Adverse Impact Analysis

BCGi: Adverse Impact & Test Validation Book Series

Overview of Biddle Consulting Group, Inc.

Affirmative Action Plan (AAP) Consulting and Fulfillment	• Thousands of AAPs developed each year • Audit and compliance assistance • AutoAAP™ Enterprise software
HR Assessments	• AutoGOJA™ online job analysis system • TVAP™ test validation & analysis program • CritiCall™ pre-employment testing for 911 operators • OPAC™ pre-employment testing for admin professionals • Video Situational Assessments (General and Nursing)
EEO Litigation Consulting /Expert Witness Services	• Over 200+ cases in EEO/AA (both plaintiff and defense) • Focus on disparate impact/validation cases
Compensation Analysis	• Proactive and litigation/enforcement pay equity studies • COMPare™ compensation analysis software
Publications/Books	• EEO Insight™: Leading EEO Compliance Journal • Adverse Impact (3rd ed.) / Compensation (1st ed.)
BCG Institute for Workforce Development	• 4,000+ members • Free webinars, EEO resources/tools
Nation-Wide Speaking and Training	• Regular speakers on the national speaking circuit

Biddle Consulting Group Institute for Workforce Development (BCGi)

- **BCGi Standard Membership (free)**
  - Online community
  - Monthly webinars on EEO compliance topics
  - *EEO Insight* Journal (e-copy)
- **BCGi Platinum Membership**
  - Fully interactive online community
  - Includes validation/compensation analysis books
  - EEO Tools including validation surveys and AI calculator
  - *EEO Insight* Journal (e-copy and hardcopy)
  - Members only webinars, training and much more…

www.BCGinstitute.org

Adverse Impact Presentation Outline

- Adverse Impact Overview & Background
- The Concept of Adverse Impact & Statistical Significance
- Adverse Impact for Hires, Promotions, Terminations: **Single & Multiple Events**
- Availability Comparisons

Why is this Topic Important to HR/EEO Professionals?

- *Why do I need to know about this topic?*
  - Federal law calls adverse impact that is not justified by validity evidence "disparate impact discrimination"
  - 90%+ of OFCCP settlements are related to adverse impact
  - Over the last few years, the EEOC has focused more on "systemic investigation" and enforcement
- *What are the key essentials I need to know about this topic?*
  - For federal contractors, one of the most critical parts of the AAP has to do with recordkeeping and adverse impact
  - Proper adverse impact analyses need to reflect the reality of your employer's hiring and promotional process, not just "push button, aggregated" data

Why is this Topic Important to HR/EEO Professionals? (cont.)

- *What are the key essentials I need to know about this topic?*
  - Adverse impact analyses should be conducted annually
  - Adverse impact can take several different forms, and many different types of proven procedures exist for computing each
- *What are the consequences surrounding these issues?*
  - Every conciliation agreement, consent decree, or legal case has the possibility of leading to negative press.
  - Employers don't want to unfairly discriminate, which is what unjustified adverse impact can possibly lead to
  - Adverse impact can be an automated "audit trigger"
  - The typical "start up" cost for an adverse impact case exceeds $30k

