appear to the casual observer that minorities were being underpaid relative to their non-minority counterparts. Under these circumstances, organizations should investigate their job titles and, perhaps, create new, more descriptive titles, to not mislead the Agency into believing that the employees are similarly situated.

The bottom line is that organizations should periodically examine the linkage between job classification systems and compensation systems. The Agency is going to equate the two, and absent a clear understanding of how the two work together, the contractor is at risk of missing potential cases of pay inequities and may be subjecting themselves to unnecessary and undue regulatory scrutiny.

3. Pay Analysis Groups (PAGs): This is Where Things Will Get Ugly

By-and-large, most attorneys and consultants will agree that in the event of an audit, compensation data should be submitted to the Agency in the "most appropriate unit for analysis purposes." More often than not this is job title, or perhaps, even a combination of job title and some type of organizational hierarchy variable (*e.g.*, department), if organizations use broad job titles across multiple units. This is because you can generally have the most confidence that those employees who are within these groupings share similar skills, have similar qualifications and responsibilities, and by-and-large will be performing similar job duties. Because the employees in groupings such as these are likely to be "similarly situated," any analysis comparing the compensation received by different gender or race/ethnic groups is likely to be more legitimate. This is because the statistical analysis can "control" for legitimate factors that impact pay like prior related experience (quality and quantity), education (quality and quantity), time in job, and job performance, which, in turn, allows for determining whether any *other* differences are due solely to the employee's gender or minority status.

While this is arguably true, there is a larger context. It is theoretically possible for an employer to create two job titles that are highly similar in *most* respects, except that one is more highly compensated that the other. If males tend to be assigned to the higher paid job title and females to the lower paid job title, then a potential problem may exist. What is distressing about this scenario is that a compensation analysis that focuses solely on job title will not uncover this disparity. It is for this reason that the new Directive also discusses the need to combine employees into larger Pay Analysis Groups (PAGs).[6]

It is important to note that the Agency clearly sees PAGs as a tool to combine jobs across locations as well. This is particularly worrisome for larger organizations with multiple locations. Consider the big-box retailer with job titles such as Merchandiser, Checker, and Warehouse Worker. In most organizations, the compensation for these titles is governed by a similar set of standardized and uniformly-applied rules, with the only (primary) difference being location. It is not a stretch to imagine a scenario where the OFCCP sees the opportunity to greatly broaden the scope of their investigation by combining all employees in the same job title across all locations.[7]

Even under the previous guidelines, it has always been the goal of the OFCCP to ensure that compensation discrimination is not hidden behind the claim "but we didn't have enough employees in that job title to conduct a statistical analysis!" The goal of the Agency has been to encourage, coerce if necessary, contractors into grouping employees into larger groups with larger sample sizes that allow for an appropriate statistical analysis that can generate informative results.

As every contractor knows, this is much easier said than done. Although the new Directive does outline many of the factors the Agency will consider when combining employees for analysis purposes, it also states very clearly that employees need not be similar in *all* factors to be combined into an analysis. This begs the question, "On which factors *do* they need to be similar?" For example, is similar skill enough, even if the employees are not similar in effort, responsibility, level of supervision required, or types of duties performed? Or do they simply need to be similar in duties, even if they are dissimilar in effort, responsibility, and skill? It's probably a safe bet that organizations will come to a very different conclusion than the Agency on this matter.

What, then, should the proactive employer do? To date, there is no single correct answer. However, it is almost universally agreed upon that, for a variety of reasons, employers should not create voluntary PAGs for the OFCCP.[8] Beyond that, it really depends on the individual employer. As it relates to PAGs, probably the best approach is to conduct a comprehensive study of the job classification system and eliminate duplicate job titles, collapse similar job titles, and create separate titles for those jobs that really are, in fact, different. For most employers this will be a large undertaking. However, as with all things, prioritize! Start with the high-employee count, high legal exposure job titles and branch out from there.

In the end, it's important to remember that during the course of an investigation, the OFCCP can choose to group employees however they want. It is the employer's role, as the keeper of all documentation and knowledge regarding the true content of each of their jobs, to help "guide" the Agency toward the most appropriate grouping methods and stand behind them regardless of what the OFCCP does. If organizations have done their homework, and they have taken the time to properly define and document the content of their jobs, then they will be in much better shape than the organization that didn't.[9]

4. Compensation Isn't Only About Pay . . . It Can Also Be About Opportunity (Which Can Lead to Differences in Pay)

A new frontier outlined in the Directive is the investigation of the *distribution of opportunities*. This approach is not foreign to EEOC and Title VII investigations. However, up to this point, it has not been a strategy widely (if at all) utilized by the OFCCP. The belief (and rightly so), is that if salaries are the only factor investigated, then rampant discrimination in the distribution of opportunities can go largely unnoticed. Is there a glass ceiling? Do only whites and/or men receive the "good routes?" Do only men work the high paying shift? Is the vast majority of the "high potential" list white and/or male? Do most training opportunities go to one group as opposed to another? The list of positive "opportunities," where discrimination may be hidden, is long and unique to each and every organization.

The good news is that analyzing the distribution of opportunities is relatively straightforward (at least when compared to analyzing compensation). Structurally, an analysis of the distribution of opportunities is exactly like analyzing for potential adverse impact hiring, where a two-by-two (2x2) table is created and a standard deviation is generated and interpreted, (see examples below in Table 1 and Table 2).

It is likely that the Agency will see these types of analyses as being very similar to a typical adverse impact analysis, and enforce in a similar fashion as well. What is unknown is whether the Agency will conduct this type of analysis and use significant results to simply bolster their argument for disparities in compensation and push for solely compensation-based remedies, or whether the Agency will use significant results with this type of analysis to push for *both* compensation-based remedies as well as "make-whole relief" for the disparity in the distribution of opportunities (*i.e.*, the need to provide remedies to the impacted group not only in compensation, but also in the impacted opportunity). On this note, only time will tell.

5. Conduct Proactive Statistical Analyses

"The best defense is a good offense." In this case, the best defense against an OFCCP audit or other compensation-related investigation is to be proactive,

TABLE 1: An Analysis of the Distribution of Training Opportunities

	Invited to Training	Not Invited to Training	Rate (%)
Men	50	50	50.0%
Women	25	75	25.0%

SD = 3.58

TABLE 2: An Analysis of the Distribution of "Good" Sales Routes

	Allocation of "Good" Routes	Allocation of "Not Good" Routes	Rate (%)
White	60	40	60.0%
Minority	40	60	40.0%

SD = 2.75

NOTE: Standard Deviations (SD) greater than or equal to 1.96 are considered significant.

be your own worst critic, and find and rectify as many potential issues as you can while conducting analyses under attorney-client privilege. Identifying issues proactively, behind the curtain of attorney-client privilege, offers many advantages: 1) it avoids the negative press associated with a public announcement of discrimination, 2) it helps to (potentially) avoid the time-consuming, lengthy, and expensive process of defending oneself throughout the audit/litigation, 3) it allows the organization to address the identified issues when and how it chooses, and 4) perhaps most importantly (at least from a financial perspective), it often allows the organization to avoid back-pay and/or punitive damages which can often dwarf any current pay adjustments.

First, and this may sound counter intuitive, when conducting proactive analyses of compensation, it's important to assume that your organization is, in fact, discriminating. If you start with this assumption then it will be critical. You will leave no stone unturned. As a result, you will be much more likely to uncover any problems that do, in fact, exist.

Second, because the greatest amount of legal exposure (as well as the path of least resistance and greatest reward for the OFCCP) is in homogeneous, high-volume job titles where all employees are doing largely the same thing (*e.g.*, Customer Service Reps, Checkers, etc.), start your proactive investigation with regression analyses of all the individual job titles with greater than 30 employees. Begin at the location level, but if your organization has job titles with large numbers of employees doing largely the same thing across multiple locations, then it may be in your best interest to conduct analyses by combining just these titles across locations. Of course, the OFCCP can find other issues by combining job titles, but given their apparent wide latitude to combine dissimilar jobs, taking this approach proactively would be like "throwing darts at a dartboard."

Third, if your organization has additional desire and/or resources then, by all means, combine into PAGs those job titles that are most likely to be seen as equivalent by the OFCCP. For example, combine all job titles within a "family" (*e.g.*, Customer Service Rep 1, 2, and 3, or Civil Engineer 1, 2, 3, etc.), or combine the same job title across multiple locations. Once again, focus first on the high-incumbency positions.

As for conducting proactive "cohort" analyses on small comparison groups, in an ideal world, employers would have unlimited time and resources to ensure all differences in employee compensation are properly documented and the reasons are clearly identifiable. However, in the real world, this is quite often unrealistic. As a result, the recommendation is to only conduct proactive cohort-level analyses when necessary to investigate issues with statistically significant disparities (see recommendation 6, below). Of course, the Agency can, and often will, identify issues between individuals. However, it is often impossible to predict from where this inquiry will come, and any attempt to identify these issues proactively often ends up being wasted effort.

6. Conduct Proactive Qualitative Analysis

After conducting statistical analyses, uncovering instances of statistically significant disparities in pay (if any), and calculating pay adjustments[10] necessary to remove those disparities, it is critical that organizations *not* proceed directly to adjusting employee pay. In cases where significant pay disparities have been observed, it is absolutely necessary for employers to then carefully evaluate data/information that was *not* included in the statistical analysis. This type of analysis is commonly referred to as a "cohort" or "file-by-file" analysis. Often there is a rational, job-related reason why employees are paid differently, but that information didn't make it into the statistical analysis. This can be because the information isn't stored within your organization's Human Resource Information System (HRIS), or because the information is anecdotal in nature. Either way, this information *must* be considered before making any final determination of discrimination. This is because multiple regression, for all of its power to identify systemic disparities, is only as good as the data included in the analysis. To the degree important information is missing from the analysis, the utility and accuracy of the results may be impacted.

