

Evaluating & Selecting Your AI Vendor

# Vetting A.I. Vendors

Brian Marentette, Ph.D. & Dan (dk) Kuang, Ph.D.



# Disclaimer

- This presentation is NOT LEGAL ADVICE.
  - We will touch on topics that may have legal ramifications.
  - The presentation is “technical” in nature: the mechanics and methods of personnel selection and adverse impact analysis.
  - We will lightly touch on a variety of topics, springboard for further learning



# Contact Information



## Brian Marentette, Ph.D.

Director of People Insights

[bmarentette@biddle.com](mailto:bmarentette@biddle.com)

[www.linkedin.com/in/brianmarentette](http://www.linkedin.com/in/brianmarentette)



## Daniel (dk) Kuang, Ph.D.

Director

[dkuang@resecon.com](mailto:dkuang@resecon.com)

[www.Linkedin.com/in/dankuang/](http://www.Linkedin.com/in/dankuang/)



# Purpose

- AI-based decision making in the employee lifecycle
  - Examples: Hires, Promotions, Terminations,
- AI-Driven Employment Decisions are “regulated”
  - Laws, e.g., NYC (LL-144), CO (SB-205), IL (HB-3773)
  - OFCCP: <https://www.dol.gov/agencies/ofccp/ai/ai-eeo-guide>
  - EEOC: <https://www.eeoc.gov/newsroom/eeoc-releases-new-resource-artificial-intelligence-and-title-vii>
  - ODEP: <https://www.peatworks.org/ai-inclusive-hiring-framework/>
  - Uniform Guidelines





# Purpose

- A LOT of AI-Tools
- A LOT of AI-Vendors
- How do you choose?



# Overview: Vetting AI Vendors

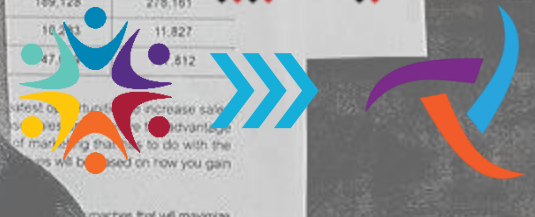
- General Questions to Consider
- Quality of Training Data
- Analyzing & Evaluating the AI Tool
- Use Case—How Will it Be Used?





# General Questions

## The Basics





# Overview: General Questions

- The Scientist(s)
- The Science
- Maturity





# The Scientist(s)

## Who developed the AI Tool?

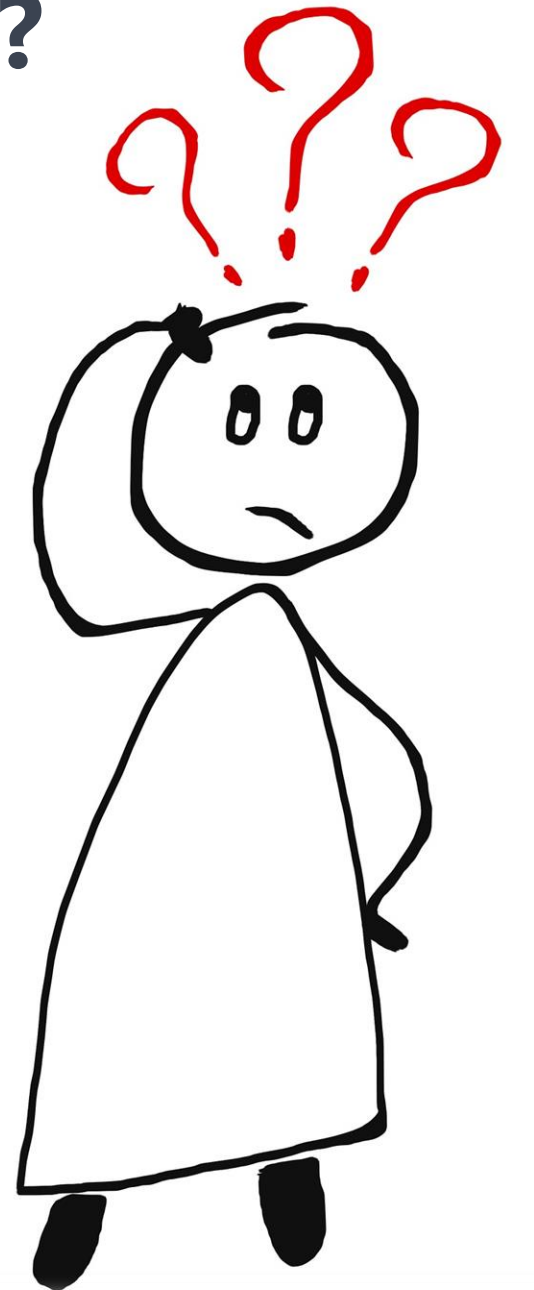
- What is their expertise?
- Are they assessment experts (I/O Psychologists)?
- How much experience with assessment?