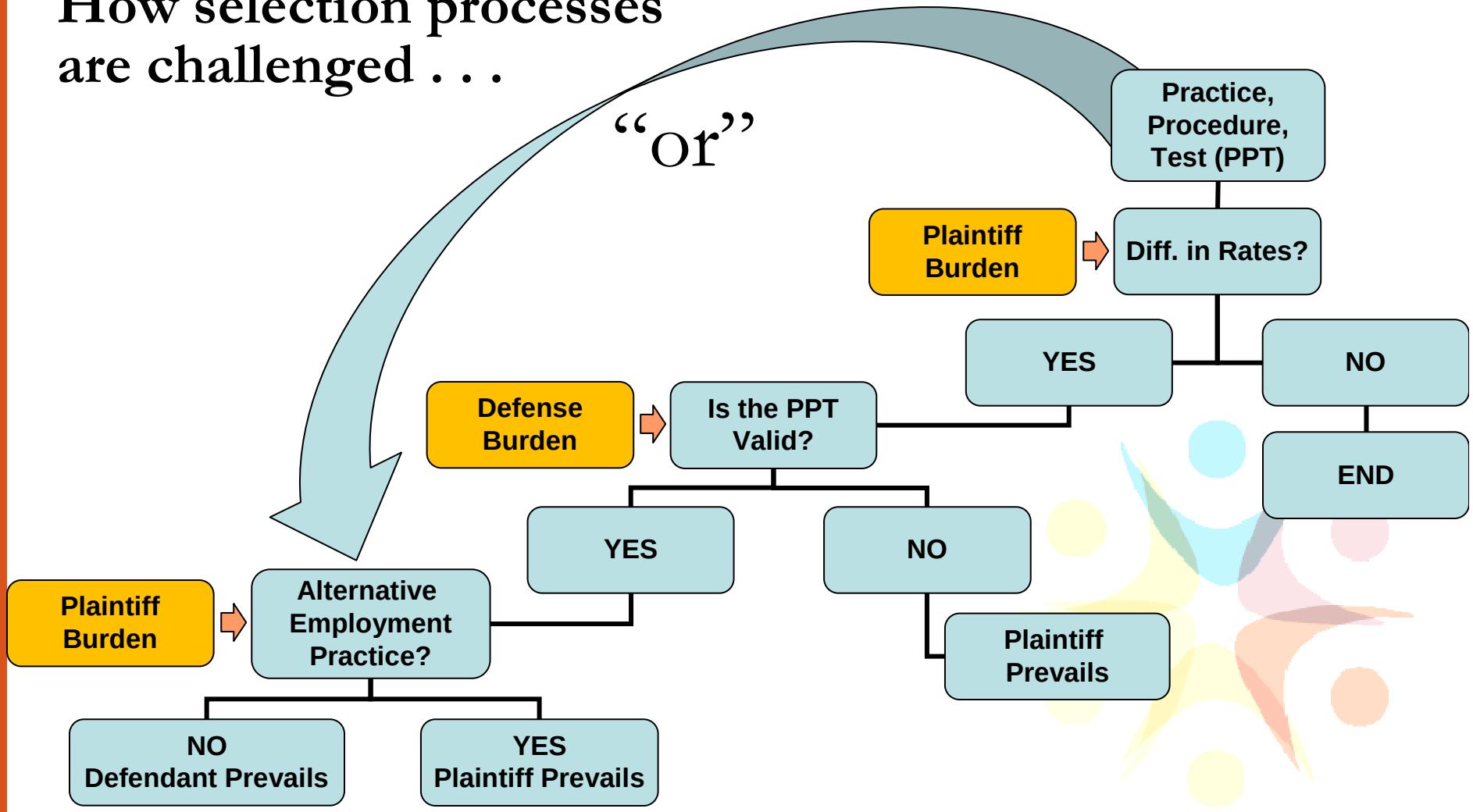
Adverse Impact Overview & Background

History and Development

- 1964 Civil Rights Act
- 1971 Griggs v. Duke Power Company
- 1972 TACT Committee
- 1972-1978 Uniform Guidelines Development
- 1978 Uniform Guidelines
- 1989 Wards Cove v. Atonio
- 1991 Civil Rights Act
- Courts & "Statistical Significance"

Current Legal Context: Adverse Impact Discrimination Flowchart

How selection processes are challenged . . .