CONCLUSION

A close read of the Directive, coupled with an understanding of how the OFCCP has already been investigating compensation over the past few years, will lead most to the conclusion that not much has changed. While this may be true, this notion is missing the point. The true impact of the Directive is not in what it says, but in the tremendous amount of pressure it puts on the Agency to succeed. This single factor alone will lead to a dramatic increase in the focus on compensation, with an equally dramatic increase on the burden for employers to defend themselves in the face of what will undoubtedly be an onslaught of data and document requests (both relevant and irrelevant). Under this pressure, Federal contractors may be tempted to conciliate simply to "see it all go away." However, be warned that this too is likely part of the "unwritten" compensation Directive . . . to "start big and see what sticks." The Agency is going to be emboldened by the new Directive and, as a result, will (likely) be more than willing to take novel issues such as broad Pay Analysis Groups to trial simply to establish precedence. Good. Bad. Indifferent. It is this precedence that will help to define the playing field for enforcement efforts to come.

Originally, when the Notice of Proposed Rulemaking (NPRM) rescinding the 2006 compensation *Standards* and *Guidelines* was announced, contractors cried, "How will we know how the OFCCP is going to investigate compensation?" Well, we got our answer. It just may not be what we wanted to hear.

END NOTES

1. *Interpreting Nondiscrimination Requirements of Executive Order 11246 With Respect to Systemic Compensation Discrimination. Fed. Reg. 71. No. 116 (Part VI)*
2. *Voluntary Guidelines for Self-Evaluation of Compensation Practices for Compliance With Nondiscrimination Requirements of Executive Order 11246 With Respect to Systemic Compensation. Fed. Reg. 71. No. 116. (Part V)*
3. *Congressional Budget Justification: Office of Federal Contract Compliance Programs*. FY 2011, 2012, 2013
4. Investigations will be led (at least initially) by Compliance Officers, many of whom will likely not have the necessary training and experience to properly and efficiently guide what will undoubtedly be complex compensation investigations. District/Regional/National Agency personnel will be, at times, involved, but the initial determination regarding whether and how to investigate deeper and make additional data/information requests will be spearheaded by the CO. On this note, it is important for readers to understand that the skills necessary to perform proper, efficient, and ultimately meaningful compensation analyses take years to amass.
5. Yes, we said it. *Burden*. Contractors by-and-large try to do the right thing. They might not like enforcement but they see it as a necessary burden for the greater good. What they have a problem with is "inefficient, unfocused, and unrealistic" enforcement characterized by voluminous data/document requests (*e.g.*, "send us all data, with all factors, for all employees . . . and we need it by Friday").
6. Noticeable in its absence from the Directive is the term "Similarly Situated Employee Group (SSEG)." It is this author's opinion that the OFCCP saw this term as too limiting, and has since opted for the more broadly-applicable (and less limiting) "Pay Analysis Group" which is consistent with the Agency's desire to combine employees in a more aggressive fashion.
7. In fact, we have heard that this exact strategy is currently being utilized by the OFCCP in audits.
8. Although conducting an internal investigation by employer-created PAGs, under attorney-client privilege, can be beneficial in the right circumstances.
9. Remember, it's not always about being "right." Sometimes it's simply about *not* being an easy target.
10. Regression analysis should be used to generate the amount needed to eliminate statistically significant disparities. This amount can then be "adjusted" through consideration of other factors not included in the statistical analysis.

How to Compute Liability

WHEN STATISTICAL EVIDENCE OF PAY DISPARITIES EXIST *Part I*

Daniel Kuang, Ph.D. | *Biddle Consulting Group, Inc.*
Dan A. Biddle, Ph.D. | *Biddle Consulting Group, Inc.*

Differences in average compensation between two groups (e.g., men and women) are not necessarily evidence of pay discrimination. However, if those differences are large enough to be statistically significant, and if they cannot be justified by job-related factors, they may be used as evidence of pay discrimination. Under these circumstances, non-defensible differences should be remedied through compensation adjustments.

This article will focus on how to address the gaps between the amounts that individual employees are paid and how much the regression model indicates they "should" be paid. However, it will not discuss the variety of options employers have regarding exactly how or when to make these adjustments, though that is just as important as identifying who needs an adjustment and how much.

The primary purpose in making compensation adjustments is to close the gap in compensation between the Focal (*e.g.*, women/minority) and Reference (*e.g.*, men/white) groups after controlling for differences in job-related factors. Despite how seemingly straightforward this should be, the massive volume of literature on compensation liability models provides sobering evidence of the complexity involved in properly modeling compensation adjustments.[1]

Fortunately, there is a very general and flexible method that may be applied in almost all circumstances. It is a methodological framework that is comprised of two steps:

Step 1: Compute Liability — the total amount of money to be paid to the negatively impacted group.

Step 2: Determine Distribution — the method of identifying how much of the total liability to distribute to each individual within a negatively impacted group.

This article will detail *Step 1—Compute Liability*. Step 2 will be detailed in a separate article.

COMPUTING LIABILITY—DESCRIPTION

The key to a general and flexible compensation liability model is a commonality in the underlying statistical model. The typical compensation analysis falls into one of two categories when it tests for differences between groups: 1) without controlling for explanatory variables (*e.g.*, *t*-Test, ANOVA – analysis of variance) or 2) with controlling for explanatory variables (*e.g.*, multiple regression). Despite differences in the analytical strategies, their underlying statistical model is the same; they are all variants of multiple regression.

Evidence of a significant between-group difference in compensation exists when the regression coefficient

(b) for the group variable (*e.g.*, men/women, white/minority, etc.) remains significant after controlling for differences in job-related variables. The following statistics are needed from the regression output to compute the amount needed to "eliminate" group differences (to varying degrees):

1. b_{Group}: Regression Coefficient for the group variable (*e.g.*, gender/race).
2. SE_b: Standard Error for the Regression Coefficient (b_{Group}).
3. N: Total sample size.
4. k: Total number of independent variables in the model.

If the group variable is coded properly (0=Focal/Women/Minority and 1=Reference/Men/White), the b is the mean difference in pay between Focal and Reference group members, shown in Equation 1 as:

$$b = \text{Mean}_{Focal} - \text{Mean}_{Reference} \quad \text{(Eq. 1)}$$

- If $b < 0$, then the Focal group is negatively impacted.
- If $b > 0$, then the Reference group is negatively impacted.

In a Title VII context, a significant b, irrespective of its directionality (positive or negative) is an indication of potential compensation discrimination.

When the regression model does not contain explanatory variables, then the b obtained in Eq. 1 can be interpreted literally: raw mean difference in compensation between Focal and Reference groups.

However, the typical regression model will include one or more explanatory variables. In such instances, the b in Eq. 1 is the mean difference between the Focal and Reference group *after* controlling for differences in the explanatory variables. This is often referred to as the "adjusted mean."

The statistical test to determine whether b is significant is:

$$t = b \,/\, SEb \quad \text{(Eq. 2)}$$

With

$$df = N - k - 1 \quad \text{(Eq. 3)}$$

METHOD

Once these statistics are obtained, computing compensation adjustments becomes a fairly straightforward mathematical exercise. The steps for computing the amount needed to "eliminate" significance are as follows:

1. Determine the desired legal defensibility: In standard deviation units, what is the tolerable pay disparity (*i.e.*, 2, 1, or 0)? Once the desired standard deviation difference in pay disparity is determined, compute the *p*-value. Common thresholds are computed and presented in Table 1.

   For the advanced analysts who desire to apply specific standard deviation units, they may convert standard deviation units into 2-tail *p*-values with a statistics table or apply the following formula in Excel:

   =2*(1-NORMSDIST(Standard Deviation)).

TABLE 1: Establishing Acceptable Levels of Legal Defensibility

Standard Deviation	*p*-value[a]
2	40
1.95	25
1	4
0	2

Note: [a] 2-tail *p*-value.

It is important for employers to understand that reducing salary differences to two (2) standard deviations will cost less than reducing the disparities to one (1) or zero (0) standard deviations. However, this gives the employer very little "cushion." Meaning, that even small changes in salaries or workforce composition can/may cause the statistically significant disparity to reappear.

2. Determine the non-significant *t*-value: Once the desired and tolerable *p*-value is determined (Step 1), the next step is to compute the *t*-value for the available degrees of freedom (*df*). The non-significant *t* may be obtained from a statistics table or Excel with the following formula:

$$t_{non\text{-}significant} = \text{TINV}(p\text{-value}, df) \qquad \text{(Eq. 4)}$$

3. Compute compensation adjustment: Once the non-significant *t*-value is determined (Step 2), the next step is to insert this value and component from the original *t*-test formula into the following formula:

$$\text{Liability}_{\text{individual}} = (t_{non\text{-}significant} \times SE_b) - |b_{\text{Group}}| \qquad \text{(Eq. 5)}$$

This computed compensation liability is at the individual level. Specifically, it is the amount that needs to be adjusted for each individual in the impacted group to reduce the pay disparity to the desired level. The total liability for the impacted group is:

$$\text{Total Liability} = N \times \text{Liability}_{\text{individual}} \qquad \text{(Eq. 6)}$$

To confirm the validity of these adjustments, a "what-if" simulation analysis can be performed. In such an analysis, calculated adjustments are added hypothetically to the appropriate employees in the database and the pay disparity between Focal and Reference members is re-evaluated. If the results of the statistical test matches the desired pay disparity (*e.g.*, 0, 1, 2 standard deviations), then the computed liability is valid.

CONCLUSION

This paper detailed a general method of computing liability within a multiple linear regression framework. Although the mechanics of computing liability is fairly straightforward, it is important that analysts understand the concepts of this method prior to making any pay adjustments.

