



Transporting Validity Using the O*NET GWAs

Joseph Abraham

IPAC 2024



PSYCHOLOGY + TECHNOLOGY

We're not just a technology platform,
and we're not just psychology – we're **both**,
and with good reason.

Your presenter

Joseph D. Abraham, Ph.D.
Vice President,
Assessment Solutions



- Special thanks to Amanda Allen and Mark LoVerde of DCI Consulting* and Dawn Lambert from Talogy

Overview

- Transportability Concepts
- Review of Abraham et al. (2003)
 - *How similar is similar enough? Job profile similarity benchmarks using occupational information network data*
- Extending the method
 - Using O*NET GWA survey directly and specifically for transporting validity

Background*

■ Validity Transportability

- Validity evidence for the use of a selection procedure collected in one situation is applied to the use of the selection procedure in a new situation.
- A method of generalizing validity evidence across settings.
- Research findings are determined to be applicable based on similarities across the jobs and contexts

Professional Guidance*

- *Principles for the Validation and Use of Personnel Selection Procedures* (SIOP, 2018)
 - Transportability of validity evidence is based on “a preponderance of key observable and/or underlying similarities with other validity evidence”
 - Comparability is defined “in terms of job content or job requirements, job context, and job applicant group...(p. 19).”
- *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014)
 - Validity evidence may be transported if there is evidence that the new job is “highly comparable” (see Chapter 11, page 173)

Regulatory Guidance*

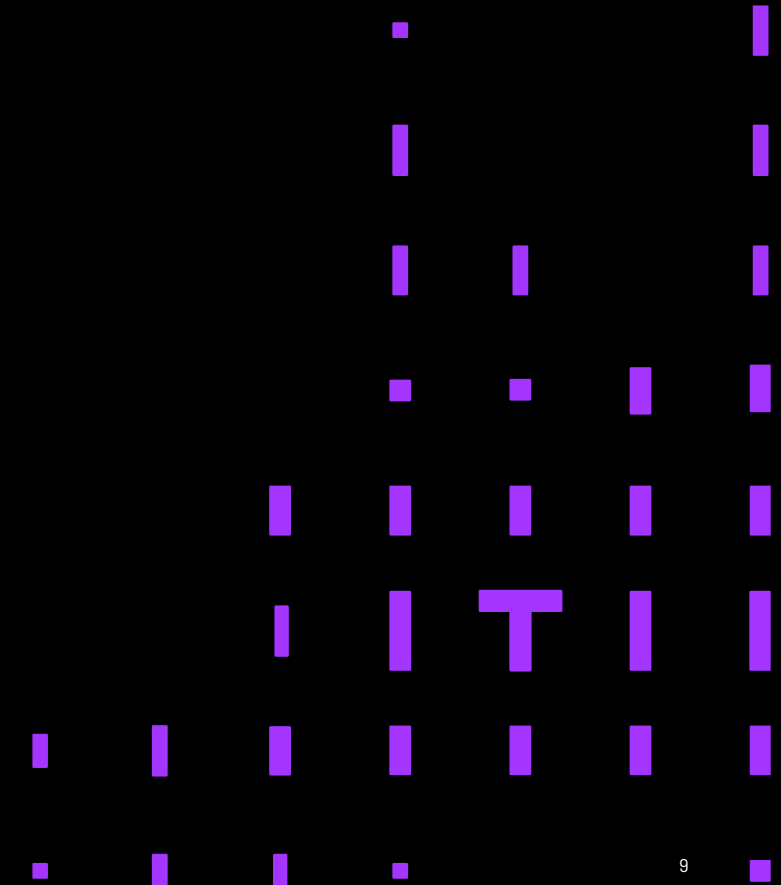
- *Uniform Guidelines on Employee Selection Procedures (EEOC, 1978)*
 - Allow for validity evidence for one job to be applied to another job
 - Establish sufficient similarity between the two jobs
 - Meets standards for criterion-related validity
 - Should consider available evidence of fairness/unfairness, other factors that might affect validity
 - Section 7B – Job similarity is established by showing incumbents *“perform substantially the same major work behaviors, as shown by appropriate job analyses both on the job or group of jobs on which the validity study was performed and on the job for which the selection procedure is to be used.”*
 - Groups of tasks? What about modern definitions of work behaviors? What about abilities or competencies? What is “substantially the same?”