Adverse/Disparate Impact: Legal Overview

DISPARATE IMPACT

An unlawful employment practice based on disparate impact is established only if:

1. A complaining party demonstrates that a respondent uses a particular employment practice that causes an adverse impact

and

2. the respondent fails to demonstrate that the challenged practice is job-related for the position in question and consistent with business necessity

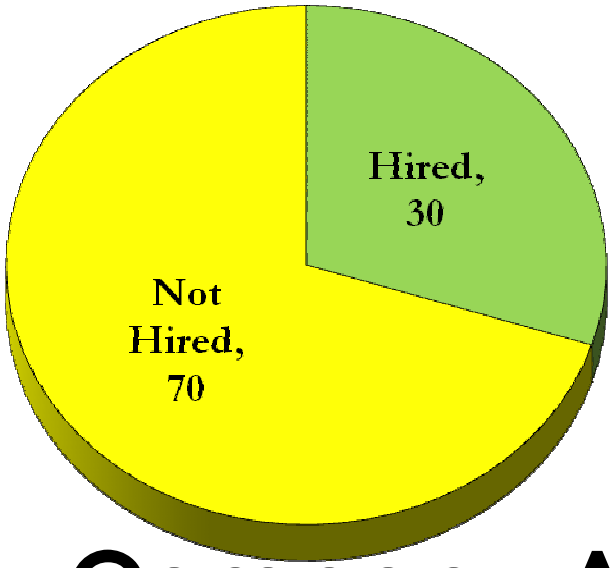
or

3. the complaining party makes the demonstration described above with respect to an alternate employment practice, and the respondent refuses to adopt such alternative employment practice.

Forms of Adverse Impact

Adverse Impact – Basics

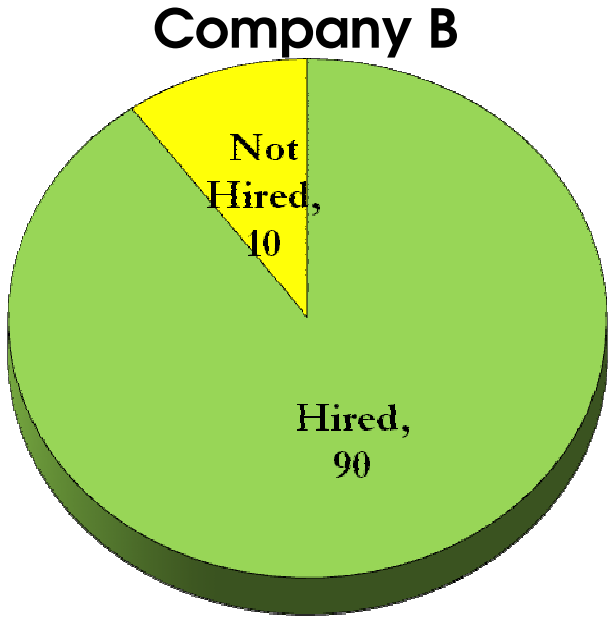
- 100 African Americans applied for a job



Company A Company B

- Which has AI against African Americans?
- Insufficient information

Adverse Impact – Basics



- What if 100 White Applicants Applied and:

Situation	Hired	Not Hired
A	15	85
B	99	1

Adverse Impact – Basics

- Descriptive statistics (percentages and counts) are insufficient
- Descriptive statistics are only ½ the picture.
- Adverse Impact conclusions are based on comparisons.
  - Comparisons against "Reference Comparator"
  - Comparisons help to provide interpretable meaning to observed percentages.

Adverse Impact – Basics

- There are 2 types of Reference Comparators
  - **Selection Rate**
    - **Example:** 30/100 White applicants were hired
  - **Availability**
    - **Example:** Of available workforce, 80% are African American
- Interpretation of African Am. Hires:

Company	Hired	Not Hired
A	30	70
B	90	10

Two Types of Adverse Impact

SELECTION RATE COMPARISON

- 2 X 2 Table Comparison
- Hires, promotions, terminations
- "Hypergeometric"

Men Pass	Women Pass
Men Fail	Women Fail

AVAILABILITY COMPARISON

- Utilization Analysis
- Single Group Test
- "Binomial"

Availability %
Women
Total

When Does Adverse Impact Result in "Disparate Impact Discrimination"?