# The Science

## What is the theoretical framework for the assessment?

- Why should the AI be effective???
- Examples
  - Math test for Accountants
  - Lifting test for warehouse stocker
  - Extroversion (outgoing-ness) for Sales
- Face Validity—Rational
- Non-theoretical framework
  - ~ok



# Maturity

**How long has the instrument been in use?**

- AI technology is pretty new

**Why does maturity matter?**

- Experience with Use Cases
- Development Depth
- Longevity and Durability
- Reliability
- Assurance



This Photo by Unknown Author is licensed under CC BY





# Training Data Quality

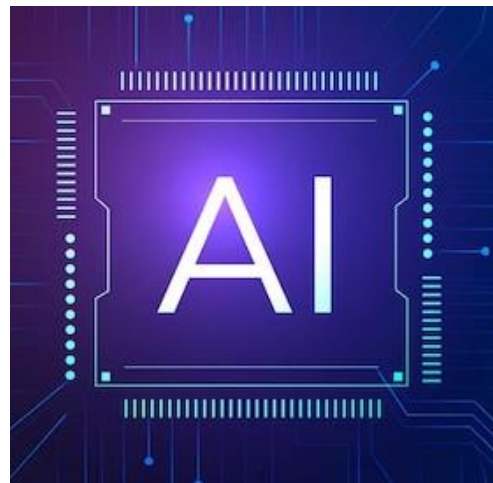
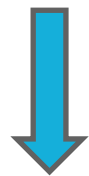
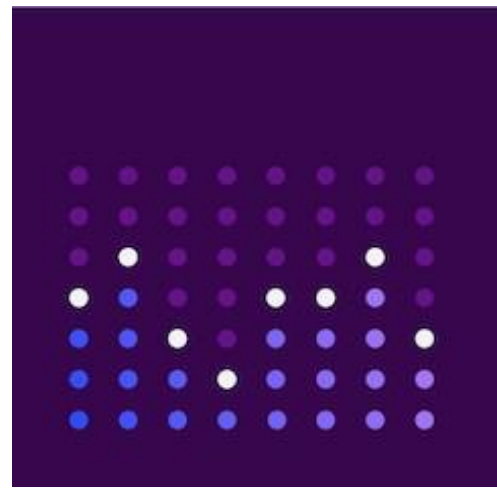




# Training Data – How Its Used

## TRAINING DATA

- Resumes
- Interview responses
- Personnel data



## MACHINE LEARNING

- Pattern recognition
- Identify relationships among data



## NEURAL NETWORK

- Create models
- Make predictions



## PERFORM TASKS

- Apply models
- Produce content
- Make decisions
- Produce inferences



Image credits: Freepik.com



# Training Data Considerations - Sources

- LOTS of data is needed – where will it come from?
- Normative data
  - Other organizations
  - Publicly available data
  - “Ground Truth” data
- Local data
  - Your applicants
  - Your employees





# Training Data Considerations - Qualities

- Work/jobs of individuals
- Industries (e.g., retail, biomedical, finance)
- Age, race, and gender composition
- Geography (country, region, area)
- Disability status and type (visual impairment, speech impairment, neurodiversity/cognitive impairments, etc.)
- Tenure of employees
- Qualifications (knowledge, skills, abilities)



# Training Data Considerations - Transportability

## Patterns May Differ by:

- Job levels
- Gender, Race
- Disability status
- Geography (speech rate, word choice by region)
- Industry
- Others

## Predictions may:

- Be only accurate for one group
- Not incorporate nuances of certain populations
- Miss critical information

