2020 ICE EXCHANGE

ICE Workshop
Basics of Psychometrics: Deciphering “It Depends”

Alpine Testing Solutions

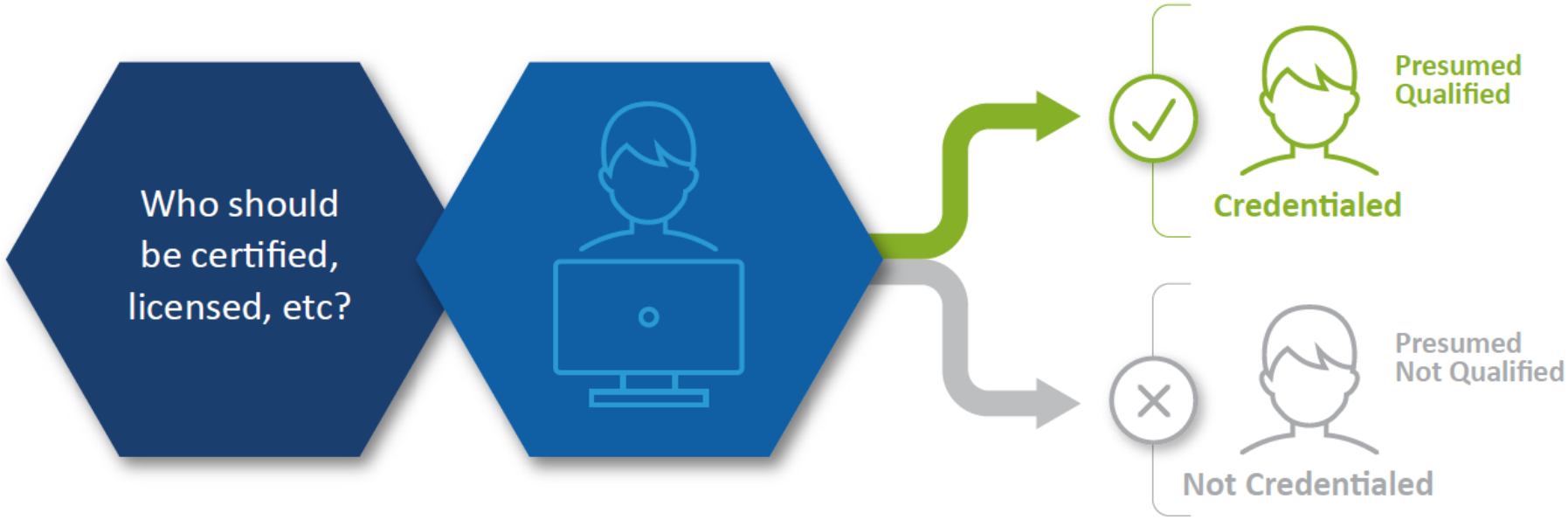
Brett P. Foley, Ph.D., Sr Psychometrician, Dir. Professional Cred.

Corina M. Owens, Ph.D., Psychometrician

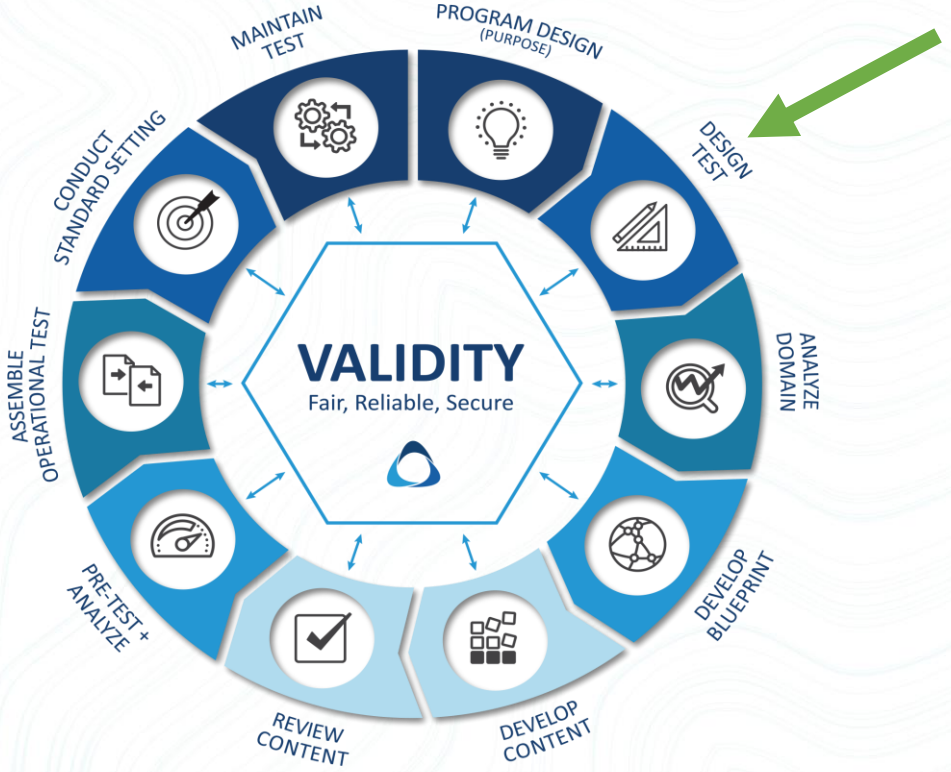
[#EXCHANGEWITHICE]

Validity

How do we know the decisions we are making about candidates are valid?