Summarizing

- Professional guidance is broad
 - UGESP is more prescriptive, but open to future developments in the field
 - Safe to assume evidence of both job similarity (new and prior setting) and test validity (prior setting) are required
- But how to assess similarity?
 - Traditional way is customized job analysis task survey administered in both settings, with overlap statistics using thresholds
 - Can we leverage existing and accumulated research to help inform similarity judgments?

Job Similarity Metrics

- Thresholds/overlap most common
- Profile Similarity Metrics: An alternative or complementary approach with three elements
 - Elevation – On average, how do the jobs compare on the mean of profile elements? For example, is mean importance higher for one job than another?
 - Dispersion – On average, how do the jobs compare in the variance (or standard deviation) across profile elements? Are they both flat profiles? Both with peaks and valleys?
 - Shape – How strongly do the profiles correlate with one another (e.g., Pearson r)? Same pattern of profile elements? Important that this is an effect size metrics as well.



Abraham, et. al. (2023)¹ summary

- O*NET 25.1 provides rich database for job similarity benchmarks
 - Matrix of job profiles contains 380,628 job pairs
 - Developed correlation matrices across five attributes
 - Represents effect size reference tables
 - Generalized Work Activities (GWAs), Work Styles, Knowledge, Skills, Abilities
 - Tables can be paired with threshold-based comparisons
 - Adds information to critical overlap statistics
 - Can help inform transportability judgments

¹Abraham, J. D., Lambert, D. D., Mihalecz, M. C., Elcott, M. D., Asbury, H. S., & Palmer, P. C. (2023). How similar is similar enough? Job profile similarity benchmarks using occupational information network data. *International Journal of Selection and Assessment*, 31, 469–476. <https://doi.org/10.1111/ijsa.12430>

More on O*NET GWAs

- *A Generalized Work Activity is an aggregation of similar job activities/behaviors that underlies the accomplishment of major work functions. (O*NET technical manual)*
- *Work Activities summarize the kinds of tasks that may be performed across multiple occupations. (O*NET Online Help)*
- *A work activity is a set of similar actions that are performed together in many different jobs. (O*NET GWA Survey Instructions)*
- Can be linked to task categories if desired

Job Profile Correlation Distribution, GWAs

	GWAs
n	380,628
Mean	0.43
SD	0.28
Min	-0.56
Max	0.98

Percentile	GWA Profile r
0	-0.56
1 st	-0.20
5 th	-0.06
10 th	0.03
15 th	0.10
20 th	0.16
25 th	0.22
30 th	0.27
35 th	0.32
40 th	0.37
45 th	0.41
50 th	0.45

Percentile	GWA Profile r
55 th	0.50
60 th	0.54
65 th	0.58
70 th	0.62
75 th	0.65
80 th	0.69
85 th	0.73
90 th	0.78
95 th	0.83
99 th	0.90
100 th	0.98

Job Comparison Procedure

1. Common-sense review of similarity based on available data, experience, and expertise

Ex: Threshold-based comparison to examine overlap in critical attribute(s)

- may use Similarity Index (SI) formula, Gibson and Caplinger (2007)²

$$SI = \frac{N_C}{\sqrt{N_S * N_T}}$$

- where N_S and N_T = # of critical tasks for source and target jobs
- N_C = number of critical tasks common to both jobs
- $SI > .5$ indicates two jobs are more alike than different
- Suggest $SI \geq .75$ as initial overlap threshold

² Gibson, W. M., & Caplinger, J. A. (2007). Transportation of validation results. In S. M. McPhail (Ed.), *Alternative validation strategies: Developing new and leveraging existing validity evidence* (pp. 29–81). John Wiley & Sons Inc.