SELECTION RATE COMPARISON

- 2 X 2 Table Comparison
- Evaluates hires, promotions, terminations
- "Hypergeometric"

Statistically Significant Result

+

No Job Relatedness / Validity

=

Disparate Impact Discrimination

AVAILABILITY COMPARISON

- Utilization Analysis
- Single Group Test
- "Binomial"
- See p. 58955 of Int. App Regs

Statistically Significant Result

+

6 "Possible Ingredients"

=

"Adverse Inference" or Evidence for Disparate Treatment Cases

Adverse Impact Analysis – Road Map

- There are 2 types of Adverse Impact Analysis
  - Selection Rate
  - Availability
- Each type can be structured in 2 forms
  - Single Event, e.g. one job, test, decision
  - Multiple Events, e.g. multiple jobs, years, decisions
- Road Map

	Selection Rate	Availability
Single	**A**	**C**
Multiple	**B**	**D**

Section 1:

The Concept of Adverse Impact & Statistical Significance

The Concept of Statistical Significance

- Statistical Significance (Thresholds):
  - 5%
  - 0.05
  - 1 chance in 20
  - 2.0 Standard Deviations *(actually 1.96)*
- Statistical Significance (Outputs)
  - Lower *p*-values=higher SD (or "Z") values
  - For example:
    - *p*-value: 0.05 = 1.96 SDs
    - *p*-value: 0.01 = 2.58 SDs

Statistical Significance and Power

- Statistical significance: The point at which differences become large enough that one can claim a trend exists.
- Statistical power: The ability to see those trends if, in fact, they do exist.
- Statistical power is directly related to effect size and sample size:
  - **Effect size**: The size of the difference in selection rates between two groups . . . the larger the difference the less number of transactions necessary to detect statistical significance
  - **Sample size**: With larger numbers of transactions it becomes much easier to detect statistical significance

Statistical Power

	Men		Women		Statistical
	Counts (#)	Percent (%)	Counts (#)	Percent (%)	FET (*p*)
A	100	50%	90	45%	**0.343**
	100	50%	110	55%	
B	200	50%	180	45%	**0.168**
	200	50%	220	55%	
C	300	50%	270	45%	**0.088**
	300	50%	330	55%	
D	400	50%	360	45%	**0.048**
	400	50%	440	55%	

- Larger counts can lead to artificially inflated statistical power

Statistical Power

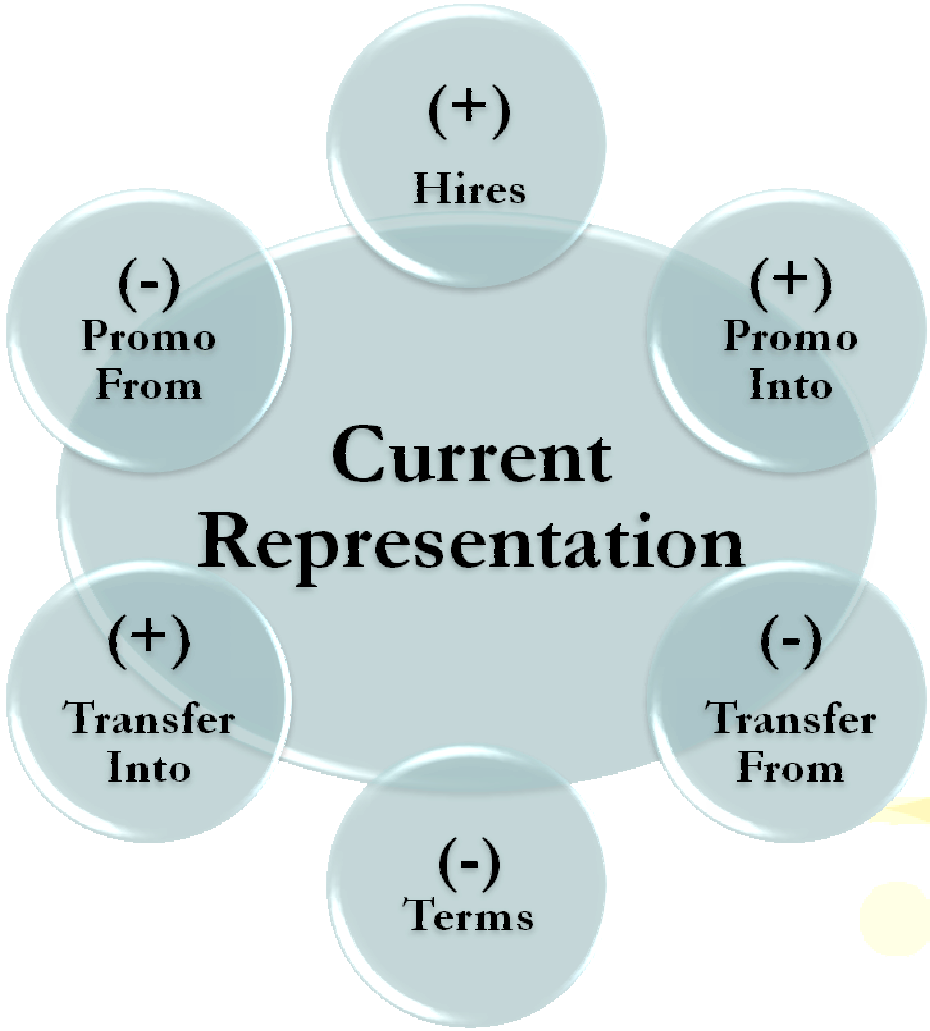
- Enforcement agencies have no control over effect size (i.e., the difference in selection rates), but they do have some control over sample size . . . which is why they often request two (2) years worth of data to analyze.
- However, simply aggregating all applicants and all hires across strata (as is typically done), can sometimes result in incorrect/misleading findings.