## Outcomes may:

- Be inaccurate
- Introduce possible bias
- Pose legal risks





# Analyzing & Evaluating AI Tool





# Overview: Psychometric

## Scientifically Analyze AI Tool to Evaluate Performance

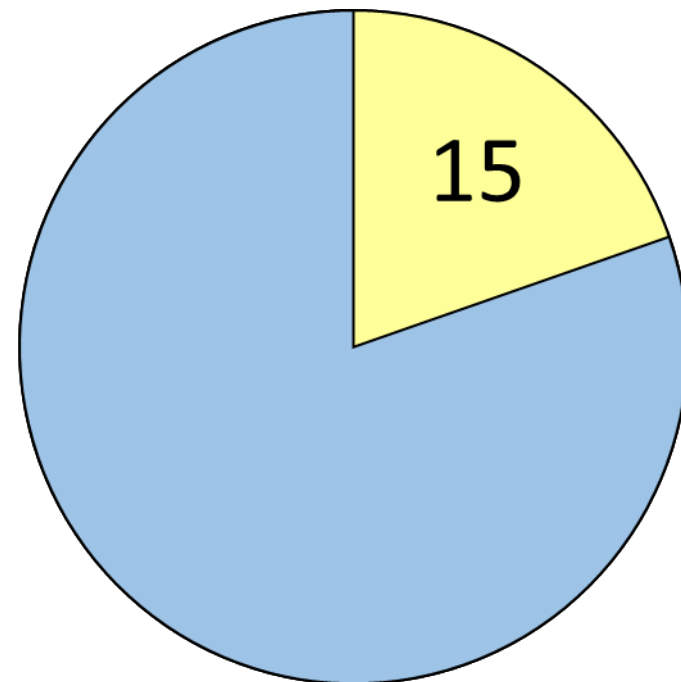
- Compliance Analytics
- Uniform Guidelines
- AI Maintenance



# Compliance Analytics

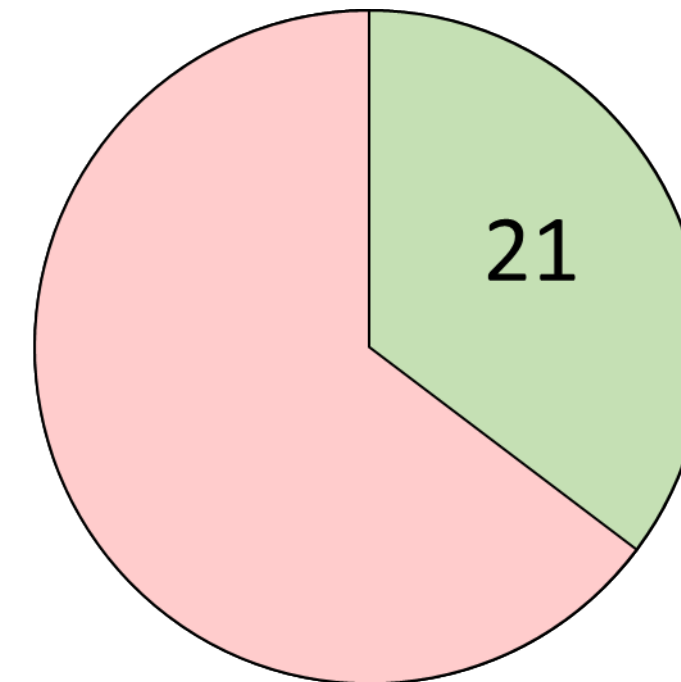
## “AI Bias Audit”