Job Comparison Procedure

2. Examine job profile correlations using O*NET benchmarks

Percentile Threshold	Suggested Interpretation
≥ 90 th percentile	Very good match
≥ 80 th percentile	Good match
≥ 70 th percentile	Moderate match
< 70 th percentile	Questionable match (additional investigation recommended)

A GWA Distance Matrix Alternative

- Correlation matrix represents effect sizes; good for benchmarking
- Distance matrix provides some advantages if using the O*NET GWA survey directly
 - Summarizes shape, elevation, dispersion in one measure
 - A similar approach can be used to assess job similarity
 - Threshold-based overlap statistic
 - GWA distance matrix of job profiles contains 380,628 job pairs
 - Can help inform transportability judgments
- Squared Euclidian Distance formula

$$D^2 = \sum_{i=1}^{82} (C_i - V_i)^2$$

- C_i = Importance and Level ratings on GWA i for the target job
- V_i = Importance and Level ratings on GWA i for the Talogy validation job

GWA Profile, Squared Euclidian Distance Distribution

	GWA Profile Distance
n	380,628
Mean	69,400
SD	53,600
Min	0
Max	371,900

Smaller distance value = stronger similarity

Percentile	GWA Profile Distance
0	371,900
5 th	182,600
10 th	155,600
15 th	135,100
20 th	118,500
25 th	104,600
30 th	93,100
35 th	83,500
40 th	75,400
45 th	68,600
50 th	62,500

Percentile	GWA Profile Distance
55 th	57,000
60 th	51,700
65 th	46,700
70 th	41,700
75 th	36,900
80 th	32,100
85 th	27,200
90 th	22,000
95 th	15,700
100 th	0

Job Comparison Procedure Revisited

1. Common-sense review of similarity based on available data, experience, and expertise

Ex: Threshold-based comparison to examine overlap in critical attribute(s)

- may use Similarity Index (SI) formula, Gibson and Caplinger (2007)²

$$SI = \frac{N_C}{\sqrt{N_S * N_T}}$$

- where N_S and N_T = # of critical tasks for source and target jobs
- N_C = number of critical tasks common to both jobs
- $SI > .5$ indicates two jobs are more alike than different
- Suggest $SI \geq .75$ as initial overlap threshold

² Gibson, W. M., & Caplinger, J. A. (2007). Transportation of validation results. In S. M. McPhail (Ed.), *Alternative validation strategies: Developing new and leveraging existing validity evidence* (pp. 29–81). John Wiley & Sons Inc.

Job Comparison Procedure Revisited

2. Examine job profile distance using O*NET GWA benchmarks

Percentile Threshold	Suggested Interpretation
$\geq 90^{\text{th}}$ percentile	Very good match
$\geq 80^{\text{th}}$ percentile	Good match
$\geq 70^{\text{th}}$ percentile	Moderate match
$< 70^{\text{th}}$ percentile	Questionable match (additional investigation recommended)

An example in practice

- Collected GWA profile data for a **Maintenance** position that involves preventive and corrective maintenance of electrical and mechanical machinery
- N = 30 GWA questionnaire responses received from SMEs
- Overall reliability = .91 across respondents

GWA Questionnaire

- O*NET questionnaire asking about 41 GWAs (e.g., Getting Information, Working with Computers, Handling and Moving Objects)
- Respondents rate **Importance** and **Level** of each GWA
- GWA profile therefore contains 82 ratings, aggregated across respondents