Section 2:

AI for Hires, Promotions, Terminations: Selection Rate Comparisons for Single Events

	Selection Rate	Availability
Single	**A**	**C**
Multiple	**B**	**D**

Comparison of Selection Rates: How We Got to Where We Are Today



"Impact Ratio Analysis" (IRA)

- The Impact Ratio Analysis (IRA) provides a single metric describing one group's success rate compared to another
- IRA evaluates whether a practice, procedure or test (PPT) results in disproportionate selection rates by gender, race/ethnic, or age group.

"Impact Ratio Analysis" (IRA)

- 2×2 Table
- Example

	Pass	Fail	
Female	4	6	10
Male	7	7	14

  - Female *passing rate*: 4/10=40%
  - Male *passing rate*: 7/14=50%
  - Impact Ratio = 40%/50%, is 80%
- Is this statistically significant?

Statistical Evaluation of 2×2 Tables

- Statistical tests determine if observed difference is:
  - Random chance
  - Significant
    - Probability (p) ≤ 0.05
- Statistical tests for 2×2 Tables:
  - Fisher Exact Test (FET)[1]
  - Chi-Square (χ^2)

Note: [1]BCG recommends FET with Lancaster's mid-p correct

Proper Statistical Test for 2×2 Tables

- Chi-Square (χ^2)
  - Appropriate for larger sample sizes
  - Too powerful for small sample sizes
- Fisher's Exact Test (FET)
  - Appropriate for small sample sizes
  - Too conservative
  - Appropriate for fixed margin 2×2 Tables
- Fisher's Exact Test (Lancaster's Mid-p)
  - In between FET and χ^2
  - Is a good all around statistical for 2×2 Tables

Statistical Evaluation of 2×2 Tables

- Example

	Pass	Fail	
Female	4	6	10
Male	7	7	14

  - Female *passing rate*: 4/10=40%
  - Male *passing rate*: 7/14=50%
  - Impact Ratio = 40%/50%, is 80%
  - FET mid-*p* = 0.55

Section 3:

AI for Hires, Promotions, Terminations: Selection Rate Comparisons for Multiple Events