- Selection Rate Analysis



100-Applicants

Selection Rate:  $\frac{15}{100} = 15\%$



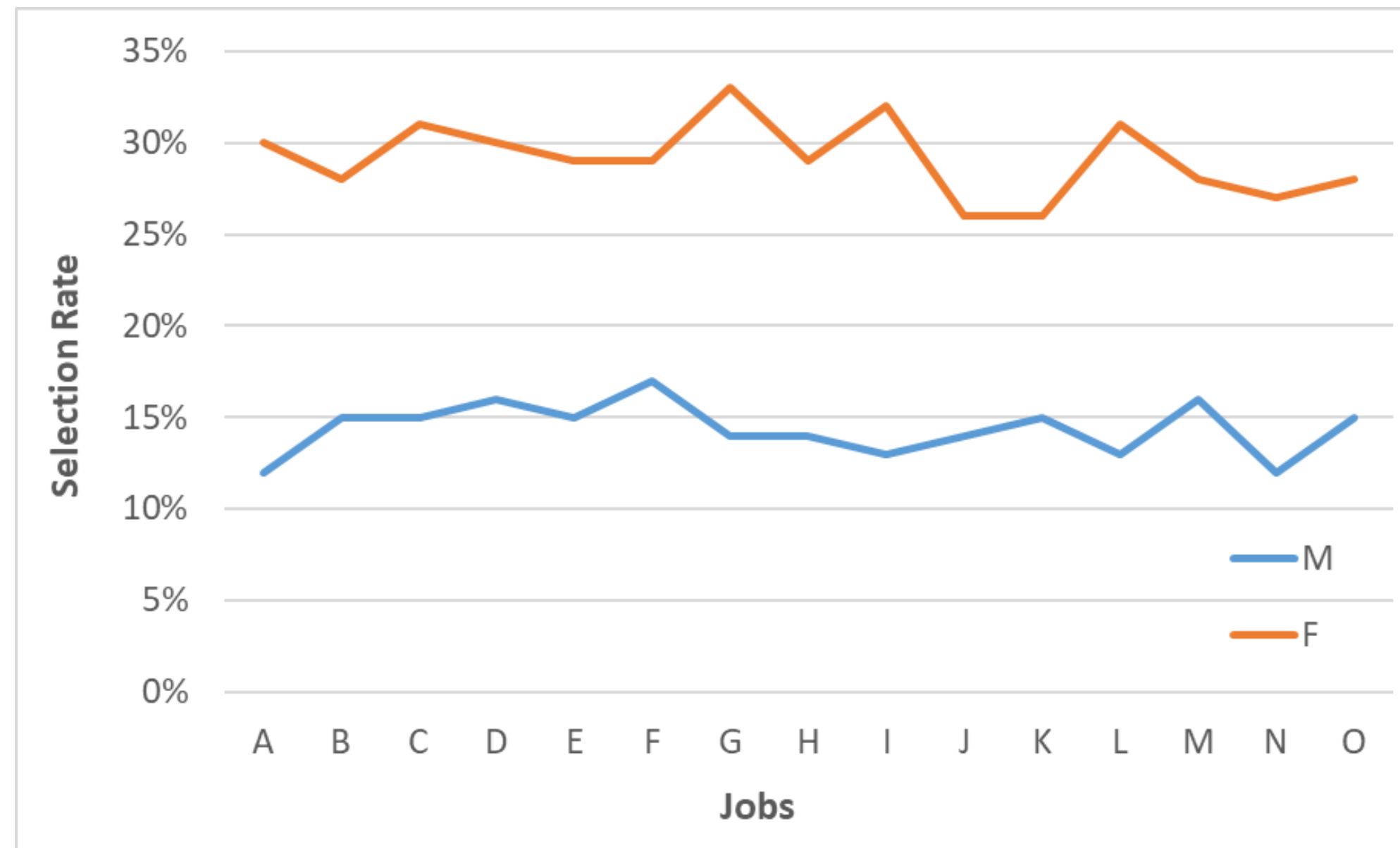
70-Applicants

Selection Rate:  $\frac{21}{70} = 30\%$



# Compliance Analytics

- Selection Rate Analysis