Please note that the method detailed in this paper is only one of two steps in a comprehensive pay adjustment study. This first step details how to compute the total amount necessary to diminish the pay gap between focal and reference members in a group. A future article will detail methods of distributing the computed liability.

ENDNOTES

1. Colquitt, J., Conlon, D.E., Wesson, M.J., Porter, C., & Ng K.Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86*(3), 425-445.

How to Compute Liability WHEN STATISTICAL EVIDENCE OF PAY DISPARITIES EXIST *Part II*

Daniel Kuang, Ph.D. | *Biddle Consulting Group, Inc.*
Dan A. Biddle, Ph.D. | *Biddle Consulting Group, Inc.*

In Part I, we detailed the process for computing liability within a multiple linear regression framework. This article covered the statistical mechanics of computing liability, as well as important concepts and steps that analysts should understand and consider prior to making any pay adjustments. The prior article detailed only one of two steps in a comprehensive pay adjustment study—how to compute the total amount *(i.e., the amount still outstanding after accounting for differences that may exist in job qualification factors) necessary to diminish the pay gap between focal and reference members in a pay study. In this article (Part II), we describe the methods of* distributing *the computed liability to the individuals in the affected class.*

After the total pay liability has been determined (using the procedures outlined in the first article), the next step is to *distribute* the liability among impacted group members. This component of the process, while absolutely critical, is not well understood and often ignored. This is understandable because the typical analysis of compensation focuses on group mean (average) differences. Since differences among individuals within a group do not alter the group mean, liability distribution is often largely ignored.

In practice, however, liability distribution is an *essential* component of correcting systemic compensation imbalances. This is because the legitimate variables (*e.g.*, tenure, education, experience) that contributed (legitimately, free of discrimination) to making up the pay differences that exist between individuals in the study need to be taken into account. And, because the individuals in the study will possess these factors in varying levels, they need to be taken into account when determining how the remedial pay is distributed among focal group members based on how far each person is below their predicted pay. Compensation adjustments based on individual employee salaries and individual differences in job-related factors will:

1. ensure optimal and stable pay equity for all individuals;
2. create a more coherent and strong statistical compensation model;
3. increase perceptions of organizational fairness; and
4. reduce (potential) legal exposure associated with making compensation adjustments.

If these steps are not done correctly, the employer can be left open to various types of liability (examples of this are provided below).

OPTIMAL AND STABLE COMPENSATION ADJUSTMENTS

Compensation is often analyzed within a static framework (*e.g.* a 12/31 snapshot dataset, amount of pay increases), because this is a required constraint when testing for group differences or computing pay disparities. However, this places the analyst in a very tenuous situation. Compensation disparities are constantly changing and vary as a function of time, workforce changes, and individual attributes (*e.g.*, gender, race, age and job-related criteria). Proper liability distribution strategies will take into account these influences to ensure that the adjustments are as fair and optimally stable as possible.

Compensation varies as a function of dynamic influences across time. Two major events occur as time passes: First, compensation naturally (and typically) increases as a function of tenure. Second, workforce composition changes due to such events as promotions, terminations, transfers, and hires. Consequently, compensation adjustments that may eliminate group differences at one point in time may unravel after one cycle of pay raises and personnel changes if the liability adjustments are not optimally distributed to employees based upon their individual level of "underpayment" or impact. Compensation follows a growth curve; if compensation adjustments are distributed blindly without regard to individual levels of "underpayment," pay disparities that are eliminated at one point in time at the group level may easily resurface because the underlying disparities still exist.

FAIRNESS OF COMPENSATION ADJUSTMENTS

There are volumes of literature on the importance of the perception of organizational justice.[1] Fairness (and *perceived* fairness) has been tied to positive organizational citizenship behavior and a decrease in counter-productive behaviors. Moreover, a perception of unfairness is a primary trigger for individuals to seek litigation against their employers. In brief, it is to the benefit of the employer to ensure that the distribution of compensation adjustments is fair. Proper liability distribution strategies take into account individual differences and are therefore, arguably, most fair.

LIABILITY DISTRIBUTION MODELS

Important Note: *There are several liability distribution models, each with specific strengths and weaknesses. Competent practitioners may differ in their opinions of which are the most appropriate even under similar circumstances. For these reasons, we believe it is critical to consider the context, cohort review results, data (e.g., sample size), and regression model before deciding which liability distribution model to apply. In most circumstances, however, we believe the Proportionate Distribution Model should be used for reasons explained below.*

Liability distribution models can be divided into

TABLE 1: Distributing Compensation Adjustments Equally to All Impacted Group Employees

ID	Compensation ($)	Liability Distribution ($)
1	$12.00	$0.80
2	$11.00	$0.80
3	$9.00	$0.80
4	$16.00	$0.80
5	$7.00	$0.80

two categories: dual regression models (where a regression model is developed for each group) and single regression models (where a single regression model is used for the entire group and the gender/race status is dummy-coded). Because the single regression model methods are more common, these are discussed first and in more detail.

Three of the most common liability distribution models that are based on a single regression model include:

1. Even Distribution for All;
2. Even Distribution for Individuals Below the Mean; and
3. Proportionate Distribution (Based Upon Model Prediction)

Liability Distribution—Even Distribution for All

This is the simplest of the three distribution models. As the name of this model implies, the total liability is evenly distributed to all individuals within the impacted group. Consider the following example of five negatively-impacted women where the computed liability for the group is $4.00 (see Table 1). Liability is evenly-distributed by dividing total liability ($4.00) by the number of individuals in the group (5): $4.00 / 5 = $0.80.

The authors do not recommend this method because it does not take into account *individual* employee differences and may require liability payments to employees already paid more than their predicted salary.

Liability Distribution—Even Distribution for Individuals Below the Mean

One of the major limitations of the *Even Distribution for All* model is that it ignores individual differences. Extending the example above (see Table 2), the average[2] salary is computed ($11.00) and each individual's salary is compared against this mean. Among the five records, three are at or above the mean (1, 2, and 4). Notably, the 4th person is overpaid by $5.00 when compared to the mean. Only two individuals' salaries fall below the mean (ID #3 and #5). Given this, it is inappropriate to evenly distribute the liability across all individuals. As an improvement to the Even Distribution for All model, the liability is evenly distributed for those individuals who fall below the mean. This is a four-step process:

1. Compute the overall group mean (including *both* Focal and Reference).
2. Compute the difference from the mean for all individuals in the negatively-impacted group.
3. Identify and count the number of impacted group members below the mean ($n_{Below\ Mean}$).
4. Compute the even distribution for impacted group members below the mean: Total Liability / n (below mean)

TABLE 2: Distributing Compensation Adjustments Equally to All Impacted Group Employees with Below Mean Compensation

ID	Compensation ($)	Difference from Mean ($)	Liability Distribution ($)
1	$12.00	$1.00	
2	$11.00	$0.00	
3	$9.00	-$2.00	$2.00
4	$16.00	$5.00	
5	$7.00	-$4.00	$2.00

Note: Total Group Mean = $11.00

When applied to the example:

1. Overall group mean = $11.00
2. Compute the difference from mean for each individual (see Table 3).
3. Count the number of individuals below the mean: $n_{\text{Below Mean}} = 2$
4. Compute the even distribution for individuals below the mean: $4.00 / 2 = $2.00.

While an improvement over the *Even Distribution for All* model, the authors do not recommend this method because it does not take into account *individual* employee differences and may require liability payments to employees already paid more than their predicted salary.

Liability Distribution—Proportionate Distribution

Although *Even Distribution for Individuals Below the Mean* is an improvement over the first model, there is a noticeable weakness—the distribution is not proportional to individual pay disparity (*i.e.*, the difference between what each underpaid employee is actually paid, and what they should be paid, based upon the regression model). In this example, Person #5 is twice as far from the mean as Person #3 ($4.00 vs. $2.00, respectively), but both received the same amount ($2.00). In addition, similar to the *Even Distribution* method, it does not take into account individual employee differences in job-related factors.

One variant of the third distribution model, *Proportionate Distribution*, serves to address these concerns. We believe that this method is ideal in most circumstances because it simultaneously considers both group- and individual-level pay disparities. At the group level, this method focuses on reducing the significant coefficient (*b*) for the group variable (*e.g.*, men/women, white/minority) to the specified level (*e.g.*, 0 for parity, to 1 standard deviation [SD]). In this way, the amount of the regression model that is directly attributable to race or gender (after controlling for differences in job-related variables) is addressed in the most direct manner possible. And, on the individual level, rather than splitting the liability evenly for those paid less than the mean, the *Proportionate Distribution* model considers individual differences when determining liability distribution. This is accomplished by first creating a regression model *without* the protected variable (*e.g.*, dummy-coded men/women, whites/minorities). By leaving out the protected group variable, an overall model of compensation is created without any potential discrimination based on gender and/or race affiliation. By applying this approach, it is possible to obtain a predicted compensation for each individual based on their unique job-related attributes (*i.e.*, explanatory variables) only.

The mechanics of this method are detailed below:

■ **Step 1:** Compute the *Predicted Compensation* ($\hat{Y}$) for each employee:[3]

Eq. 1.

$$\hat{Y}_{predicted} = \alpha + b1X1 + b2X2 + \ldots + biXi$$

Apply this model in computing the predicted compensation ($\hat{Y}_{predicted}$) for each individual, given their unique attributes (*i.e.*, explanatory variables). When no explanatory variables are specified in the model, the regression model simplifies to $\hat{Y}_{predicted} = \alpha$, which is the average compensation for all members (Focal and Reference together).