Example Rating						
Inspecting Equipment, Structures, or Materials: Inspecting equipment, structures, or materials to identify the cause of errors or other problems or defects						
How important is the activity to the XYZ job?						
1	2	3	4	5		
Not Important	Somewhat Important	Important	Very Important	Extremely Important		
What level of the activity is needed to perform the XYZ job?						
1	2	3	4	5	6	7
Checks that doors to a building are locked			Inspect equipment in a chemical processing plant		Conduct a thorough inspection of an aircraft engine	

Job Comparison Process in Practice: Example 1

- Example of job we expected to be similar to Maintenance role under study: **Maintenance Specialist**
- We have existing validity evidence for such roles in our database
- Validity evidence is for an assessment that:
 - Includes both cognitive and non-cognitive components
 - Incorporates behavioral, personal beliefs, situational judgment, and simulation item types
 - Measures numerous competencies such as Attention to Detail, Accountability, Teamwork, Problem-Solving, and Work Tempo

Job Comparison Process in Practice: Example 1

- Comparison Step 1:
 - Threshold-based comparison: Similarity Index = .93
 - This value is $\geq .75$; meets overlap threshold
- Comparison Step 2:
 - GWA distance measure = 4,349
 - Corresponds to 99th percentile
- Interpretation: VERY GOOD MATCH
- Conclusion: Validity evidence can be transported from the Maintenance Specialist job to the target job

Job Comparison Process in Practice: Example 2

- Example of job we expected to be dissimilar to Maintenance role:
Quality Assurance Inspector
- We have existing validity evidence for the Quality Assurance Inspector role in our database, for the same assessment

Job Comparison Process in Practice: Example 2

- Comparison Step 1:
 - Threshold-based comparison: Similarity Index = .58
 - This value is < .75; below overlap benchmark
- Comparison Step 2:
 - GWA distance measure = 59,948
 - Corresponds to 53rd percentile
- Interpretation: QUESTIONABLE MATCH
- Conclusion: Validity evidence should NOT be transported from the Quality Assurance job to the target job

Job Comparison Summary

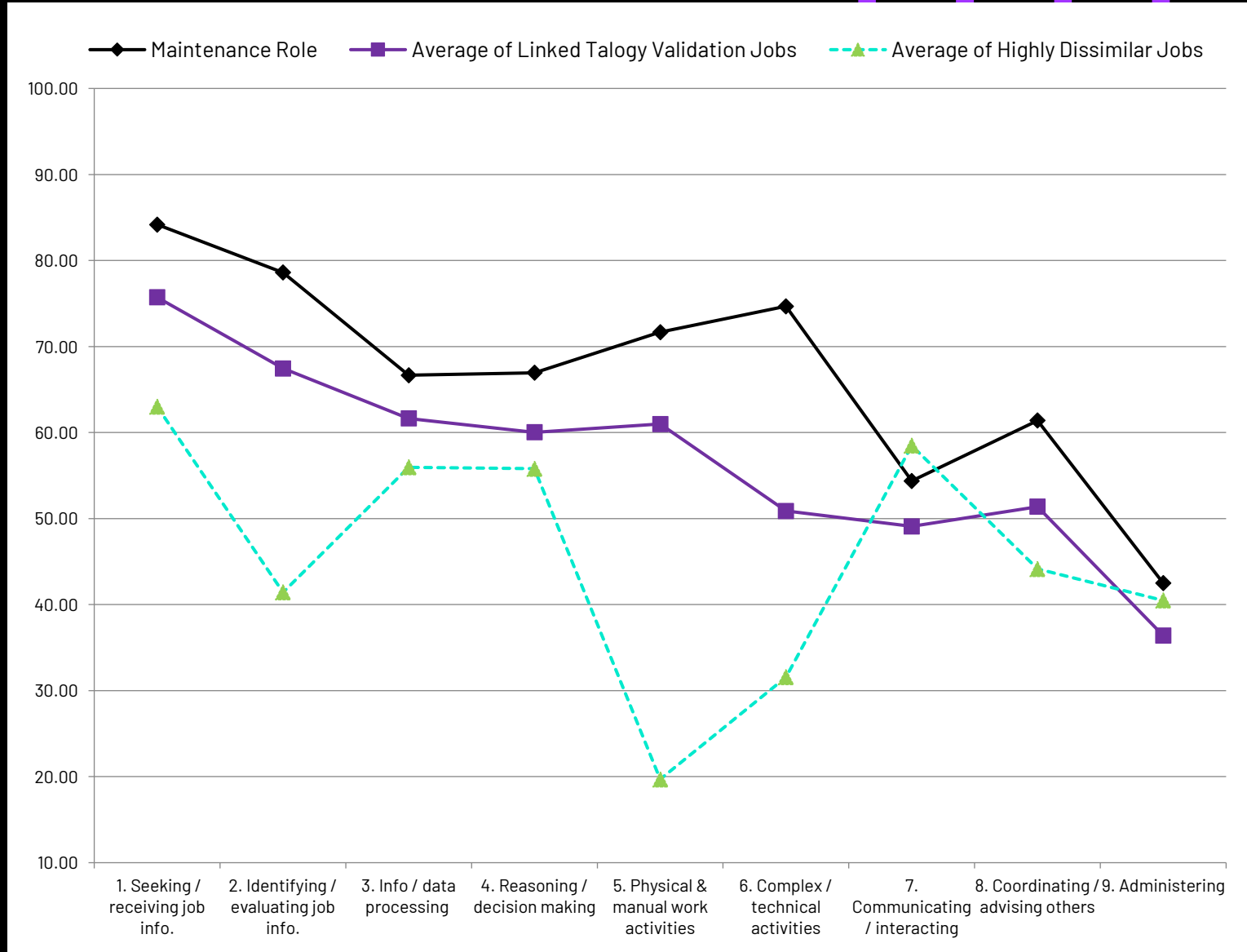
Maintenance Role (linked to 20 Talogy jobs with validation data)

