

Assessment Centers in the Federal Government

Suzanne Tsacoumis, Dan J. Putka, & Michelle Davis King

34th International Congress on Assessment Center Methods

September 24, 2008



Bureau of Alcohol, Tobacco, Firearms and Explosives

- Federal law enforcement agency within the Department of Justice
- Dual responsibilities for enforcing Federal criminal laws and regulating firearms and explosives industries
- Identifies, investigates and recommends for prosecution violators of the Federal firearms, explosives, arson, tobacco, and alcohol diversion laws.



Project Background

- 1996 Settlement agreement to change human resource practices for special agents
- Promotion systems for
 - 1st-level supervisor (grade 14)
 - 2nd-level supervisor (grade 15)
- Oversight committee
 - Plaintiff and defendant representatives
 - Two external “non-partisan” members
 - ATF representative



Sample Competencies

- General investigative knowledge
- Knowledge of relevant laws, regulations, and policies
- Knowledge of administrative procedures
- Judgment and problem solving
- Plan, organize, and prioritize
- Relate to others
- Lead others
- Communicate orally



Assessment Center Exercises

The candidate plays the role of the supervisor

- **Role Play #1:** Meeting with an agent to help plan an upcoming operation
- **Role Play #2:** Meeting with a U.S. Attorney to discuss a case
- **Analysis Exercise:** Review of a case file with recommendations for how to proceed with the investigation
- **Oral Presentation:** Presentation to local neighborhood group
- **In-Basket Exercise**



Implementation: Assessors

- ATF managers
 - Technical knowledge
 - Credibility
- 2 assessors per candidate per exercise
 - Different team for each exercise
 - Rotate every day



Implementation: Room Set-up

- Director's room
- Two main administration rooms
 - Candidates prepare for simulations
 - Computers
- Pairs of assessors in break-out rooms
 - Candidates interact with assessors
 - Video cameras



Implementation: Assessors/Candidates

- 1 to 1 ratio
 - Staggered schedule
 - One group interacts with assessors while other group is preparing
- Familiarity and objectivity ratings
 - ≤ 3 familiarity (on 5-point scale)
 - 1 objectivity (on 3-point scale)
 - Computer program to match candidates to assessors



Implementation: Scoring

- Two assessors make independent ratings immediately following interaction with candidate
- Compare ratings and ensure identical or 1 point apart
- No end-of-day consensus meeting
- Scores mechanically combined



AC Design and Results

- Overview of AC design
- Overview of AC scoring
- Insights from a multiyear assessment program



Overview of AC Design

- Grade 14 AC

- 2002 $N = 170$

- 2003 $N = 100$

- 2006 $N = 275$

- 2008 $N = 400+$

- Grade 15 AC

- 2003 $N = 71$

- 2004 $N = 41$

- 2005 $N = 13$

- 2007 $N = 85$



Overview of AC Scoring

- Multitude of options afforded by AC design
 - Each candidate completes four exercises
 - Each exercise measures ~ 4-10 competencies
 - Competencies assessed partially overlap across exercises
- Multiple “levels” of scores are possible
 - Overall AC score
 - Exercise-level scores
 - Competency-level scores
 - Competency-within-exercise scores



The Overall AC Score

- Examined over 30 different scoring algorithms
 - Unit, reliability-based, job-analysis based weights
- Final score based on job-analysis weighted composite of competency-within-exercise scores
 - The greater number of critical job tasks a given competency is linked to, the greater weight it is given
 - Produces an overall AC score on a 100pt scale



...but does weighting *really* matter?

- Empirically, no.
 - All options resulted in similar reliabilities, subgroup differences, and rank-ordering of candidates
- Conceptually, yes.
 - Linking algorithm to job analysis data conceptually contributes to content validity evidence



Insights from a Multiyear Assessment

- Able to explore several interesting experience-related effects
 - **Retest effects**
 - Some candidates completed the AC twice
 - **Experience-as-an-assessor effects**
 - Some candidates previously served as an assessor
 - **Experience-as-a-supervisor effects**
 - Some candidates had “acting” supervisory experience

...



Retest Effects

- Q1. Did candidates improve their AC performance upon retesting?
 - Within-subjects question
 - Retesters' 1st time AC scores vs. their 2nd time AC scores
- Q2: Was the performance of “retesters” better than those completing the AC for their first time?
 - Between-subjects question
 - Restesters' AC scores in 2006 vs. 1st time testers' AC scores in 2006
-]



Retest Effects

- Q1. Did candidates improve their AC performance upon retesting?
 - Yes, $d = 0.53$
 - 2nd Time Score -1st Time Score: $N = 41$, $M = 6.02$, $SD = 11.42$
 - 1st Time: $M = 63.10$, $SD = 7.95$; 2nd Time: $M = 69.13$, $SD = 9.91$
- Q2. Was the performance of “retesters” better than those completing the AC for their first time?
 - No, $d = -0.11$
 - Retesters’ AC scores in 2006
 - $N = 41$, $M = 69.13$, $SD = 9.91$
 - 1st time testers’ scores in 2006
 - $N = 234$, $M = 70.37$, $SD = 11.74$



Experience-as-an-Assessor Effects

- Q3. Was the performance of candidates who previously served as an assessor for the AC process better than those candidates who had no such experience?
 - Between-subjects question
 - Candidates with assessor experience vs. those without
 - Yes, $d = 0.70$, but this is based on a very small sample...
 - Candidates w/assessor experience
 - $N = 11$, $M = 79.0$, $SD = 6.80$
 - Candidates w/o assessor experience
 - $N = 74$, $M = 71.8$, $SD = 10.6$