■ **Step 2:** Compute the *Difference from the Model* for each underpaid employee. For members in the negatively-impacted group, compute the difference between observed compensation ($Y_{observed}$) and predicted ($\hat{Y}_{predicted}$):

Eq. 2.

$$\text{Difference from Model} = Y_{observed} - \hat{Y}_{predicted}$$

■ **Step 3:** Identify those employees paid less than the model predicts. For members in the negatively-impacted group, select only those who are paid below their predicted salary ($Y_{observed} < \hat{Y}_{predicted}$), (*i.e.*, negative *Difference from the Model*).

■ Step 4: Compute the *Total Model Shortfall*. For members in the negatively-impacted group who fall below their predicted salary ($Y_{observed} < \hat{Y}_{predicted}$), sum the *Difference from the Model* (item 2, above). This is the *Total Model Shortfall*—the total amount that is under the predicted model.

■ Step 5: Compute the *Proportion of Impact*. For each member in the negatively-impacted group who fall below predicted, compute their *Proportion Impact* using:

Eq. 3.

$$\text{Proportion of Impact} = (Y_{observed} - \hat{Y}_{predicted}) \,/\, \text{Total Model Shortfall}$$

■ Step 6: Compute *Proportionate Distribution*. For each member in the negatively-impacted group who fall below the predicted value, compute their *Proportion of Distribution* using:

Eq. 4.

$$\text{Prop. Dist.} = (Y_{observed} - \hat{Y}_{predicted}) \,/\, (\text{Total Model Shortfall X Total Liability})$$

This formula computes the proportion of the *Total Liability* that each individual should receive as a function of their individual attributes ($\hat{Y}_{predicted}$) and how far their observed compensation is from their predicted compensation ($Y_{observed} - \hat{Y}_{predicted}$) relative to all members who fall below their predicted compensation.

A detailed (and realistic) case study of the *Proportionate Distribution* method is provided below.

Case Study: An Example of the Proportionate Distribution Method

Company Z conducts a proactive compensation analysis that is not in response to litigation or government enforcement agency investigations. Company Z has 100 employees in the at-issue Similarly Situated Employee Groups (SSEG) (*i.e.*, job title): 50 women and 50 men. The average compensation for men is $52,260 and $48,520 for women (a $3,740 mean difference, or about 7.16%). Tenure is the only explanatory variable included in the model, and no interactions are found between tenure, pay, and gender.

After conducting the regression analysis, both tenure and gender are statistically significant, with tenure having a correlation of .56 to pay and a corresponding coefficient ($b1$) of $1,631.96 and the gender variable having a .34 correlation to pay and a coefficient ($b2$) of $2,205.95. The t-values are 5.83 ($p < .01$) and 2.36 ($p = .02$) respectively; indicating that tenure is highly significant and that evidence of possible pay discrimination exists because the gender t-value exceeds 2.0 (indicating that $p < .05$) after controlling for tenure.

The tenure coefficient indicates that for every single-unit increase in tenure (*i.e.*, for every year), the expected pay of each employee in the model goes up by $1,631.96. Because men are coded as 1 and women as 0, the gender coefficient indicates that the effect of being a man (after controlling for tenure) adds $2,205.95 to an individual's predicted pay. Adding the constant (a = $42,514.37) to the model allows for pay predictions to be made for each employee in the SSEG.

These coefficients can readily be used to predict pay using the following standard regression formula ($\hat{Y}_{predicted} = a + b_{tenure}X_{tenure} + b_{tenure}X_{tenure}$). For example, a woman (dummy-coded 0) with five years experience has a predicted pay of: $42,514.37 + ($1,631.96 x 5 = $8,159.80) + 0 = $50,674.17. A man (dummy-coded 1) with the same five years experience has a predicted pay level that is exactly $2,205.95

higher ($52,880.12) because the 0 would be replaced by $2,205.95 x 1, which adds $2,205.95 to the pay prediction.

Because the tenure coefficient is still significant after controlling for job qualification factors, Company Z desires to utilize the information from the regression model to eradicate the possible pay discrimination using the *Proportionate Distribution* method. Company Z has already completed an extensive cohort analysis including manager interviews to determine whether the statistical evidence of possible pay discrimination that has been identified by the regression study would be confirmed with additional evidence. As a result, they are choosing to correct the pay differences from the current *t*-value of 2.36 (with a corresponding *p*-value of .02) down to a *t*-value of 1.0 (with a corresponding *p*-value of .0.317, obtained by using the formula: =2*(1-NORMSDIST(1)) in Excel. (Remember, *t*-values are about the same conceptually as "standard deviations" referred to in the OFCCP regulations when interpreting the probability levels of the analysis results, with *t*-values exceeding values of 2.0 as statistically significant).

Company Z then generates a regression model *without gender* (including the tenure variable only) to compute predicted pay values for each employee, and subtracts each employee's actual pay from their predicted pay. Because there are 50 women in the at-issue group and the coefficient associated with gender is $2,205.95, the *maximum* liability amount is $110,297.67 (50 women x $2,205.95 each). In other words, the $2,205.95 effect associated with gender multiplied by the total number of women equates to the *total gender effect* identified by the regression model—*after giving each employee credit for their tenure*. If Company Z desired to reduce this gender effect to 0, this total amount ($110,297.67) would be allocated to the subgroup of women in the job title who were under-paid based upon the regression model—in proportion to how far each was away from their predicted salary (in this case 29 of the 50 women—see Table 3).

Important Note: *In ideal situations, liability computations should be made for employees for whom complete data exists for the variables in the model. However, in practical settings, this is not always possible. In these situations, data can be* imputed *for the missing variables using the average from the regression model. Without imputing data values for subjects who have missing data, the liability computations would simply compute 0 values for each—working detrimentally to the employees. In other words, without imputing data, the impact of not having the data for a certain variable—say job performance score—will actually treat the employee as if their score was 0.*

However, because Company Z desires to correct the pay disparity down to a *t*-value of 1.0 (and not make the assumption that 100% of the pay gap is due to possible discrimination), they will use the *p*-value that corresponds to a *t*-value of 1.0 (p = 0.317, using the formula above) and compute the associated gender coefficient: $t = b \, / \, SE_b$, which translates to $2,205.95 / $934.44 = $1,266.72. Multiplying this value by the total number of class members ($1,266.72 x 50) results in a modified total liability value of $63,336.02. This amount is then proportionately distributed to each of the (29) women whose actual pay ($Y_{observed}$) is below their predicted pay ($\hat{Y}_{predicted}$), by dividing each "Difference" value by the liability total (see values provided in Table 4 in the column titled, "Pay Adjustment to 1.0 *t*-value").

To confirm the validity of these adjustments, a "what-if" simulation analysis can be performed. In such an analysis, calculated adjustments are added hypothetically to the appropriate employees in the database and the pay disparity between focal and reference members is reevaluated. If the results of the statistical test match the desired pay disparity (*e.g.*, 0, 1, 2 standard deviations), then the computed liability is valid. However, unless the sample sizes are very large and the regression model is perfect (or nearly perfect—which, of course, is never the case), the desired *t*-value will not be exactly obtained.

TABLE 3: Identifying Compensation Liability Adjustments

ID #	Tenure	Curr. Pay ($Y_{observed}$)	Pred. Pay ($\hat{Y}_{predicted}$)	Diff.[a]	Weigh Prop.[b]	Pay Adjust. to 0 *t*-value[c]	Pay Adjust. to 1.0 *t*-value[d]
51	2	$36,000	$46,481	$-10,481	9.03%	$9,955	$5,716
53	4	$41,000	$50,117	$-9,117	7.85%	$8,659	$4,972
52	2	$38,000	$46,481	$-8,481	7.30%	$8,055	$4,626
72	4	$42,000	$50,117	$-8,117	6.99%	$7,710	$4,427
94	4	$42,000	$50,117	$-8,117	6.99%	$7,710	$4,427
56	4	$44,000	$50,117	$-6,117	5.27%	$5,810	$3,336
73	4	$44,000	$50,117	$-6,117	5.27%	$5,810	$3,336
83	4	$44,000	$50,117	$-6,117	5.27%	$5,810	$3,336
58	5	$46,000	$51,935	$-5,935	5.11%	$5,637	$3,237
55	3	$43,000	$48,299	$-5,299	4.56%	$5,033	$2,890
54	2	$42,000	$46,481	$-4,481	3.86%	$4,256	$2,444
95	3	$44,000	$48,299	$-4,299	3.70%	$4,083	$2,345
91	7	$52,000	$55,571	$-3,571	3.08%	$3,392	$1,948
77	5	$49,000	$51,935	$-2,935	2.53%	$2,788	$1,601
90	6	$51,000	$53,753	$-2,753	2.37%	$2,615	$1,502
66	7	$53,000	$55,571	$-2,571	2.21%	$2,442	$1,402
74	3	$46,000	$48,299	$-2,299	1.98%	$2,184	$1,254
85	3	$46,000	$48,299	$-2,299	1.98%	$2,184	$1,254
96	3	$46,000	$48,299	$-2,299	1.98%	$2,184	$1,254
76	4	$48,000	$50,117	$-2,117	1.82%	$2,011	$1,155
97	4	$48,000	$50,117	$-2,117	1.82%	$2,011	$1,155
89	5	$50,000	$51,935	$-1,935	1.67%	$1,838	$1,055
57	2	$45,000	$46,481	$-1,481	1.28%	$1,407	$808
84	2	$45,000	$46,481	$-1,481	1.28%	$1,407	$808
75	3	$47,000	$48,299	$-1,299	1.12%	$1,234	$709
86	3	$47,000	$48,299	$-1,299	1.12%	$1,234	$709
63	4	$49,000	$50,117	$-1,117	0.96%	$1,061	$609
88	4	$49,000	$50,117	$-1,117	0.96%	$1,061	$609
65	6	$53,000	$53,753	$-753	0.65%	$715	$411

Notes: aThese values are computed by subtracting each employee's actual pay from their predicted pay. bThese values are computed by dividing each employee's Difference value by the total of all values in the Difference column. cThese values are computed by multiplying each employee's Weighted Proportion value by the total amount of liability identified by the regression model (computed by multiplying the gender coefficient by the total number of impacted class members). dThis column is identical to the "Pay Adjustment to 0 t-value" column, but is set to correct pay to a t-value of 1.0 (computed using Eq. 2), then multiplying this value by the total number of impacted class members.

