Moving Beyond DOT and O*NET:

How Do We Solve the Challenge of Linking Work Activities and Worker-Trait Requirements?

RJ Harvey
Virginia Tech; PSTC

Today's Goals

- 1. Review the sad history of O*NET
- 2. Discuss some other disturbing trends and widely-held "truths"
- 3. Propose a strategy for dealing with the loss of DOT, inadequacies of O*NET
  - How do we collect accurate JA data?
  - How do we link it to worker-trait specifications?

---

1. The Sad History of O*NET

- Since 1939, DOT served many users well
- Definitive listing of occupational titles
- At least some task-based content on each
- Common-metric of general work activities (GWAs) – aka "work dimensions" – based on FJA's Data-People-Things taxonomy of worker functions
- Plus some rationally rated worker-trait requirements (the "green monster")
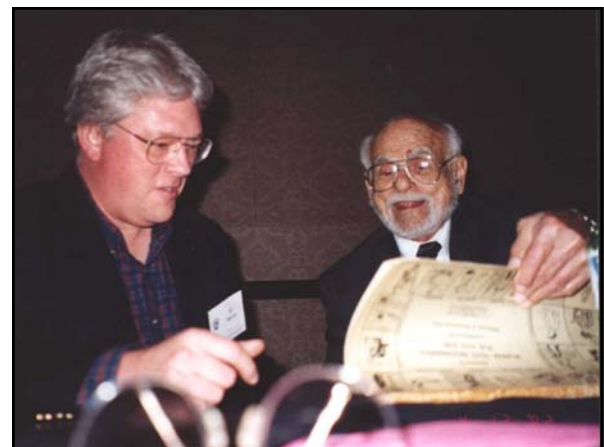
1990's: Developing O*NET

- Rather than revise it again, DOL decided to terminate DOT, with extreme prejudice
- O*NET Goal: do everything DOT did, plus much much more…
- Infamous APDOT "content model" designed to include everything imaginable regarding an occupation
- Who's Who (e.g., Fleishman, Jeanneret, Campion, Borman, Pearlman, Campbell, Peterson) enlisted to provide the intellectual guidance and direction
- No question its goals included serving as a source of information on employee competency requirements, selection information

---

- Hubbard, McCloy, Campbell, Nottingham, Lewis, Rivkin, & Levine (2000):
  - "O*NET will be the most comprehensive standard source of occupational information in the United States. O*NET will be at the center of an extensive network of occupational information used by a wide range of audiences, from individuals making career decisions, to public agencies and schools making training investment decisions, to employers making job structure and hiring decisions. O*NET will also be widely used for administration of federal programs" (p. v).

The Result: A Bridge to Nowhere

- Tens of millions of $$s and a decade invested
- We get a system based almost entirely on single-item holistic ratings of highly abstract traits emerged
- 4 short questionnaires (Abilities, Skills, Knowledge, GWAs)
- Many years of withering criticism from researchers and practitioners ensued
- Former DOT users (FDUs) were particularly upset, especially folks working in voc rehab, disability determination (e.g., SSA)

What's Wrong with O*NET?

- O*NET title taxonomy has only 5-10% of DOT's titles (13,000+ to 700-1200): WAY too abstract
- Construct validity: can't tell difference between *work* versus *worker-trait* constructs
- No moderate- or high-specificity data
- No mod-spec survey that provides a common metric for comparing jobs (e.g., PAQ, CMQ level)
- No "crosswalk" to worker traits rated by DOT
- Future DB updates come from small samples of untrained, volunteer, unaccountable incumbents
- No evidence of *validity* of worker-trait inferences
- No evidence of *accuracy* of JA ratings
- Terrible interrater agreement, even using trained analysts

O*NET interrater *r*s
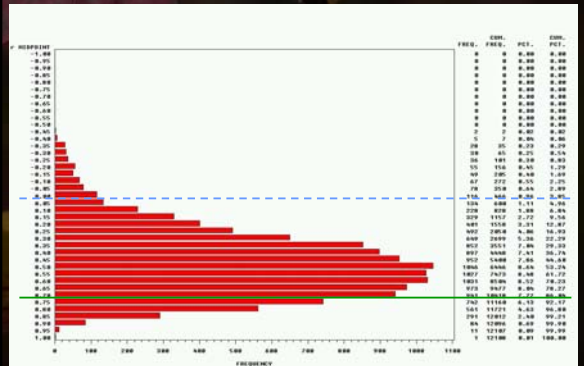
Current Status:

- After many years of trying to discredit its critics, DOL finally did a 180, admitted that O*NET is so seriously flawed it can't be used for selection
- DOL backed off in a big way, now saying it's only good for "career exploration"
- Jim Woods @ DOL: "don't you dare use it for selection or any litigious application!