# Compliance Analytics

- Selection Rate Analysis



# Uniform Guidelines

## Reliability

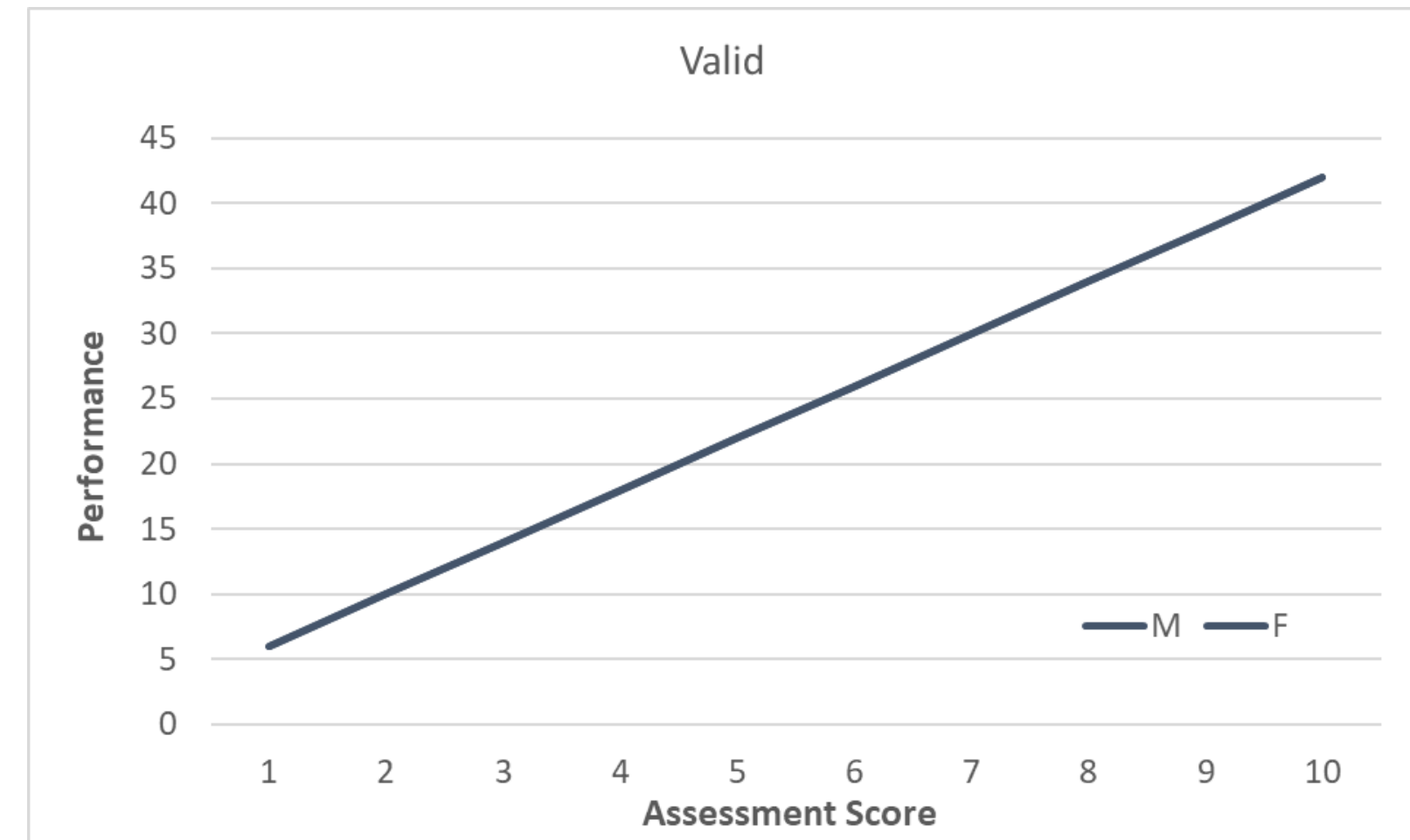
- Guidelines 15.B.8
- Consistency
- Measurement Error