Some of the jobs in Talogy's validation database found to be **most similar** to the Maintenance role include...

- Maintenance and Repair Workers
- Maintenance Specialists
- Electrical and Electronics Repairers
- Electric Motor, Power Tool, and Related Repairers
- Multiple Machine Tool Setters, Operators, and Tenders
- Service Technicians and Mechanics

Some of the jobs found to be **most dissimilar** include...

- Telemarketers
- Chief Executives
- Human Resources Managers
- Sales Representatives



Aggregating Validity Evidence

- Can aggregate across multiple “linked” jobs with validation data
- Aggregated mean (sample-weighted) validity in relation to job performance*

Competency Measured	# of linked studies	Total N	Uncorrected r with Job Performance	Corrected r with Job Performance
Accountability	38	5,904	.19	.26
Attention to Detail	38	5,960	.17	.24
Dependability	30	4,304	.17	.23
Positive Attitude	38	6,037	.13	.18
Problem-Solving	24	5,453	.18	.25
Process Monitoring	37	5,904	.15	.21
Safety Orientation	36	5,791	.17	.24
Teamwork	36	5,854	.14	.19
Work Ethic	36	5,567	.17	.23
Wok Tempo	35	5,721	.18	.25
Overall Composite Score	34	4,822	.29	.37

*Can also compute meta-analytic estimate across linked jobs, including credibility intervals

Summary

- In a previous article (Abraham et. al., 2023), we discuss how the O*NET correlation matrices for various job attributes can help in assessing job similarity
- The O*NET Generalized Work Activities (GWAs) are particularly helpful when investigating transportability
- If the researcher is using the O*NET GWA survey directly, the squared Euclidian distance matrix among O*NET GWA profiles can be applied as a frame of reference when assessing similarity
 - Includes information about job profile elevation, dispersion AND shape
 - Pairs well with threshold-based comparisons
- Aggregation of criterion-related validation evidence across linked studies may be possible



THANKS FOR
YOUR TIME

Learn more at [Talogy.com](https://talogy.com)