Liability Distribution—Dual Regression Models

A "dual" regression model for calculating liabilities simply implies developing one regression model to identify the existence of an unexplained statistically significant disparity between two groups, and another regression model to calculate the liability. The dual regression model is also commonly referred to as the "Peters-Belson" (P-B) method named after the two authors.[4] The P-B method (also sometimes referred to as the "Blinder–Oaxaca" method[5]) simply builds the liability model using only the reference group members, then applies the model to the focal group members. The resulting pay differences are said to constitute the "difference due to discrimination" (at least in the context of compensation analysis where the facts would support this conclusion). For example, a male-only regression model could be developed using the relevant job qualification factors, the resulting constant and regression variable weights could be used to compute predicted pay values for each of the women, and the resulting differences between their actual and predicted pay treated as the liability amounts.

While this method has been used in some litigation settings[6] it has not been met without criticism.[7] Perhaps the most significant limitation with the P-B method is that it substantially reduces the sample size used in the analysis. Because the regression model is developed using only the reference group members, the resulting model is less "conditioned"—and therefore possibly less accurate—than a regression model developed using the entire sample. Unless the strong assumption that "the focal group members can offer no useful information for building an accurate regression model" can be met, the predictive accuracy of the model will typically be reduced by using only part of the sample to build it.

This is especially true when conducting regression analyses on smaller samples which will typically result in a wider Standard Errors of Estimate (*SEE*). Wider SEEs result in decreased accuracy when using the model to make predictions regarding pay. While techniques do exist (*e.g.*, "jackknifing" and "bootstrapping") to help accommodate for these limitations,[8] we view this particular limitation as a serious one that applies to many regression situations.

An additional limitation that pertains to the P-B method has to do with the *statistical distributions* of the job qualification factors used in the model. For the P-B method to work accurately and reliably, both the *range* and the *variance* of the job qualification factors should be similar between the two at-issue groups. For example, if the regression liability model (*i.e.*, the model used in the dual regression approach to identify the dollar liability amount) is developed using a male-only model, and most of the men in the model have mid- to high-levels of the job qualification factors, the model might not predict well when applied to the females if they only have low- to mid-levels of the same factors. This is because the regression model may not be able to make accurate predictions for individuals in the focal group (*e.g.*, women), who may have lower levels of the job qualification factors, if there is a meaningful floor in the job qualification factor where the correlations were observed in the reference group (*e.g.*, men only) model.

If there are no major range differences between the groups on job qualification factors, then the variance between the two groups should be similar enough so that building the model on the reference group will provide regression weights that can be validly applied to the focal group members. If one group's variance on a job qualification factor is wide (shown by a large SD), while the other group's variance is narrow (shown by a small SD), the regression weights may not translate accurately between models.

While the extent of these limitations can be evaluated on a case-by-case basis, we do not view the P-B method as a typical *starting place* to use when computing liabilities. Rather, we recommend the single regression model (using the Proportionate Distribution Method) described above because it does not have these limitations. In addition, the preferred method typically *increases* the robustness of the model (*i.e.*, by increasing R^2) after pay corrections are made, whereas the P-B method lowers the same. Given these limitations, the P-B method may still be an effective way of computing the "upper bound limit" of liability in some circumstances (because the liability amounts will almost always be higher when using the P-B method as compared to others).

EVALUATING THE LEGAL DEFENSIBILITY OF REGRESSION MODELS DESIGNED TO INVESTIGATE PAY DISPARITIES

Employers that make pay adjustments to certain group members by relying on weak or flawed statistical evidence can open themselves to legal challenge. One of the most widely-cited cases dealing with this issue is *Rudebusch v. Hughes*.[9] In *Rudebusch*, the employer made $278,966 in corrective pay adjustments to women and minorities based on a limited (and flawed) regression study. After several years of litigation and a review (and remand) by the Ninth Circuit Court of Appeals, the Federal District Court ultimately ordered the defendants to pay $2 million to the whites and men who were adversely impacted by the earlier decision to increase the pay of women and minorities based on the limited and flawed regression study. On remand, the District Court made a careful review of the original regression study that was used to make the pay increases to women and minorities and identified several issues that undermined the validity of the regression model—making the resulting pay changes to women and minorities unjustified.

When looking back over the last 30 years of pay disparity cases, it becomes quite clear that the most fundamental requirement for substantiating pay discrimination (and therefore making changes to the disadvantaged group) is a showing of *statistical significance* associated with the gender/race variable. This is precisely where the employer went awry in *Rudebusch*—they made changes to the salaries of minorities and women without first clearly proving that a *statistically significant* pay disparity existed between groups. When evaluating whether making pay adjustments to disadvantaged group members is justified, the courts first typically evaluate whether a *manifest imbalance* exists in the pay between groups. Absent clear evidence of disparate treatment, demonstrating that a manifest imbalance exists in the pay between two groups *requires* a statistically significant finding. In the context of regression analysis, this means that the gender or race variable must be statistically significant *after* controlling for job qualification factors. This was not the situation in *Rudebusch*, and was one of the reasons that the employer ultimately had to redress their decision to make pay changes to the minorities and women in the case.

In fact, both the defense and plaintiff regression studies reviewed by the Court in *Rudebusch* revealed that the differences in pay were not statistically significant. The defendant's regression revealed that the difference attributable to ethnicity was only $87 and was not statistically significant. The difference between men and women was also not statistically significant.[10] The plaintiff expert's regression analysis found that the differences between men and women "would not even remotely be statistically significant" and both "gender and minority status do not come close to being statistically significant."[11] The District Court further clarified that "if 'manifest imbalance' requires a 'statistically significant disparity,' then there is no 'manifest imbalance' in this case."[12]

In addition to not demonstrating that a manifest imbalance existed between groups (through a showing of statistical significance), the opposing expert analysis and the court noted several internal flaws with the original regression analysis that was used as a basis for making pay changes.

There are two major lessons that can be learned from the *Rudebusch* case. First, before making pay adjustments to a group, be sure the regression model clearly shows that the gender or race variable is statistically significant after controlling for job qualification factors. Second, make sure that the regression model is sound, accurate, and reliable. To help employers address these key requirements, as well as the core requirements from other related pay discrimination cases, we offer the following guidelines:

1. Do not make pay adjustments unless multiple regression analyses are used (opposed to other techniques) to control for realistic differences in job-related factors that may exist between groups. In most situations, using multiple regression is the only clearly acceptable way to model compensation

decisions, and has decades of support in the federal courts and recent endorsement from both federal enforcement agencies that investigate and enforce pay equity cases.[13]

2. Do not make pay adjustments unless the gender or race variable is statistically significant after controlling for job qualification factors.

3. Be sure that the pay equity analysis was designed to identify significant pay disparities that may exist for *any* group (whites and men included). In addition, determine (preferably in advance) how pay disparities will be addressed if discovered to insure that the criteria and rules will be *uniformly applied* across all gender and race/ethnic groups.

4. Do not make pay adjustments unless *the regression model itself is statistically significant*. This can be accomplished by evaluating the ANOVA associated with the model.

5. Insure that the *strength* of the regression model is adequate for making reliable predictions. The strength of the regression model can be evaluated by referencing the Adjusted R^2 value, with Adjusted R^2 values that are statistically significant passing a minimum threshold.[14] In addition, the degree of multicollinearity among the variables should be evaluated (high multicollinearity tends to inflate standard errors associated with predictions, which can make predictions less reliable).

6. Do not make pay adjustments until after performing a "*cohort*" analysis whereby additional variables (*i.e.*, those not included within the regression analysis) are investigated.

7. Be sure that the *fundamental factors* relevant to compensation have been included in the regression analysis or evaluated in the cohort analysis. This has been one of the key factors reviewed when regression studies are contested in litigation settings. In *Bazemore v. Friday*,[15] the U.S. Supreme Court addressed this issue by evaluating the validity of statistical evidence that is necessary to support an inference of discrimination, but fails to consider *all possible variables*. In *Bazemore*, the Court reversed the lower court's refusal to accept the plaintiff's regression analysis as proof of pay discrimination, noting that "discrimination need not be proved with scientific certainty." The Court rejected the lower court's conclusion that "an appropriate regression analysis should include *all measurable variables thought to have an effect*" (478 U.S. at 399, 400) (emphasis added). Thus, in *Bazemore*, the Court ruled that statistical evidence may prove discrimination provided that it accounts for the major measurable factors causing the disparity. Rather than requiring the "perfect regression model," the courts typically require the opposing party to prove that the omitted variables would have *substantially changed the outcome of the study*, and they typically do not allow an inference of discrimination (based on statistical evidence) to be rebutted by simply pointing out unaccounted variables that might have affected the analysis.[16]

8. Be sure that the compensations adjustments made to the disadvantaged group are no more than necessary to attain a balance. As noted in *Rudebusch*:

> *"In addition to existence of a manifest imbalance, the pay equity plan must not unnecessarily trammel the rights of others, and it must be designed to do no more than 'attain a balance' (citing Johnson v. Transportation Agency, 480 U. S. at 637-39, 1987). It is logical that, since pay equity plans are, at least theoretically, implemented to eliminate a pre-existing manifest imbalance, Title VII requires that they must not be designed to go beyond correcting the imbalance, or unnecessarily trammel the rights of others."*[17]

When dealing with the important issue of "attaining balance" and "not trammeling the rights" of other groups not part of the pay adjustments, the Ninth Circuit noted in *Rudebusch* that "while pay equity plans resemble affirmative action, they are not concerned (as affirmative action usually is) with providing an ultimate advantage, such as providing preferences in hiring and promotion plans. Though sometimes labeled as affirmative action, "a pay equity plan such as that implemented by [the defendants] seeks to eliminate *existing* salary disparities for *particular individuals* due to race and sex (emphasis added)."[18] The Federal District Court also clarified this matter by stating: "In other words, where salary is already skewed due to discrimination (as prohibited by Title VII, on account of race and sex equalization results in the *elimination* of the preferences—it does not create a preference."