# Uniform Guidelines

## Validity

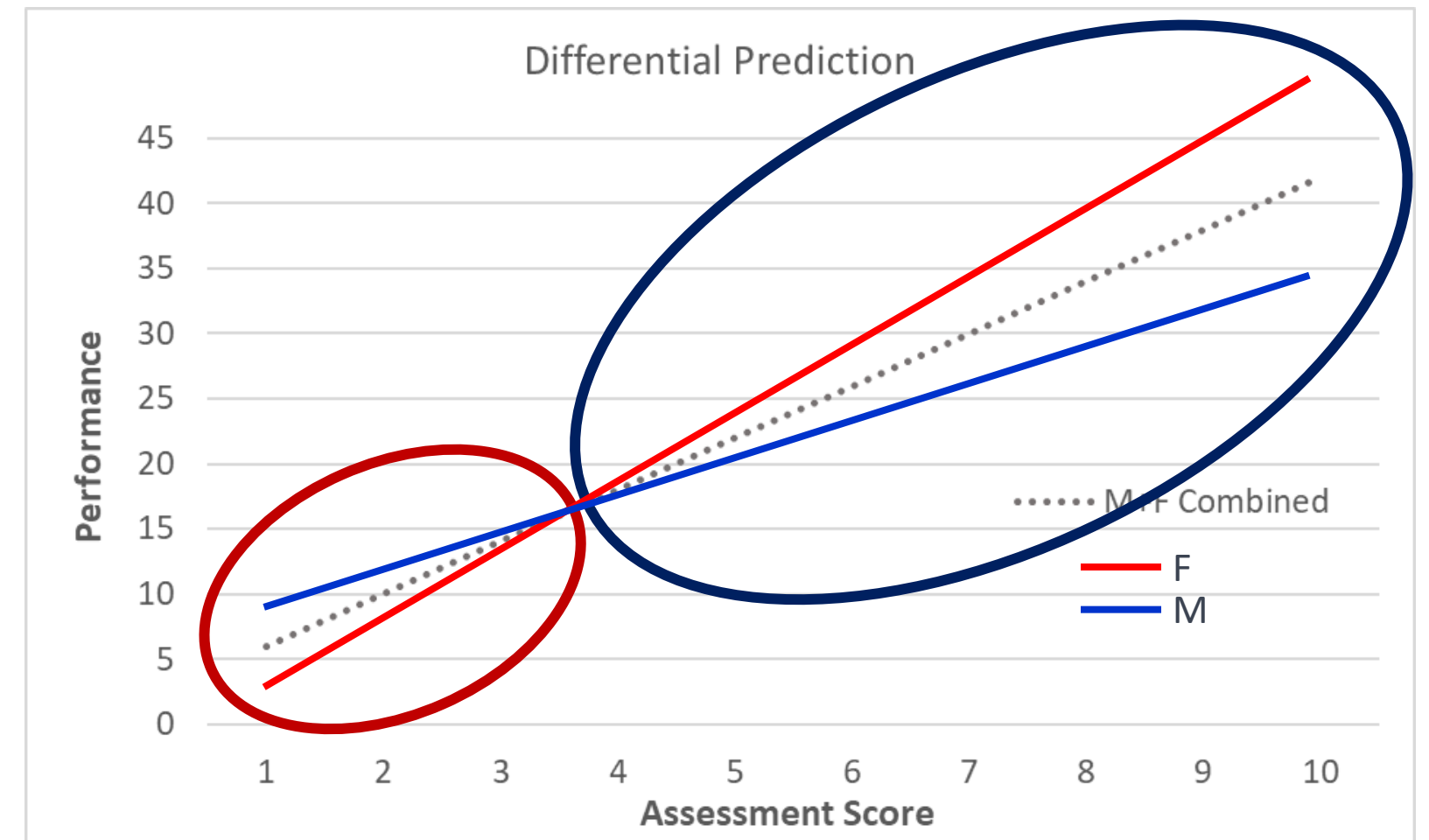
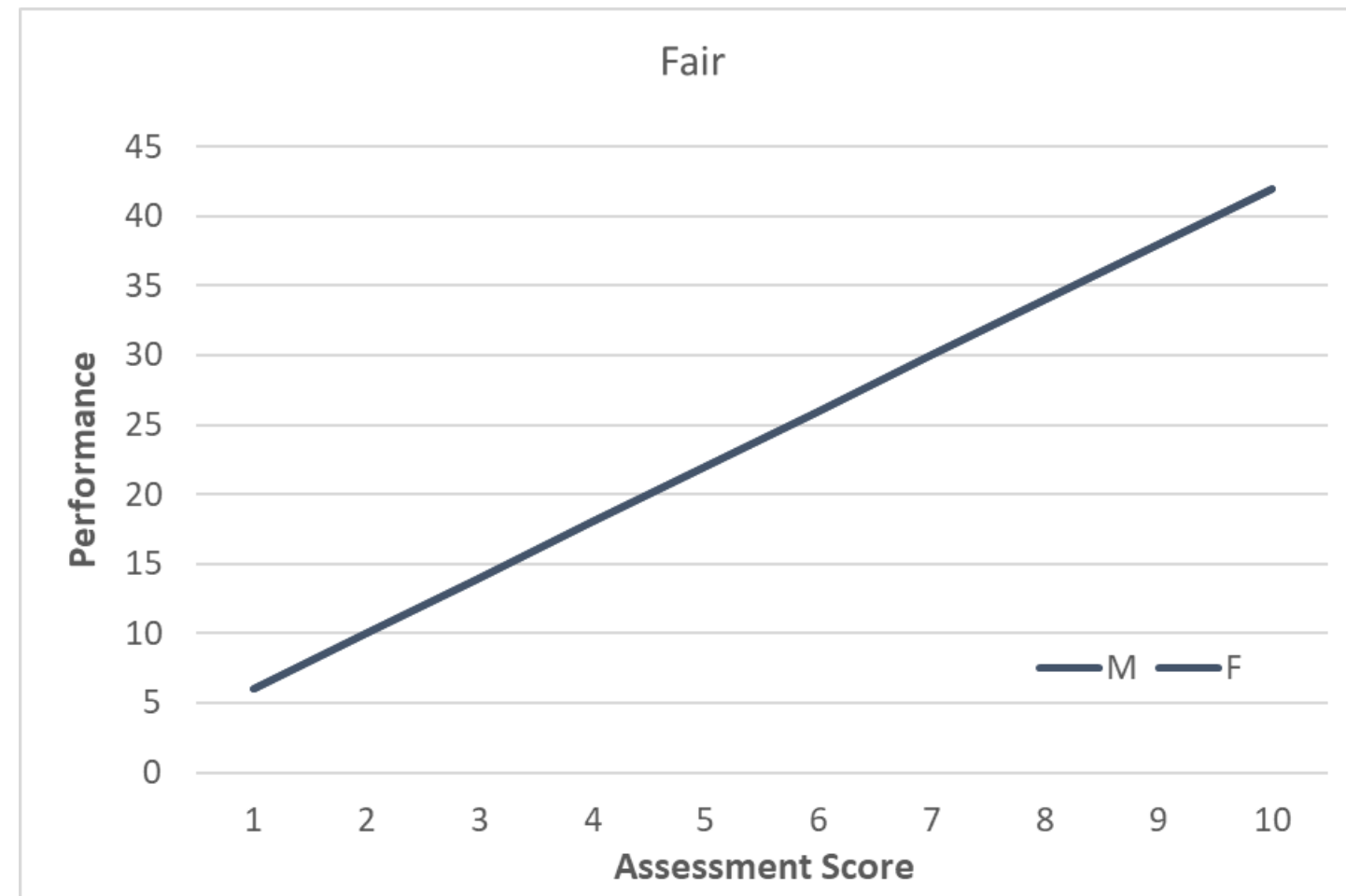
- Job Related and Consistent with Business Necessity
  - Is AI Scores related to a Criterion?
- Situation Specificity
  - Data Quality (previous section): Is the AI Model Relevant?
- Cutscore Analysis (Validity)
  - Test Use (next section)