	Selection Rate	Availability
Single	A	C
Multiple	B	D

Single Event v. Multiple Event Analyses

ALL applicants and ALL hires throughout the time period

Men Pass	Men Fail
Women Pass	Women Fail

= Chi-Square or Fisher's Exact

Event 1 + **Event 2** + **Event 3**

Men Pass	Men Fail
Women Pass	Women Fail

Men Pass	Men Fail
Women Pass	Women Fail

Men Pass	Men Fail
Women Pass	Women Fail

Adverse Impact Across Years: Simpson's Paradox

EXAMPLE

Testing Year	Group	# Applicants	# Selected	Selection Rate %
2004 Test	Men	400	200	50.0%
	Women	100	50	50.0%
2005 Test	Men	100	20	20.0%
	Women	100	20	20.0%
2004 + 2005 Tests Combined	Men	500	220	**44.0%**
	Women	200	70	**35.0%**

Mantel-Haenszel

Mantel-Haenszel (MH) Defined

- In the context of selection rate comparison analyses (UGESP 4D), the MH:
  - is a statistical tool that allows researchers to appropriately combine separate and distinct selection processes into a single analysis
  - appropriately allows for the benefits of increased sample size while controlling for Simpson's Paradox
  - can be used to analyze an overall selection process over time **OR** an individual practice, procedure, or test over time
- The MH is a useful tool for evaluating whether the employer has a "pattern and practice" that is possibly discriminatory

Mantel Haenszel v. FET

EXAMPLE

Testing Year	Group	# Applicants	# Selected	Selection Rate %
2004 Test	Men	400	200	50.0%
	Women	100	50	50.0%
2005 Test	Men	100	20	20.0%
	Women	100	20	20.0%
2004 + 2005 Tests Combined	Men	500	220	44.0%
	Women	200	70	35.0%

- **Single Event Method:**
  - **FET mid-*p*: SD = 2.16 (Significant)**
- **Multiple Events Analysis**
  - **Mantel-Haenszel: SD = 0.02 (NOT Significant)**

Section 4:

Availability Comparisons for a Single Event

	Selection Rate	Availability
Single	A	**C**
Multiple	B	D

Comparison of Incumbency to Availability: What We Do Look Like Compared to What We "Should" Look Like

Comparison of Incumbency to Availability

- Regulations require contractors to compare the percentage of minorities and women in each job group with the availability for those job groups determined in the availability analysis

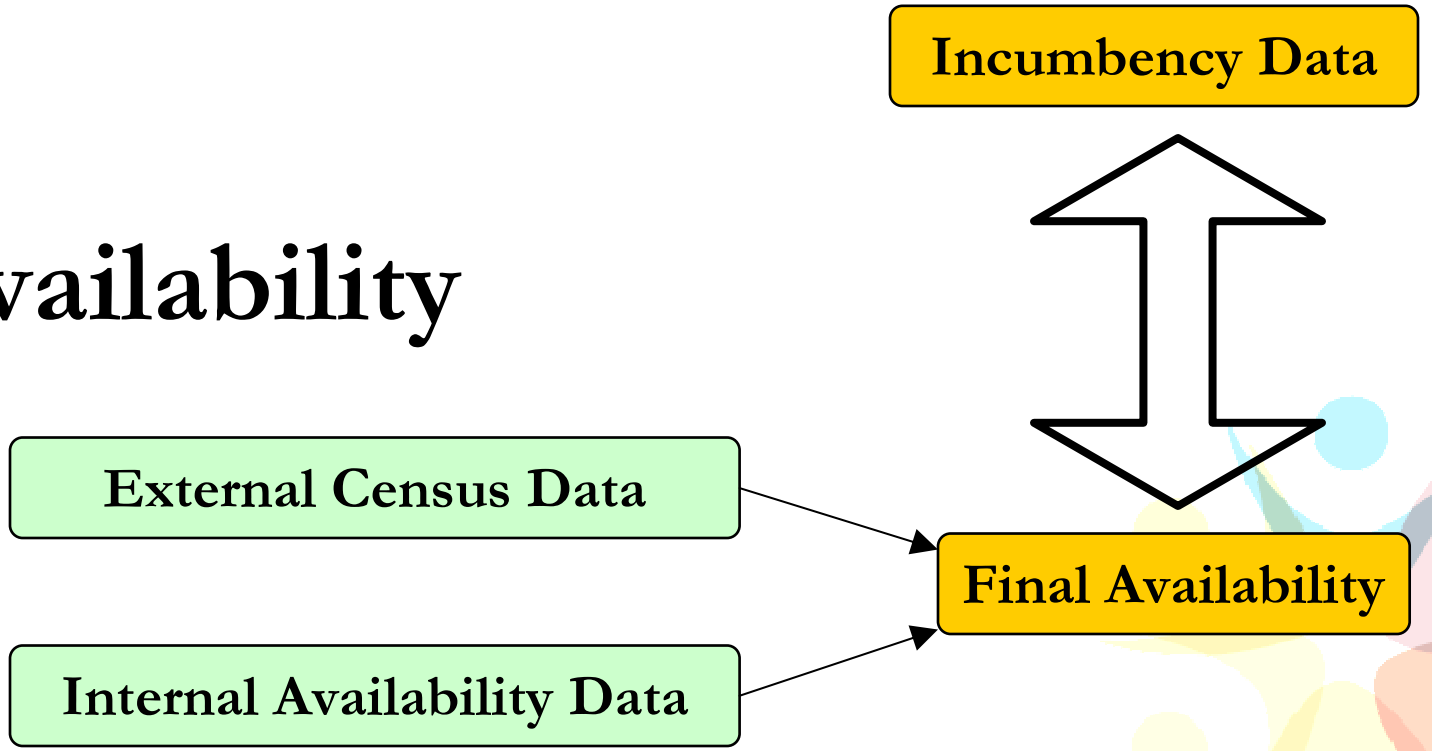
- When the percentage of minorities or women employed in a particular job group is ***less than would reasonably be expected*** . . . the contractor must establish a placement goal and create action-oriented programs associated with that goal