9. Thoroughly discuss with legal counsel and executive staff how adjustments to compensation will be made (*e.g.*, incrementally, lump-sum, as part of a yearly compensation/performance review).

SUMMARY

Based upon the limited discussion of the three above distribution methods, it should be apparent that liability calculations and the distribution of those monies is a dynamic and complex issue. It is important for employers to remember that regression analyses are only as good as the data they include. Given this, it is important for employers to also realize that, to the degree a regression analysis is lacking due to small sample sizes, missing data, or a myriad of other factors, the liability calculations and the distribution of those monies should be used only as a guide. Regression analyses and statistical liability calculations should never be used as the sole determinant for compensation adjustments.

Lastly, employers are cautioned against making compensation adjustments based upon weak, incomplete, or flawed regression analyses.

Important Note: *It is important for employers to remember that* all *employees are protected from unlawful discrimination. Making unjustified salary adjustments only for certain groups may lead to findings of unlawful compensation discrimination elsewhere. It is highly recommended that* all *significant disparities, regardless of impacted group, be thoroughly investigated, documented, and addressed.*

ENDNOTES

1 For example, Colquitt, J., Conlon, D.E., Wesson, M. J., Porter, C., & Ng K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology, 86*(3), 425–445.

2 Average salary computation includes individuals from both the Focal and Reference groups.

3 Exclude the grouping variable (*i.e.*, Men/Women, Whites/Minority).

4 Peters, C. C. (1941). A method of matching groups for experiment with no loss of population. *Educational Research, 34*, 60; Belson, W. A. (1956). A technique for studying the effects of a television broadcast. *Applied Statistics, 5*, 195.

5 Blinder, A. S. (1973). Wage discrimination reduced form and structural variables. *Journal of Human Resources*, 8, 436–455; Oaxac, R. (1973) Male-female differentials in urban markets. *International Economic Review*, 14, 693–709.

6 Sinclair, M. D. & Pan, Q. (2009) Using the Peters-Belson method in equal employment opportunity personnel evaluations. *Law, Probability and Risk, 8*, 95–117.

7 Greiner, J. D. (2008). Causal inferences in civil rights litigation. *The Harvard Law Review, 122* (2), 540–597.

8 Sinclair, M. D. & Pan, Q. (2009) Using the Peters-Belson method in equal employment opportunity personnel evaluations. *Law, Probability and Risk, 8,* 95–117; Efron, B. & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman and Hall, New York.

9 *Rudebusch v. Hughes*, No. 95-CV-1313-PCT-RCB, No. CV-1077-PCT-RCB (D. Ariz. 2007).

10 Plaintiffs/Appellants Reply Brief, Exhibit 46, p. 2-3, T. 130). Rudebusch v. Hughes, 313 F.3d 506 (9th Cir. 2002). See also Rudebusch v. Hughes, No. 95-CV-1313-PCT-RCB, No. CV-1077-PCT-RCB (D. Ariz.). Judgment entered June 28, 2004.

11 Plaintiffs/Appellants Reply Brief (Appendix 1 to Opening Brief, p. 8; 2 T. 84.). *Rudebusch v. Hughes*, 313 F.3d 506 (9th Cir. 2002).

12 Plaintiffs/Appellants Reply Brief (Citing Excerpt of Record, 147, pp. 25–26). *Rudebusch v. Hughes*, 313 F.3d 506 (9th Cir. 2002).

13 U.S. Department of Labor (DOL), Employment Standards Administration, Office of Federal Contract Compliance Programs (June 16, 2006). Voluntary Guidelines for Self-Evaluation of Compensation Practices for Compliance With Nondiscrimination Requirements of Executive Order 11246 With Respect to Systemic Compensation Discrimination; *Federal Register*, Vol. 71, No. 116. See also EEOC Compliance Manual, Section 10-III A.3.c. (*http://www.eeoc.gov/policy/docs/compensation.html#3*).

14 For example, if the R^2 of a given regression is 0.28 and the corresponding p-value is .03, one would also desire the Adjusted R^2 (*e.g.*, 0.24) to be statistically significant ($< .05$).

15 *Bazemore v. Friday*, 478 U. S. 385 (1986).

16 See for example, *Contractors Association v. Philadelphia*, 6 F.3d 990, 1007 (3d Cir. 1993); *EEOC v. General Tel. Co.*, 885 F.2d 575, 582 (9th Cir.), cert. denied, 498 U.S. 950 (1989); *Sobel v. Yeshiva Univ.*, 839 F.2d 18, 34 (2d Cir. 1988), cert. denied, 490 U.S. 1105 (1989); *Catlett v. Missouri Highway and Transp. Comm'n*, 828 F.2d 1260 (8th Cir. 1987), cert. denied, 485 U.S. 1021 (1988); *Palmer v. Shultz*, 815 F.2d 84, 101 (D.C. Cir. 1987).

17 *Rudebusch v. Hughes*, No. 95-CV-1313-PCT-RCB, No. CV-1077-PCT-RCB (D. Ariz.). Judgment entered June 28, 2004 (p. 16).

18 *Rudebusch v. Hughes*, 313 F.3d 506 (9th Cir. 2002) (at 520).

The Difference Between Differential Validity and Differential Prediction: TIME TO CLEAR UP THE MESS

Clifford R. Haimann | *George Mason University*
Alok Bhupatkar | *American Institutes for Research*

Differential validity and prediction are relevant for any industrial psychologist validating a selection device. In terms of psychological theory, these two concepts are important, but legal regulations also stress that differential prediction and validity should be examined. While distinguishing these two concepts is meaningful, examination of literature from the psychological and legal domains suggest that these concepts are not clearly differentiated. This work, therefore, highlights some of the confusion surrounding these two terms and describes a simple solution that could help anyone trying to either learn about these concepts or use them in their validation work.

Differential validity and differential prediction are two concepts that are relevant to researchers and practitioners concerned with validation analyses. While they are important for research, they are especially significant for Industrial-Organizational (I/O) psychology practitioners in a legal context because the Uniform Guidelines on Employee Selection Procedures (EEOC, 1978) posit that fairness analyses should be conducted whenever they are technically feasible. Based on the authors' initial understanding, differential validity entails differences in regression slopes and correlations between a predictor and criterion for different sub-groups. Differential prediction, in contrast, indicates that one common regression line for all subgroups will produce different residuals when one examines each sub-group using the common regression line; however, an exploration of the literature on psychological testing in order to verify the authors' conceptions has revealed that there are many often conflicting definitions of these two terms (see Table 1 for a summary of the different descriptions). An examination of court opinions has also revealed definitions that do not resolve, but only add to the confusion.

PURPOSE

Consequently, this work will highlight the inconsistencies and lack of differentiation between differential validity and prediction by drawing from the psychological literature and court opinions. The authors will then stress that the current time is an opportune moment to clear up the confusion. Other researchers have noted concern over the use of differential prediction, and they offered suggestions regarding the proper utilization of methods to test for it (Meade & Tonidendel, 2010); but, these authors did not focus on differential validity because the concept appears once in their article. Further, the current research explores court opinions about the topic and emphasizes different solutions that could help the legal and psychological fields.

PSYCHOLOGICAL PERSPECTIVE

The psychological examination of differential validity and prediction is very meaningful because the Uniform Guidelines (EEOC, 1978) specifically state that researchers should review the APA's views regarding fairness, which can be found in the APA Standards for Psychological and Educational Testing (American Educational Research Association/American Psychological Association /National Council on Measurement in Education, 1999). These standards can also be supplemented by the Principles for the Validation and Use of Personnel Selection Procedures (Society for Industrial and Organizational Psychology, 2003).

In Standard 7.6, the APA provides guidance about differential prediction. They say that differential prediction analyses involve the use of regression equations (or an equivalent) for different subgroups. They also state that correlations do not provide enough evidence for or against differential prediction and that researchers must consider both slopes and intercepts (American Educational Research Association/American Psychological Association /National Council on Measurement in Education, 1999). While the APA hints at differential validity by identifying the idea that correlations may vary by subgroup, they do not explicitly state what these different correlations mean. The Standards for Educational and Psychological Testing also identify the fact that one can look at slopes for different sub-groups, which is an indicator of differential validity, but it seems that the examination of these slopes is a subset of analyses one performs for differential prediction. Ultimately, a clear distinction is not made as the Standards say that one can examine correlations and slopes, but these analyses are not clearly discriminated from the components of differential prediction.

The SIOP Principles are also relevant when examining test fairness, and these guidelines identify what predictive bias is (the Principles say that this bias is also called differential prediction). They cite Cleary (1968) and state that differential prediction occurs when for one sub-group "consistent non-zero errors of prediction are made for members of the subgroup" (Society of Industrial Organizational Psychology, 2003, p. 32). They further note that testing for predictive bias involves the use of moderated regression where some criterion is regressed on a predictor, group membership, and the interaction of the two previous independent variables (Society for Industrial and Organizational Psychology, 2003). The predictive bias concept is clear in the Principles, however, the phrase differential validity does not occur within the document. Hence, a distinction is not made between the two terms.