Can I use the O*NET Ability Profiler for selection?

No, the O*NET Ability Profiler was designed specifically for career exploration, career counseling, and career planning purposes. Validity studies have not been conducted for using the O*NET Ability Profiler for selection purposes. The O*NET Ability Profiler must not be used for personnel selection.

A number of sources are available to help organizations locate suitable assessments and explore the reliability and validity of various assessments for specific uses. Some of the available sources include:

This is progress?

- DOL has totally dropped the ball
- There is now no government-provided, economy-wide source of accurate, current, defensible JOA data
- Or, perhaps more important, a way of collecting it in local situations
- Unfortunately, CRA, ADA, *Griggs*, *Uniform Guidelines*, etc., are all still there
- We're arguably worse-off than back in 1939 before the first edition of DOT
- How far have we really come since Hawthorne days?
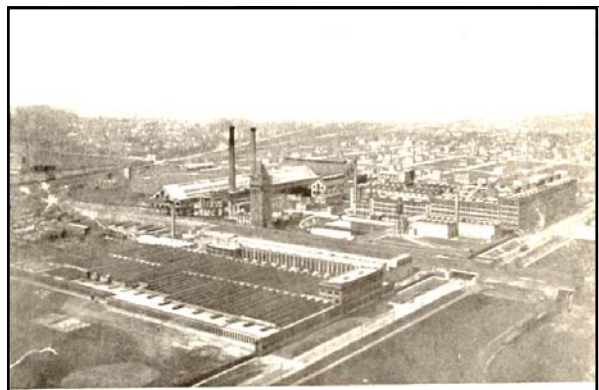
HAWTHORNE WORKS, FROM AN AEROPLANE

Root problem # 1: cheap & easy = "job 1"

- O*NET fatally flawed from the drawing board due to insistence it be dramatically less costly than DOT
- Peterson et al. (2001):
  - "The O*NET provides a highly usable and inexpensive methodology for analyzing jobs. The structured self-report questionnaire format of the O*NET's rating scales is much easier to use than the analyst-based and largely narrative format of the DOT. In addition, it will be readily available for public and private sector use through information technology (e.g., Internet). This suggests that the O*NET will have a great impact on research and practice. It is certain to provide many years of good service to the public, just as the DOT did" (p. 487)

Root problem # 2: holistic ratings

- To achieve the drastic reductions in cost demanded by DOL, many corners had to be cut
- No more OAFCs with staffs of occupational analysts to collect data
- No more task-level data
- No independent review, verification
- "Solution" = make single-item *holistic* ratings of *highly* abstract work-activity, worker-trait *constructs*

Result

- A database composed almost entirely of unverified, *unverifiable* ratings
- Normally, to infer scores on a latent *construct*, we combine scores on a large number of specific, observable data points
- In O*NET, we *start* the rating process by rating *constructs*
- Many former DOT users disgusted, refuse to use O*NET to replace DOT
- The nightmare of going to court and having to defend the O*NET ratings….

Verifiably accurate?

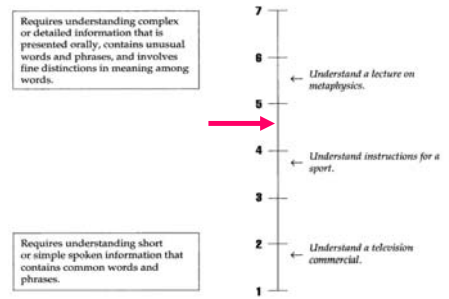- How do you defend accuracy (for JA) or validity (for JS) when…
- Items are up in the clouds in terms of abstraction
- Definitions of traits are often obtuse
- Rating process uses single-item judgment of job as a whole
- Anchors of rating scale almost guaranteed *not* to involve actual behaviors performed on job in question
- Many rating points lack any anchor at all
- Extreme anchors are kind of odd…

Let's audit the O*NET

- We'll rate a job familiar to us all – Industrial/Organizational psychologist
- Oddly, given that hundreds of DOT titles were combined to form many O*NET occupations, many separate titles for "Psychologist" exist…
- You determine the "correct" rating of each ability trait for I/O psychologist

1. **Oral Comprehension** — The ability to listen to and understand information and ideas presented through spoken words and sentences.