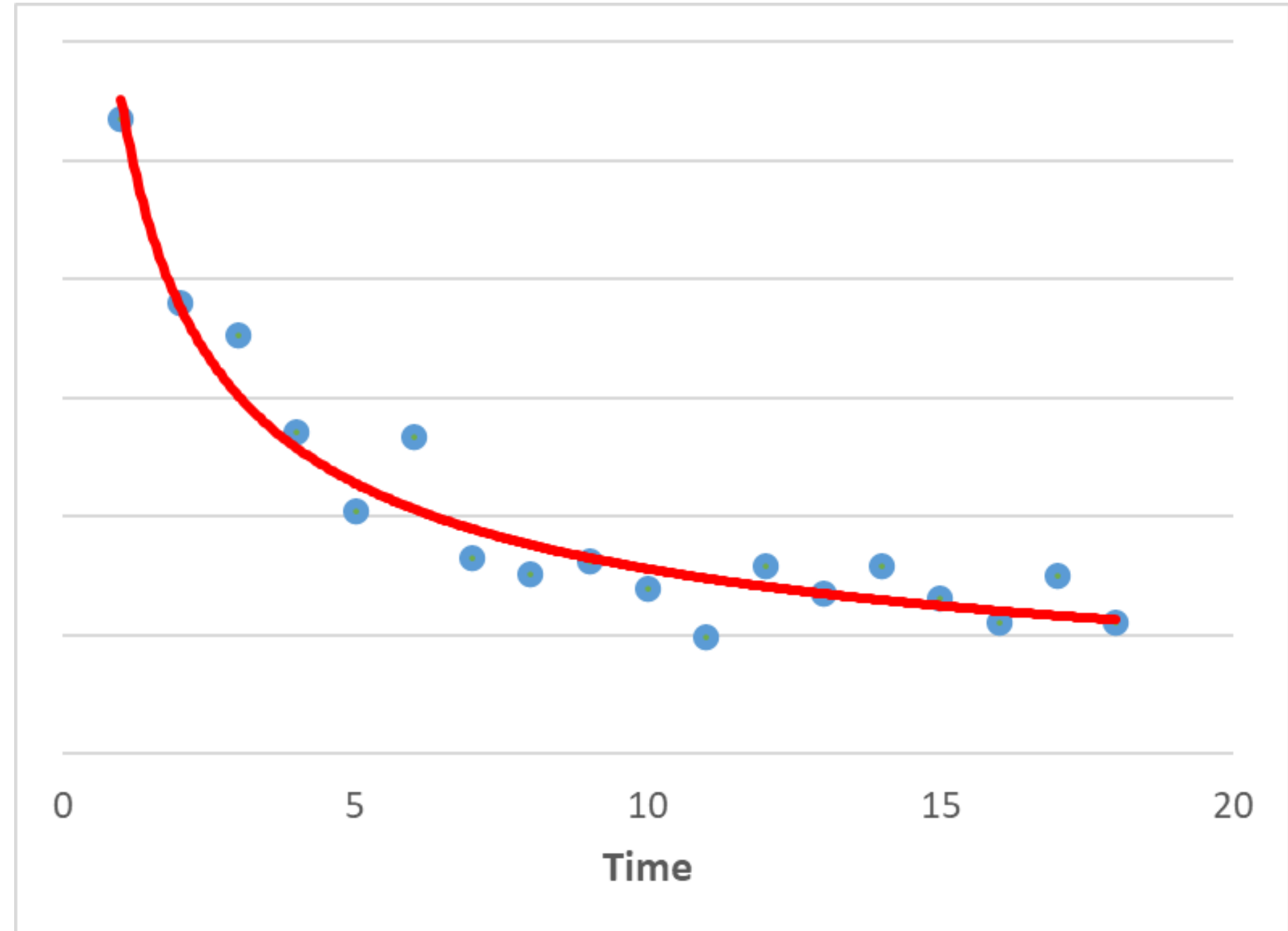
# Uniform Guidelines

- Guidelines 14.B.8
- Fairness Study (Cleary, 1968)



# AI Maintenance

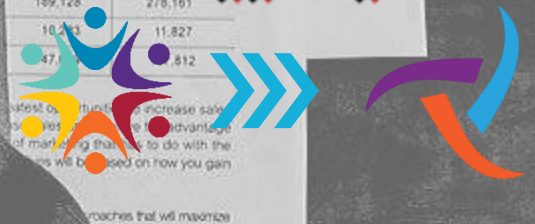
- Drift
  - Prediction Drift
  - Concept Drift
  - Data Drift
- Retraining





# Use Case

## How Will It Be Used?





# Using the AI Tool – Inferences

- How will candidates be scored?
- How is the score or classification determined, e.g. Cutscore?
- What is the ultimate purpose and inference drawn?
  - **Screen out:** remove unqualified applicants/candidates
    - Pass/fail
    - Qualified/not-qualified
  - **Select in:** identify your future highest job performers
    - Priority 1, Priority 2, Priority 3, etc.
    - Grade A, B, C, D
    - Single score



# Using the AI Tool – Validation Considerations

- **Screen Out**

- **Goal:** remove unqualified applicants/candidates
- Heavily dependent upon:
  - Valid minimum qualifications
  - Valid tool to identify minimally qualified vs not qualified
- **Key question:** How is the cutoff for qualified vs not qualified determined? Is it valid?

- **Select In**

- **Goal:** identify the "best" applicants (highest potential for strong performance)
- Heavily dependent upon:
  - Statistical prediction of performance
- **Key Question:** can you relate scores on the AI tool to job performance measures?



# Using the AI Tool – Validation Strategies

- **Validity Generalization**

- The vendor has demonstrated in many environments that the tool is valid. You assume that because it works everywhere else, it must be valid for your organization.

- **Validity Transportability**

- A different organization has demonstrated the tool is valid. You demonstrate that your jobs are the *same* as in the other organization.

- **Local Validation**

- Your organization has demonstrated the tool is valid in your organization.





# Conclusions



# Recap

- General Questions to Consider
- Quality of Training Data
- Analyzing & Evaluating the AI Tool
- Use Case—How Will it Be Used?



# Thank You!!!

## Questions/Comments, e-mail us:



**Daniel (dk) Kuang, Ph.D.**

Director

[dkuang@resecon.com](mailto:dkuang@resecon.com)

[www.Linkedin.com/in/dankuang/](http://www.Linkedin.com/in/dankuang/)

**Brian Marentette, Ph.D.**

Director of People Insights

[bmarentette@biddle.com](mailto:bmarentette@biddle.com)

[www.linkedin.com/in/brianmarentette](http://www.linkedin.com/in/brianmarentette)





# Compliance Analytics

- Fairness Study (Cleary, 1968)