The Standards and Principles do not make a well-defined distinction between differential validity and prediction and examination of some literature in psychological testing does not clear up the confusion. For example, one can compare recent articles that can be found in a short search on the PsychInfo Database. For instance, Meade and Fetzer (2009) highlight differential prediction, and the authors say that the Cleary regression-based method and similar analyses are appropriate methods to test for differential prediction. They state that one regresses a criterion on a predictor, group membership variable, and the interaction between the two independent variables to test for different prediction. Further, if the slopes are different, the authors claim that this is representative of differential validity.

This last statement about regression slopes does not completely align with Berry, Clark, and McClure's (2011) view on differential validity. These authors state that differential validity means differences between predictor-criterion correlations for different subgroups. Ultimately, the question arises as to whether differential validity means different slopes or different correlations? Do both of these signify differential validity as is noted in some articles (Hibbard et al., 2010)? If the two represent differential validity, an operational definition of the term should state that slopes and correlations can both be considered because the two are often not the same mathematically. A bivariate correlation between a predictor and criterion will only equal a regression slope in simple regression or when independent variables are completely uncorrelated. In the commonly used Cleary method described above with its three predictors, looking at bivariate correlations compared to regression slopes will not provide the same values. Hence, to avoid confusion, an operational definition should be provided that highlights both regression slopes and correlations.

Moreover, Kane and Mroch (2010) say in another piece that differential validity requires an examination of correlations and regression lines. This description creates more confusion because regression lines contain slopes and intercepts, so one could infer that correlations, slopes, and intercepts are all relevant to differential validity. Furthermore, the article gives consideration to Cleary's (1968) model of test fairness, which Meade and Fetzer (2009) and the SIOP Principles have connected to differential prediction. Given this connection between Cleary (1968) and differential prediction, it seems reasonable to presume that Kane and Mroch (2010) would mention this term in their work. It seems, however, that the two words appear only one time— in the references.

Hibbard et al. (2010) wrote another article relating to psychological measurement that potentially creates confusion about differential validity and prediction. The authors first state that differential validity can be established if one sees different correlations by sub-group. Secondly, an interaction between race (relevant in their article) and the predictor also entails differential validity. The researchers, however, list a third definition that seems incongruent with typical descriptions of differential validity. The authors write:

> *"A regression equation (with a single predictor) is formed across the entire sample predicting the criterion measure from the scale to be validated… Afterward, prediction errors (residuals) are computed as the difference between predicted scores and observed scores… if the two groups differ significantly by t test in mean prediction errors, this is evidence for differential validity (p. 353)."*

This definition aligns more with the description of differential prediction listed in the SIOP Principles, which entails examination of the errors of prediction (these errors are the same thing as the residuals mentioned by Hibbard et al., (2010)). Furthermore, this third test of differential validity highlighted by Hibbard et al. parallels work performed by other researchers when examining differential prediction. For instance, Mattern, Patterson, Shaw, Kobrin, and Barbuti (2008) tested the differential prediction of the SAT by looking at residuals for the different sub-groups under examination. Ultimately, it seems that there are inconsistencies between what the literature says and what practitioners describe in their work.

LEGAL PERSPECTIVE

Differential prediction and validity are ultimately significant in a legal context, and the Equal Employment Opportunity Commission (EEOC) looked to clarify the distinction between the two in 1978. According to this organization, differential validity is when a test instrument has different validity coefficients for sub-groups. Differential prediction, in contrast, occurs when a score on a predictor over- or under-predicts a criterion for certain groups (EEOC, 1978). This latter definition seems to parallel the text provided by SIOP, which centers on examining predicted scores as opposed to any kind of validity coefficient.

While the EEOC attempted to clarify the distinction between validity and prediction, an overview of court opinions suggests that some judges have not paid much attention. First, a search in LexisNexis Legal database for the term differential validity returns over 30 results. A search for the two words "differential prediction" produces zero results. This finding alone provides evidence that legal opinions do not distinguish between differential validity and differential prediction.

Moreover, some opinions that discuss differential validity do not correctly define the concept. For instance, in *U.S. v. The City of Erie Pennsylvania* (2005), the opinion said that differential validity meant the following in the context of physical ability testing: "If a man and a woman obtained the same score on the push-ups test, the woman's predicted job performance would be better than the man's." Such a definition aligns more accurately with differential prediction. If the slope or rate of change for one's criterion score given a unit increase in push up score was different for men compared to women, that would be differential validity. As other examples, differential validity has been referred to as: "whether it [a cognitive ability test] was as accurate a predictor for whites as for blacks" (*Dozier v. Chupka*, 1991). Such a definition is vague because the word accuracy is not clear. Is accuracy supposed to represent the validity or correlation between the test and an outcome? One cannot determine this fact.

AUTHORS' POINT OF VIEW

As this work has identified, there is still confusion regarding what is meant by differential validity and differential prediction. The APA and SIOP do not clearly differentiate the two, and academic articles relating to psychological assessment along with court opinions do not help clear confusion. Due to the fact that differential validity and prediction are terms that are rooted in psychological theory, it is the responsibility of psychologists and other professionals to accurately define the two concepts. Furthermore, while the legal field uses these terms, the Uniform Guidelines states that investigators should refer to the APA's Standards when considering fairness, and in a sense, the legal field is giving authority to the psychological domain in regards to test fairness. Also, judges and courts rely on psychologists acting as expert witnesses, and it is the responsibility of these witnesses to articulate clearly the difference between prediction and validity because the judges and lawyers do not have training in psychological theory.

Due to the fact that psychologists need to state the difference between differential prediction and validity, one resolution to the current confusion would be to describe the two in the APA Standards and note the differences between them. This is an opportune time to resolve this issue because the APA is currently revising the Standards. The SIOP Principles are written to mirror these Standards, and the Principles could also differentiate between the two terms. Currently, the Principles state what differential prediction is and also say what it is not. Specifically, they differentiate differential prediction as it is used in validation studies compared to its utilization in the job classification literature. The precedent has already been set for establishing what differential prediction is not, and the Principles could note that differential prediction is distinct from differential validity as well.

If the distinction between prediction and validity is made in the Standards or Principles, these two documents would need to address the fact that some have said that differential validity involves regression slopes or correlations. In regards to the relation between these two forms of validity and differential prediction, the common test for differential prediction, Cleary's Method, is distinct from correlational analyses for differential validity; however, significant interactions in a Cleary-moderated regression would signify different regression slopes for the groups in question. In this case, differential validity would be found as a result of analyses used to examine differential prediction. Ultimately, differential validity would be a subset of differential prediction analyses, and descriptions of these statistical concepts should be cognizant of this fact.

Anyone opposed to adding the distinction between differential validity and prediction to the Standards or Principles could argue that differential validity rarely exists (*e.g.*, Hunter, Schmidt, and Hunter, 1979) and that adding a clarification is not effective because one rarely sees differential validity. Recent evidence, however, (Berry et al., 2011) suggests that differential validity still exists, and concern for this topic is warranted. Furthermore, recent studies were conducted to test this concept (*e.g.*, Mattern et al., 2008; Callahan et al., 2010; Gardner & Deadrick, 2012), and it seems that practitioners are still examining differential validity. Lastly, the EEOC has not removed the term from the information used to support the Uniform Guidelines (EEOC, 1978), and given that Industrial/Organizational (I/O) psychologists are interested in the EEOC's documents, differential validity will likely be an important topic for years to come. Unless, SIOP and the APA note in the Principles and Standards that differential validity likely does not exist and that it should not be examined, these organizations should differentiate differential prediction from validity in order to clear up confusion for graduate students, academics, and practitioners.

TABLE 1: Different Descriptions of Differential Validity and Prediction

	Differential Validity	Differential Prediction
APA Standards	Not in the index	Use a regression model or similar tactic to analyze both slopes and intercept differences
SIOP Principles	Not in the document	"when for a given subgroup, consistent nonzero errors of prediction are made for members of the subgroup" (p. 32)
Meade & Fetzer (2009)	"differential slopes [in a regression analysis] indicate differential validity" (p. 741)	"…the term differential prediction much more accurately describes what is assessed by the regression-based procedure for evaluating the across groups equality of the relationship between the test and the criterion" (p. 749)
Berry, Clark, & McClure (2011)	"Differential validity refers to differences between subgroups in the correlation between a predictor and a criterion" (p. 881)	"Differential prediction focuses on differences between unstandardized regression slopes and intercepts relating the test and criterion across subgroups" (p. 882)
Kane & Mroch (2010)	"In evaluating the relationship between two measures across different groups (i.e., in evaluating "differential validity") it is necessary to examine differences in correlation coefficients and in regression lines" (p. 215)	Mentioned in the references only
Hibbard et al. (2010)	"A regression equation (with a single predictor) is formed across the entire sample predicting the criterion measure from the scale to be validated. Ethnicity is not a variable in this equation. Afterward, prediction errors (residuals) are computed as the difference between predicted scores and observed scores. Because these errors are theoretically randomly distributed… if the two groups differ significantly by t test in mean prediction errors, this is evidence for differential validity" (p.353)	Not mentioned

REFERENCES

American Educational Research Association/American Psychological Association/National Council on Measurement in Education. (1999). Standards for educational and psychological testing. Washington, DC: American Educational Research Association.

Berry, C. M. Clark, M. A. & McClure, T. K. (2011). Racial/ethnic differences in the criterion-related validity of cognitive ability tests: a qualitative and quantitative review. *Journal of Applied Psychology*, *96*, 881–906.

Callahan, C. A, Hojat, M., Veloski, J., Erdmann J. B., & Gonnella, J. S. (2010) The predictive validity of three versions of the MCAT in relation to performance in medical school, residency, and licensing examinations: A longitudinal study of 36 classes of Jefferson Medical College. *Academic Medicine*, *85*, 980–987.

Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement*, *5*,115–124.