Requires understanding complex or detailed information that is presented orally, contains unusual words and phrases, and involves fine distinctions in meaning among words.

7

6 ← Understand a lecture on metaphysics.

5

4 ← Understand instructions for a sport.

3

2 ← Understand a television commercial.

1

Requires understanding short or simple spoken information that contains common words and phrases.

2. Written Comprehension
The ability to read and understand information and ideas presented in writing.

Understand signs on the highway		Understand an apartment lease		Understand an instruction book on repairing missile guidance systems		
①	②	③	④	⑤	⑥	⑦

Highest Level

3. Oral Expression
The ability to communicate information and ideas in speaking so others will understand.

Cancel newspaper delivery by phone		Give instructions to a lost motorist		Explain advanced principles of genetics to college freshmen		
①	②	③	④	⑤	⑥	⑦

Highest Level

4. Written Expression
The ability to communicate information and ideas in writing so others will understand.

Write a note to remind someone to take food out of the freezer		Write a job recommendation for a subordinate		Write an advanced economics textbook		
①	②	③	④	⑤	⑥	⑦

Highest Level

7. Problem Sensitivity
The ability to tell when something is wrong or is likely to go wrong. It does not involve solving the problem, only recognizing that there is a problem.

A. What level of PROBLEM SENSITIVITY is needed to perform your current job?

Recognize that an unplugged lamp won't work		Recognize from the mood of prisoners that a prison riot is likely to occur		Recognize an illness at an early stage of a disease when there are only a few symptoms		
①	②	③	④	⑤	⑥	⑦

Highest Level

8. Deductive Reasoning
The ability to apply general rules to specific problems to produce answers that make sense.

Know that a stalled car can coast downhill		Decide what factors to consider in selecting stocks		Design an aircraft wing using principles of aerodynamics		
①	②	③	④	⑤	⑥	⑦

Highest Level

9. Inductive Reasoning
The ability to combine pieces of information to form general rules or conclusions (includes finding a relationship among seemingly unrelated events).

Requires the development of the best rule or theory to explain how very different things or groups of things are related. The things may be related in several ways.

7
← Develop the theory of relativity.
6
5
4
← Interpret a weather chart.
3
2
← Order a seafood platter at a restaurant to determine whether or not you like seafood.
1

Requires the development of a simple basic rule to explain how two similar things are related.

4

12. Mathematical Reasoning

The ability to choose the right mathematical methods or formulas to solve a problem.

Determine how much 10 oranges will cost when they are priced at 2 for 29 cents

Decide how to calculate profits to determine the amounts of yearly bonuses

Determine the mathematics required to simulate a space craft landing on the moon

① —— ② —— ③ —— ④ —— ⑤ —— ⑥ —— ⑦

Highest Level

23. Manual Dexterity

The ability to quickly move your hand, your hand together with your arm, or your two hands to grasp, manipulate, or assemble objects.

Requires very fast, skillful, coordinated use of one hand, a hand and arms, or two hands to grasp, place, move, or assemble objects.

7

6

5

4 ← Package oranges in crates as quickly as possible.

3

← Tie a necktie.

← Perform open-heart surgery.

Requires some speed, skill, and coordination to grasp, place, move, or assemble objects with one hand, a hand and arm, or two hands.

2

1

27. Response Orientation

The ability to choose quickly between *two or more movements* in response to *two or more different signals* (lights, sounds, pictures). It includes the speed with which the correct response is *started* with the hand, foot, or other body part.

Requires the extremely rapid selection of the correct movement when there are many signals and many response possibilities.

7 ← In a spacecraft which is out of control, quickly react to each malfunction with the appropriate control movements.

6

5 ← Operate a busy switchboard where you have to plug calls in and out quickly and accurately every few seconds.

4

3 ← When the doorbell and telephone both ring at the same time, quickly select which one you will answer first.

Requires some speed in selection of the correct movement when there are two signals and two response possibilities.

2

1

35. Trunk Strength

The ability to use your abdominal and lower back muscles to support part of the body repeatedly or continuously over time without "giving out" or fatiguing.