Test Development Process: Design Test

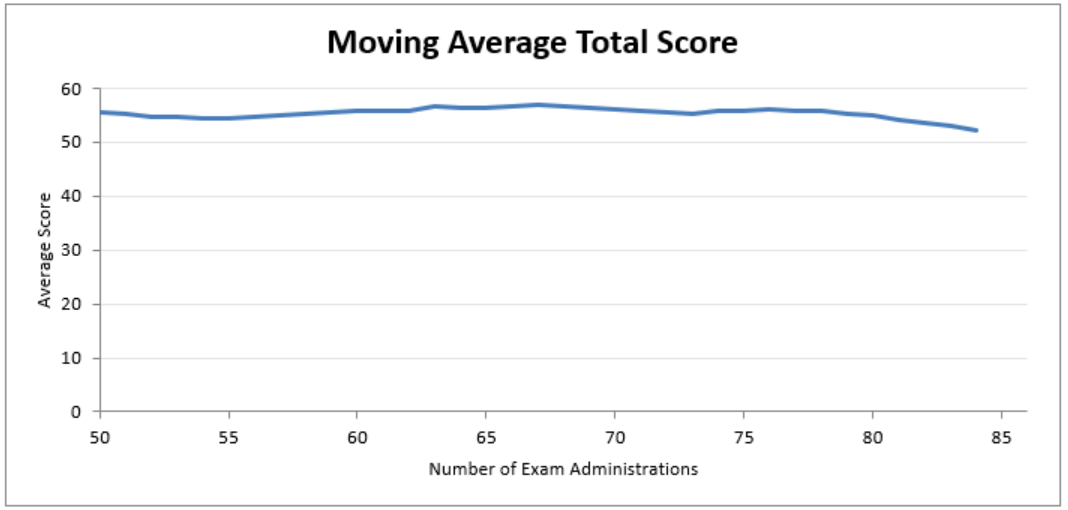


Test Design

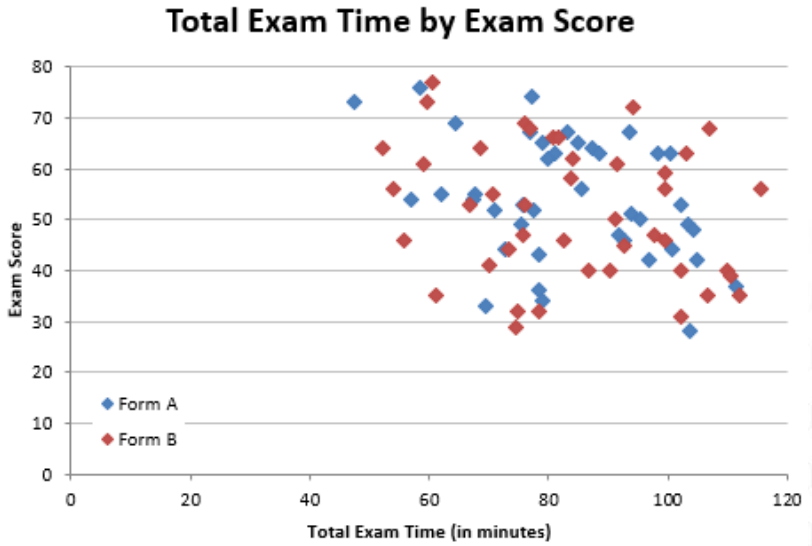
- Purpose should include information on:
  - Intended use of the test
  - Intended interpretation of the results
  - Potential misuses of the test/results
- Purpose should be informed by:
  - Intended value to stakeholders
  - Structure of the larger program
  - Requirements related to the test

Basic Analyses

- Exam/Item Performance Over Time



- Score/Time Combination



Domain Analysis

- Underlying knowledge, skills, and abilities (KSAs) of tasks/functions identified to:
  - Clarify enabling sub-tasks/sub-functions
  - Provide tools to downstream users (e.g., survey participants, item writers)
  - Inform translation of tasks into measurement objectives
- Might be referred to as:
  - Task specifications
  - Cognitive enablers
  - Descriptors

Development Activities in the Virtual Environment

- Can you do them and what does that look like?
  - You can!
  - We have successfully conducted all exam dev meetings virtually
  - How those look varies based on the needs of the clients
    - SME involvement
    - SME location
- Flexibility, flexibility, flexibility
- Tools are VERY important

Psychometric Security Analyses: Detecting Collusion & Pre-Knowledge

Security Analysis:	Brief Descriptions:	Focus:
Unscored-only Analysis	Health check analysis on unscored items only	Items
Moving Average Pass Rates	Snapshot of the average pass rate over time	Form
Differential Person Functioning (DPF)	Review of how a candidate performs on an exam – compare how the candidate did on unscored items vs. how they did on scored items	Person
Differential Item Functioning (DIF)	Review of how items perform on an exam – compare how items perform on identified groups of candidates	Items
Bivariate Score by Time	Comparison of score vs. time for a candidate	Person
Score Similarity	Comparison of candidates with similar scores/patterns	Person
with Clustering	Cluster candidates into different groups for further investigating	Person
Person Residual Correlation ("B3") – *Currently experimental, if interested, discuss w/Psych lead and Brett.*	Experimental - TBD	Person
z-Score Report	Comparison of demographic variables (e.g. test centers)	Test Center

Remote Proctoring: Top Areas of Concern

- Security
- Validity/Fairness/Consistent Experience/Equivalence
- Access/Capacity/Ease of Use
- Capacity for Alternate Item Types
- Reputation/Stakeholder Buy-in
- Cost
- Proctor Concerns
- Accreditation