Comparison of Incumbency to Availability

Incumbency

Actual Workforce Representation/Headcount

Availability



Comparison of Incumbency to Availability

- How is "less than would reasonably be expected" defined?
  - Any Difference: Is there any difference between incumbency and availability?
  - Whole Person Rule: Is the difference between incumbency and availability at least one whole person?
  - 80% Rule: Is incumbency at least 80% of availability?
  - Statistical Significance: Is the difference between incumbency and availability statistically significant?

Important Note: Identifying underutilization is NOT a declaration of discrimination. Choose a rule that best represents your organizational size/structure and how it views/perceives affirmative action.

Comparison of Incumbency to Availability

Q: When ***can*** underutilization lead to a finding of discrimination?

A: When one (1) of six (6) additional ingredients is added:

1. Failure to **keep applicant records** (sometimes referred to as an "adverse inference"—see 4D of the Guidelines)
2. Failure to **run/keep adverse impact analyses** on the selection or promotional processes (also an "adverse inference"—see 4D of the Guidelines)
3. Discriminatory **recruiting practice** (e.g., Hazelwood School District v. United States)

Comparison of Incumbency to Availability

Q: When ***can*** underutilization lead to a finding of discrimination? (cont.)

4. Discriminatory reputation **"chilled"** or **"discouraged"** certain group members from applying
5. Promoting employees through **"appointment only"** process (rather than conducting track-able promotional processes)
6. **Invalid "Basic Qualifications"** (see p. 58955 of Int. App Regs)

Utilization analyses that are significant (based on either the employer's availability or "proxy" availability data) "plus" any of these factors can ***possibly*** lead to a finding of discrimination.

Comparison of Incumbency to Availability

- Unless one or more of the 6 ingredients exist, statistically significant underutilization should not be directly equated with discrimination
- Several other factors can sometimes explain underutilization:
  - Job interest
  - Occupational qualifications
  - Labor trends
  - Traditional roles (e.g., engineering vs. clerical)
- Unless one of the "6 ingredients" exist, a specific practice, procedure, or test will need to be identified that caused the adverse impact (using statistical significance tests). The only exception is if the agency's practices cannot be "separated for analysis purposes" (see 1991 Civil Rights Act)

Section 5:

Availability Comparisons for Multiple Events

	Selection Rate	Availability
Single	A	C
Multiple	B	**D**

Single Event v. Multiple Event Analyses

Single utilization analysis

Exact Binomial

Incumbency (%)	Availability (%)

Multiple utilization analysis,

- e.g. multiple years, locations, positions.