Sit up in an office chair

Shovel snow for half an hour

Do 100 sit-ups

① —— ② —— ③ —— ④ —— ⑤ —— ⑥ —— ⑦

Highest Level

41. Near Vision

The ability to see details at close range (within a few feet of the observer).

Read dials on the dashboard of a car

Read the fine print of a legal document

Detect minor defects in a diamond

① —— ② —— ③ —— ④ —— ⑤ —— ⑥ —— ⑦

Highest Level

42. Far Vision

The ability to see details at a distance.

Read a roadside billboard

Focus a slide projector

Detect differences in ships on the horizon

① —— ② —— ③ —— ④ —— ⑤ —— ⑥ —— ⑦

Highest Level

Curiously…

- Millions are still being given to contractors (e.g., Aguirre Intl) for O*NET work
- Still going strong, according to its site

"outreach"?

---

2. Other Questionable "Truths"

- Many of our conceptual leaders in I/O
  – designed O*NET
  – remain adamant backers
- A number of other "truths" seem to be widely held in some quarters
- I find them disturbing, especially when you put them all together

---

Top-10 scary "truths"

10. Interrater agreement = reliability (Fleishman & Mumford, 1991)

9. Reliability > .70 = OK (Hunter: anything > 0 = OK)

8. Reliability = validity (Fleishman & Mumford, 1991)

7. JA = rate worker-trait requirements ("personality-oriented JA")

6. Validation standards for holistic competency ratings are much less stringent than for traditional tests

5. Anonymous incumbents, people with no direct job familiarity, are good sources of JA & competency ratings

4. JA ratings aren't faked; w/i title disagreement = noise

3. McCormick invented "worker-oriented" JA; PAQ measures it

2. Personality test faking is irrelevant; nothing moderates validity

1. We can't assess JA accuracy because there is no reality

Ontology

- The metaphysical study of the nature of being and existence
- Q: You're making this all up – does anybody really maintain that there is no "reality" in job analysis and test validation?
- A: Yes, some very influential ones…

Sanchez & Levine (2000)

- "a basic assumption of any attempt to assess JA accuracy is that there is some underlying "gold standard" or unquestionably correct depiction of the job. This assumption is problematic at best, for any depiction … is of necessity a social construction"
- "the concept of accuracy, when defined in terms of proximity to a known standard, has no legitimate meaning in psychological measurement"
- "Even if we assume for the sake of argument that there is an underlying reality in the way positions are clustered into a grouping we refer to as a job title, such reality is still open to various interpretations"
- "Conventional wisdom dictates that disagreement between two judges indicates that at least one of them must be wrong … However, as the French thinker Pascal … said, "there are truths on this side of the Pyrenees that are falsehoods on the other." In JA, just like between observers sitting on opposite sides of the Pyrenees, accuracy may be relative not absolute"
- "errors in classical reliability theory do not need to be mistakes"
- "Researchers sometimes fail to distinguish between measurement errors and mistakes, hence assuming that disagreement is a sure sign of mistaken judgment in at least one of the parties. Instead, disagreement may simply indicate systematic depictions of alternate but equally valid views. … disagreement does not always represent inaccuracy "

Schmidt, Hunter, Pearlman (1981)

- "the field of personnel psychology came to be so far off base" because during the "late 1950s and early 1960s ... behaviorist influences began to make themselves felt" (pp. 178-179).
- "There are two central claims to the modern behavioristic beliefs as they are manifested in the field of personnel psychology: that [human] abilities are not observable and that behavior is observable. Both claims are false to fact. … Consider the supposed observability of behavior. Suppose that a worker is to screw a certain bolt into a certain hole in each automobile as it passes on the assembly line. Is 'screwing in the bolt' an observed behavior in the worker? Certainly not." (p. 181).
- physical aspects of the work environment "are no more observable than [human ability] traits. Furthermore, physiology has established similar facts about the [work] response or behavior. A motor act such as screwing in a bolt is known to be a highly complex pattern of time-sequenced patterns of neural impulses to thousands of muscle fibers, to postural muscles, to eye muscles, and so forth. Thus, it is well known that no response ever repeats itself either. Thus, any equivalence of successive acts must be an internal perceptual process carried out by the brain. Therefore, responses [i.e., worker behaviors] too are events in the observing psychologist's mind and hence not directly observable. ... Thus the response (i.e., the behavior) of behaviorism is no more observable than is an ability; both are hypothetical constructs in the minds of those who use them as theoretical devices." (p. 181).