Experience-as-a-Supervisor Effects

- Q4. Was the performance of candidates who previously served as an acting supervisor better than those candidates who had no such experience?
 - Between-subjects question
 - Candidates with supervisory experience vs. those without
 - Yes, $d = 0.30$, but the effect was small
 - Candidates w/supervisory experience
 - $N = 199$, $M = 72.40$, $SD = 11.18$
 - Candidates w/o supervisory experience:
 - $N = 302$, $M = 69.04$, $SD = 11.20$



The Selection Process

- 3 Unique Aspects
 - Role of the Selection Board
 - Role of the Receiving Manager
 - Use of AC Data



Role of the Selection Board

- Selections are made by a 3-member board
 - Board does not know candidates names or scores
 - Pre-selection screening process prevents problems
 - No interviews, just a review of paper



Role of the Receiving Manager

- Input limited to...
 - Identifying critical competencies
 - Takes place before the vacancy is announced
 - Targets needed skills; not a specific individual
 - Participates as a member of the selection board
 - No interviews

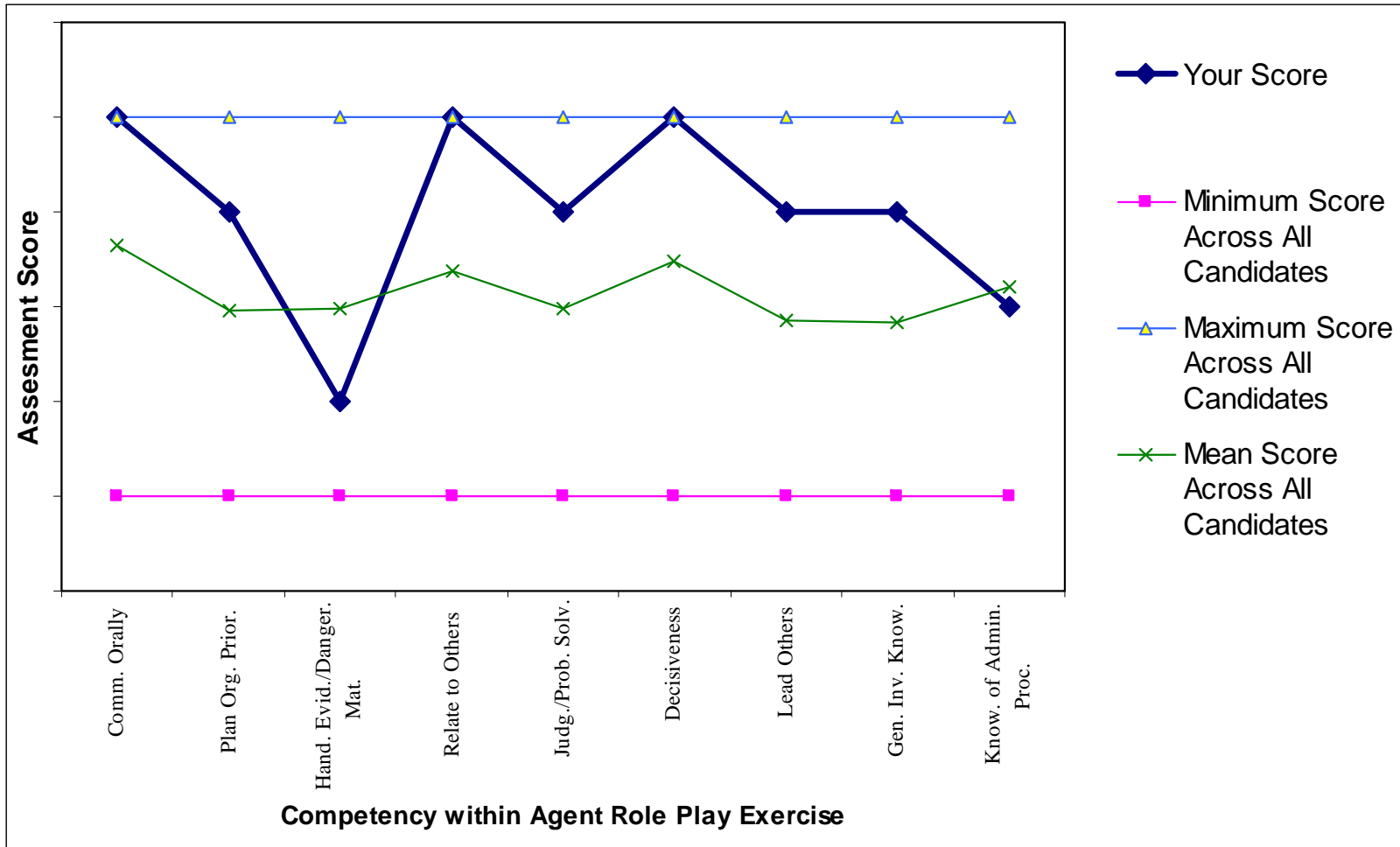


Use of AC Data

- Participants receive a score, plus profiles showing how they performed in each exercise
- Same profiles are provided to the selecting body



Sample Profile



Final Selections

- Selections based upon...
 - Critical competencies pre-identified by the manager
 - Profiles for each candidate outlining all competencies measured in AC
 - Pre-identified specialized or technical needs
 - e.g. certification as a polygraph examiner



Closing Thoughts

- A major paradigm shift for the agency
- Legal challenges drastically reduced
- Overall, viewed as an improvement
- The playing field is now level for everyone
- Still some frustrations
 - By managers, who want a bigger say in selection decisions
 - By employees, who feel that promotion decisions should be based on more than just the AC
- A living process, with ongoing reexamination and adjustment