Event 1

Incumbency (%)	Availability (%)

+

Event 2

Incumbency (%)	Availability (%)

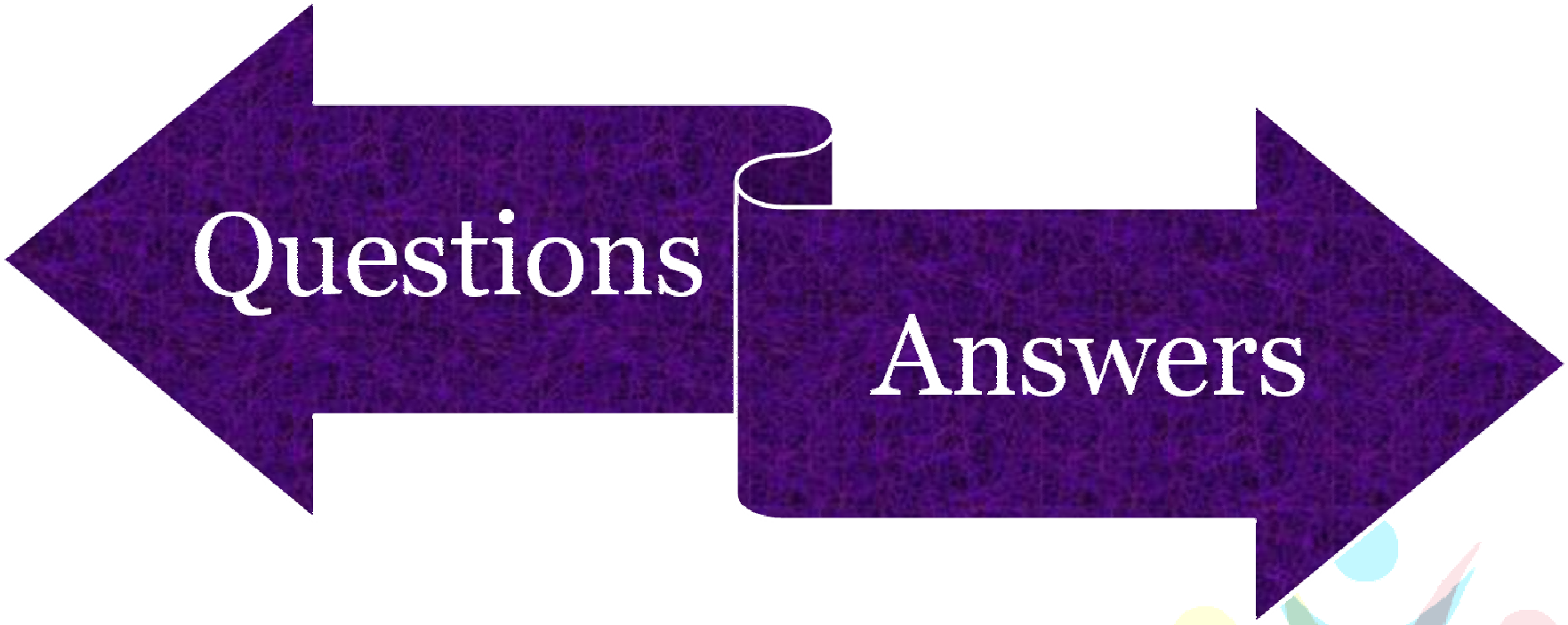
+

Event 3

Incumbency (%)	Availability (%)

Multiple Events Availability Analysis

- Use Multiple Events Exact Binomial models
  - Generalized Binomial Test
- Proper multiple events model avoids:
  - Artificially inflated statistical power: overly aggregated data can trigger with small differences
  - Lack of statistical power: overly disaggregated data lacks statistical power



THANK YOU

Contact us:

- **Biddle Consulting Group Institute for Workforce Development (BCGi)**

www.BCGinstitute.org