Why maintain that there is no reality?

- Takes us off the hook for even bothering to try to see if JA ratings are accurate
- Convenient explanation for bizarre incumbent ratings, poor interrater agreement
  - "relax, the raters are just describing alternative but equally-valid realities…"
- Lets us "prove" accuracy by working backward:
  - If the personnel function we developed seems to work…
  - then that proves the JA data driving it is accurate!
- Lets us blur the distinction between "accuracy" of JA ratings versus "validity" of worker-trait inferences

Does your job require you to…

- "Spray enamel or lacquer on automobile, using knowledge of car painting techniques, to build up thickness of paint specified in separate applications"
- "Perform manual operations to sever jugular vein of animals or poultry"
- "Use long-handled tools (rakes, hoes, shovels)"
- "Make managerial decisions to approve/deny the purchase or sale of subsidiary corporations"
- "Use firearms"
- → If you're doing behavior-oriented JA with properly written items, the answer has to be 'yes'

3. A Way Forward

- My explanation for O*NET debacle:
  - "Leaders" in I/O-HRM lost sight of history of our field
  - Forget the dumb ideas of the past, you'll repeat them
  - "Ivory-tower envy"
  - VG's antipathy for behavior-oriented JA
- Solution is to get back to basics
  - Stop trying to blur distinction between JA vs. JS
  - Do verifiable, behavior-based job analysis
  - Use empirical means of linking JA to JS inferences
- Only thing dramatically new is use of web-based infrastructure to collect, deliver it

Defensible JA essential

- Challenge 1: right balance between detail/ verifiability vs. length
  - Holistic single-item scales will never work
  - How short can we make it w/o sacrificing convergence with full-length assessment?
- Challenge 2: better ways of ensuring quality/ accuracy of ratings
  - IRT may yet be helpful as a first-cut
  - Eyeballs-on review by SMEs may be unavoidable (grounds-crew laborer using shotguns…)
- Challenge 3: what IS "worker-oriented" analysis, and how does it fit in replacing DOT/O*NET?

What is "worker oriented" JA?

- For over 30 years, "worker oriented" JA =
  - Standardized survey applicable to many jobs
  - Moderate-specificity items ("make decisions", "use long-handled tools", "operate highway vehicles")
  - Items rated on typical JA rating scale(s)
- "Job analysis" portion of O*NET (GWA survey) continues this pattern
- Basically, a simple matrix:
  - Rows = GWAs
 - Columns = same scales from a task inventory

But…

- How does that conceptually differ from a task-inventory approach?
  - Just using somewhat more abstract items
- How does that address W/O goal of describing the abstract *worker-functions* involved in performing the work?

Answer – it doesn't

- Although he's briefly cited in McCormick et al (1972) the real originator of structured, "worker oriented" JA is Sidney Fine
- Fine's FJA theory predated the Purdue work on structured JA, forming the conceptual foundation for original *DOT*
- The BIG THREE: FJA articulated a theoretical framework for the general structure of work that is theoretically elegant and supported by data
  - Data, People, Things as the general conceptual dimensions of work
  - Harvey (1987): factor analysis of PAQ → DPT higher-order factors
  - Harvey (2004): factor analysis of 1,222 CMQ items for 6,734 positions → DPT higher-order factors

FJA dates to 1930's

- The core of FJA is the goal of offering "a conceptual system defining *dimensions of work activity* and thus a way of conceiving the world of work" (Fine & Wiley,1971, p. 77)
- "what workers do, insofar as their job content is concerned, *they do in relation to three primitives: Things, Data, and People*. In relation to each primitive, workers function in unique ways. Thus in relation to Things, workers draw on physical resources; in relation to Data, on mental resources; and in relation to People, on interpersonal resources. *All jobs require the worker to relate to each of these primitives in some degree*. Although the behavior of workers or the tasks performed by them can apparently be described in an infinite number of ways, *there are only a small number of definitive functions* involved." (p. 78, emphasis added).
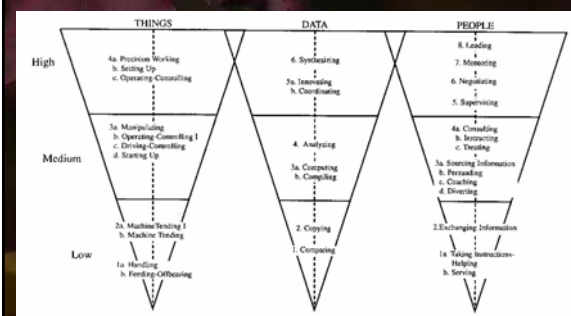
Unfortunately….