Equal Opportunity Employment Commission, Civil Service Commission, Department of Labor, & Department of Justice (1978). Uniform Guidelines on employee selection procedures. Federal Register, 43, 38295–38309. Available from *http://www.uniformguidelines.com/uniformguidelines.html*

Gardner, D. G. & Deadrick, D. L. (2012). Moderation of selection procedure validity by employee race. *Journal of Managerial Psychology*, *27*, 365–382.

Hibbard, S., Tang, P. C. Y., Latko, R., Park, J. H., Munn, S., Bolz, S. & Sommerville, A. (2000). Differential validity of the Defense Mechanism Manual for the TAT between asian americans and whites. *Journal of Personality Assessment*, *75*, 351–372.

Hunter, J. E., Schmidt, F. L., & Hunter, R. (1979). Differential validity of employment tests by race: A comprehensive review and analysis. *Psychological Bulletin*, *86*, 721–735.

Kane, M. T. & Mroch, A. A. (2010). Modeling group differences in OLS and orthogonal regression: Implications for differential validity studies. *Applied Measurement in Education*, *23*, 215–241.

Mattern, K. D., Patterson, B. F., Shaw, E. J., Kobrin, J. L., & Barbuti, S. M. (2008). *Differential validity and prediction of the SAT* (College Board Research Report No. 2008-4). New York, NY: College Board.

Meade, A. Fetzer, M. (2009). Test bias, differential prediction, and a revised approach for determining the suitability of a predictor in a selection context. *Organizational Research Methods*, *12*, 738–761.

Meade, A., W. & Tonidandel, S. (2010). Final thoughts on measurement bias and differential prediction. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *3*, 224–227.

Society for Industrial and Organizational Psychology (SIOP) (2003). Principles for the validation and use of personnel selection procedures. Bowling Green, OH: Society for Industrial and Organizational Psychology.

CASE LAW

Dozier v. Chupka, 763 F. Supp. 1430 (6th Cir. 1991)

U.S. v. City of Erie, Pennsylvania. 411 F. Supp. 2d 524 (3rd Cir. 2005)

ADDITIONAL RESOURCES

AFFIRMATIVE ACTION PLANNING AND AUDIT RESOURCES

OFCCP Website .... *http://www.dol.gov/ofccp/index.htm*
OFCCP Frequently Asked Questions .... *http://www.dol.gov/ofccp/regs/compliance/faqs/offaqs.htm*
Census Data (2000 EEO File, Details) .... *http://www.eeoc.gov/stats/census/index.html*
Census Data (2000 EEO File, Live Data) *http://www.census.gov/eeo2000*
Census Data for Puerto Rico .... *http://www.dol.gov/ofccp/regs/compliance/PRicoeod.htm*
Zip Code/County Search .... *http://www.zip-codes.com/search.asp*
Construction AAP Guide .... *http://www.dol.gov/ofccp/TAguides/ctaguide.htm*
Uniform Guidelines (Regulation Format) ...... *http://www.dol.gov/dol/allcfr/ESA/Title_41/Part_60-30/toc.htm*
Uniform Guidelines (User-Friendly Format) ... *http://www.uniformguidelines.com*
AAP Fact Sheet .... *http://www.dol.gov/ofccp/regs/compliance/aa.htm*
Compliance Manual .... *http://www.dol.gov/federalregister/DocumentList.aspx?AgencyId=10&DocumentType=3*
State Job Bank Information .... *http://www.jobbankinfo.org*

FEDERAL REGISTERS, LAWS, AND REGULATIONS

Title VII of the Civil Rights Act of 1964 ...... *http://www.eeoc.com/policy/laws/title-vii-of-the-civil-rights-act-of-1964*
E-Verify .... *http://www.dhs.gov/xprevprot/programs/gc_1185221678150.shtm*
Americans with Disabilities Act .... *http://www.dol.gov/ofccp/regs/statutes/ada.htm*
Federal Register Notices .... *http://www.dol.gov/federalregister/DocumentList.aspx?AgencyId=10&DocumentType=2*
Regulations for AAPs (41CFR Chapter 60) ..... *http://www.dol.gov/dol/allcfr/ESA/Title_41/Chapter_60.htm*
Executive Order 11246 .... *http://www.dol.gov/ofccp/regs/statutes/eo11246.htm*
Section 503 of the Rehabilitation Act of 1973, as amended *http://www.dol.gov/ofccp/regs/compliance/sec503.htm*
Vietnam Era Veterans' Readjustment Assistance Act of 1974, as amended *http://www.dol.gov/ofccp/regs/statutes/4212.htm*
VEVRAA Fact Sheet *http://www.dol.gov/vets/programs/fact/vet97-5.htm*
Policy Directives *http://www.dol.gov/ofccp/regs/statutes/eo11246.htm*
G-FIVE *http://www.dol.gov/ofccp/regs/compliance/directives/dir282.htm*
Federal Contractor's Online Application Selection System *http://www.dol.gov/ofccp/regs/compliance/directives/dir281.htm*
Corporate Management Reviews *http://www.dol.gov/ofccp/regs/compliance/directives/dir202.htm*
Active Case Management (ACM) rules *http://www.dol.gov/ofccp/regs/compliance/directives/dir285.htm*
Investigative procedures when a test(s) is one cause of adverse impact in hiring *http://www.dol.gov/ofccp/regs/compliance/directives/dir267.htm*

Retention of Documents . . . .*http://www.dol.gov/ofccp/regs/compliance/directives/dir77_26.htm*
Calculating Interest on Back Pay *http://www.dol.gov/ofccp/regs/compliance/directives/dir280.htm*
Procedures for Reviewing Contractor
Compensation Systems and Practices *http://www.dol.gov/ofccp/regs/compliance/directives/dir307.htm*

ADDITIONAL AGENCY AND ORGANIZATION RESOURCES

EEOC Website . . . .*http://www.eeoc.gov*
EEO-1 . . . .*http://www.eeoc.gov/eeo1survey*
VETS-100 . . . .*https://vets100.vets.dol.gov*
National Industry Liaison Group (NILG) . . . . . .*http://www.nationalilg.org/main.html*
National Opinion Research Center (NORC)
data for graduate surveys . . . .*http://www.norc.org/homepage.htm*
Employment Resource Directory
(OFCCP Regional Outreach) . . . .*http://www.dol.gov/ofccp/ERRD/errsrvs.htm*
National Employment Law Institute (NELI) . . . . .*http://www.neli.org*
Society for Industrial & Organizational
Psychology (SIOP) . . . .*http://www.siop.org*
American Psychological Association . . . .*http://www.apa.org*

CALENDAR *of* EVENTS

DATE	ORGANIZATION	EVENT	ADDITIONAL DETAILS	CONTACT FOR MORE INFORMATION
3/3–3/6/2013	National Employment Law Institute	Employment Discrimination Law Update	Disney World Orlando, FL	www.neli.org
3/10–3/13/2013	Society for Human Resources Management	SHRM Employment Law & Legislative Conference	Washington, DC	www.shrm.org
3/14/2013	Sacramento Regional Inclusion & Diversity Council	Sacramento Inclusion & Diversity Conference	Roseville, CA	www.sactoid.com
3/14/2013	Personnel Testing Council of Northern California	PTC-NC Annual Conference	Sacramento, CA	www.ptcnc.org
3/17–3/20/2013	National Employment Law Institute	Employment Discrimination Law Update	Las Vegas, NV	www.neli.org
4/11/2013	National Employment Law Institute	OFCCP Mid-Year Update Webinar	Online	www.neli.org
4/19/2013	Arizona Industry Liaison Group	Quad A/AZILG Compliance Conference	Phoenix, AZ	www.azquada.org
4/25/2013	National Employment Law Institute	OFCCP Mid-Year Update Webinar	Online	www.neli.org
4/28–5/1/2013	International Public Management Association for Human Resources (IPMA-HR)	Eastern Region IPMA-HR Conference	Bethesda, MD	www.ipma-hr.org
5/1–5/3/2013	International Public Management Association for Human Resources (IPMA-HR)	Western Region IPMA-HR Conference	Orange, CA	www.ipma-hr.org
5/2–5/3/2013	National Employment Law Institute	Employment Law Conference	San Francisco, CA	www.neli.org
5/7–5/10/2013	American Association for Affirmative Action	AAAA National Conference	San Antonio, TX	www.affirmativeaction.org
5/9–5/10/2013	National Employment Law Institute	Employment Law Conference	Washington, DC	www.neli.org
5/16–5/17/2013	National Employment Law Institute	Employment Law Conference	Chicago, IL	www.neli.org
6/9–6/12/2013	International Public Management Association for Human Resources (IPMA-HR)	Central/Southern Region IPMA-HR Conference	New Orleans, LA	www.ipma-hr.org

* Dates and event information are subject to change. If you would like to submit an event for inclusion in the calendar, please e-mail editor@eeoinsight.com.

6/12/2013	The Personnel Testing Council of Metropolitan Washington	June 2013 Workshop	Michigan State University	www.ptcmw.org
6/16–6/19/2013	Society for Human Resources Management	SHRM Annual Conference	Chicago, IL	www.shrm.org
7/21–7/24/2013	International Personnel Assessement Council	IPAC 2013 Conference	Columbus, OH	www.ipacweb.org
7/30–8/2/2013	National Industry Liason Group	Annual Conference	Indianapolis, IN	www.nationalilg.org
9/21–9/25/2013	International Public Management Association for Human Resources (IPMA-HR)	Joint IPMA-HR/IPMA Canada International Training Conference	Las Vegas, NV	www.ipma-hr.org
10/22–10/24/2013	Society for Human Resources Management	SHRM Diversity & Inclusion Conference	Chicago, IL	www.shrm.org
10/27–10/29/2013	College and University Professional Association for Human Resources (CUPA-HR)	CUPA-HR Annual Conference	Las Vegas, NV	www.cupahr.org