- The "victors" write the history books
- Jeanneret et al. (1999) GWA chapter in O*NET textbook – published by APA – does not cite Fine's work *even once*
- Indeed, it turns out McCormick was the one who supposedly invented the idea of measuring work using a common array of abstract work dimensions…

Jeanneret et al. (1999)

- "*Ernest J. McCormick made one of the most important contributions to job analysis research when he observed that descriptors of job content can be classified as either job-oriented or worker-oriented* (McCormick, 1979). This distinction is now especially important when considering how job analysis information will be used to document the activities of tomorrow's workforce."
- *"McCormick's vision of worker-oriented or behaviorally based job descriptors* was a very viable solution to the problems inherent in the current DOT and other systems that rely on task-based information as primary descriptors. (p. 105, emphasis added)."
- Jeanneret et al. conveniently forgot that Fine developed the same concepts 30 years before PAQ appeared

Adding insult to injury

- Jeanneret et al. (1999) did EFA of O*NET GWAs
- Using it, they claimed to have developed a novel integrative taxonomy of general work dimensions that "summarizes well the GWA domain" (Jeanneret et al., 1999, p. 125)
- Its 3 dimensions:
  - *Working With Information*,
  - *Working With and Directing the Activities of Others,*
  - *Manual and Physical Activities*
- Obviously, Data, People, and Things

Moving Forward

- Problem: how do we measure general work dimensions in a national JOA database?
- PAQ, O*NET rate things that are way too abstract to be verified
  - 1-item scale of Decision Making, Responsibility
- FJA can rate DPT for tasks, but how do you combine to get overall score for job?
- Original item ratings must be specific enough to be verified independently
- Answer: CMQ, other surveys that rate verifiable behavior, use *decomposed judgment* strategy to estimate the abstract GWA traits

CMQ Content model
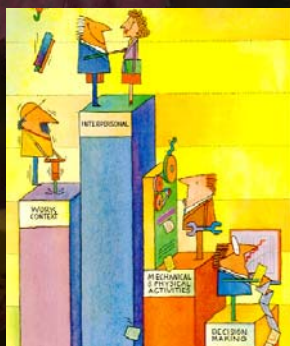
Four categories of work:

- **Work Context,** such as risks, autonomy, schedules
- **Interpersonal,** such as employee supervision and sales
- **Mechanical & Physical,** including machine, equipment, and tool use
- **Decision-Making,** including knowledge and business planning
- Owes an obvious debt to Fine's Functional Job Analysis theory "Big Three" of Data-People-Things

Data

9

People

Things

How to link Dunnette's "two worlds"?
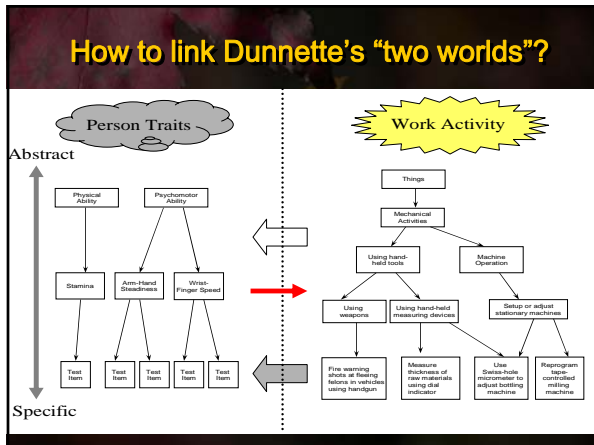
Linking JA to JS

- Biggest practical problem for practitioners
- We have many options
- Rational/non-empirical
  - SME linking panels
  - direct holistic judgment
  - Problem = very weak job relatedness
- Empirical
  - Criterion-related: local validation
  - Criterion-related: VG

Can't we just use VG??

- VG is an appealing option for large-scale OA
- Unfortunately, it's fatally flawed
- Typical uses go far beyond what Uniform Guidelines allow under validity transportability
- *Aggregation bias* as much a problem as for O*NET's title system (e.g., families = "clerical", "managerial")
- Based on questionable assumption that *sign* of correlation of predictor-criterion validity is always positive
  - At least for personality predictors, for some jobs we want the opposite pole
- Also based on highly questionable assumption that top-down is always best way to hire on a trait
  - No level-based prediction of cutoff scores possible

JCV to the rescue

- Job component validity (JCV) popularized by McCormick
  - Stunning levels of predictability using work dimensions to predict worker-trait requirements
  - Just do an accurate JA, apply equations, and out come the level-based worker-trait reqs
- Basically, policy-capturing of KSAOs given GWAs
- Very strong multiple Rs seen predicting worker-trait requirements (from GATB to MBTI) from PAQ
- Similar strong Rs seen for CMQ
  - Brown & Harvey (1996): 4 MBTI types
  - Wagner & Harvey (2004): 52 O*NET/ Fleishman ability traits
  - Wagner & Harvey (2005): DOT worker-function scales

JCV

Occupations

CMQ JA Ratings

derive

CMQ Dims

MLR

Worker KS & AO-Trait Scores

R_{Policy}

PREDICTED KS & AO Scores

$r_{current}$

KS-AO Scores in New Occs
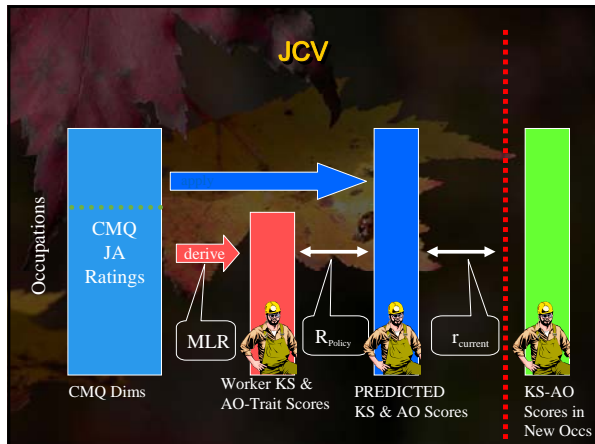
JCV advantages

- Based on defensible JA
- Get specific predictions of cut scores for each worker-trait
- Easily handle new, or substantially changed, or heterogeneous occupations
- Once robust equations developed, economy-wide system is highly practical, efficient
- BIGGEST PLUS: you can do ANY new job, even one not in database, and simply apply the JCV equations to predict its KS/AO requirements
  - No need to maintain an up-to-the minute national OA database covering every job

---

Now what?

- O*NET is dead: after tens of millions of $$s spent, DOL has given up on O*NET's original goals
- We're now in pre-1939 days
- SSA is going to have meltdown soon
- Other FDUs are in similar mess
- What's going to replace DOT?

Private sector to the rescue

- We've got the JA technology online now to collect verifiable JA data
- Web-based infrastructure is cheap, scalable
- We have sound taxonomies of dimensions of general work activity up to Big Three
- Missing link 1:
  - JCV studies to provide the equations to link the JA domain with the JS-trait domain
- Missing link 2:
  - Somebody to pull it all together

Doing what O*NET *should* have been

- Freely-available, web-based JA surveys for data collection
- Rate *verifiable* work activities at high enough behavioral specificity allow QC review
- Hierarchical GWA taxonomy to link rated job activities all the way to Fine's "Big Three" T-D-P
- Defensible "paper trail" using traditional decomposed-judgment (factor scoring) methods to estimate scores on higher-level GWAs
- "Open source" national database based on contributions of participating orgs, individuals
- Quality-control via rater certification system
- JCV equations to empirically link JA to JS/worker-traits (public-domain and private publishers)
- 3rd party add-ons address more focused HR needs
- Database and reporting available via SOA web-services

Where I'm Going with CMQ

- CMQ can be the foundation of a national JOA system
- Database over 10,000 profiles
- CMQ available online CMQ for several years
- "Open source" type of model
  - Give away the job analysis
  - Sell value-added services to make it useful
- Near-term goals
  - Updated data-collection engine
  - SOA-based access to database, reporting
  - Find development partners who can contribute KS and AO profiles for occupations to do more JCV
  - Additional languages

Your participation encouraged

- Web-based version of CMQ is getting a major overhaul:
  - http://cmqonline.com
- Need organizations to participate in JCV studies
  - Worker-trait profiles for occupations
  - Or profiles plus the JA too
- Email me at rj@pstc.com