

Possible Solutions to Content Validity Challenges in Court; An Insider's Analysis and Insights

By
Joel P. Wiesen, Ph.D.
jw@jpwphd.com

2024 IPAC Annual Conference
New Orleans, LA, 7/29/2024

Wiesen (2024) IPAC Conference

1

Who Am I?

- Independent consultant
- Worked for the MA civil service agency
- A scientist-practitioner
- 46 year member of IPAC
- 45 years as expert for Defense and Plaintiff
- Expert in *Tatum* case for 17 years

Wiesen (2024) IPAC Conference

2

Goals of this Presentation

- Identify testing weaknesses seen in *Tatum*
 - Also weaknesses seen in other cases
- Propose ways to address these weaknesses
- Propose new approach to police promotion

Wiesen (2024) IPAC Conference

3

What is the *Tatum* Case?

- Minorities challenged 8 sergeant exams
- 2 exams for Boston
- 6 annual “statewide exams”
 - MC & E&E
 - Used for 100+ PD in Massachusetts
- Plaintiffs prevailed
- Detailed court decision (75 pages)

Wiesen (2024) IPAC Conference

4

Role of a Police Sergeant

- Sergeants supervise officers (about 7)
- Sergeants spend much time in the field
 - Go to the most serious incidents
- No sergeant at the Geo. Floyd incident
 - Was a minor crime of passing a fake \$20 bill
- Officers ask sgts. if uncertain (law, SOPs)
- Sgts must answer quickly (instantaneously)

Wiesen (2024) IPAC Conference

5

Tatum Decision: Discrimination

“**Overwhelmingly persuasive evidence** proves that HRD interfered with the class members' rights to consideration for promotion to police sergeant without regard to race or national origin.”
[emphasis added]

Wiesen (2024) IPAC Conference

6

Decision: Conclusion

“...a discriminatory system that has injured qualified candidates and **deprived the public of the benefits of having the best-qualified police sergeants.**”
[emphasis added]

Wiesen (2024) IPAC Conference

7

Summary of Tatum Decision

- JK tests consistently had adverse impact
- Intent discrimination based on past impact
- JK tests measured rote memorization
- JK tests did not measure important KSAPs
- JK tests invalid, especially for ranking
- Did not use alternatives with less AI
(JK=job knowledge)

Wiesen (2024) IPAC Conference

8

Decision Example 1

“According to CP, the most critical determinant of future success as a community policing Officer is:

- A. Superior communication skills.
- B. Empathy. (key)
- C. Autonomy.
- D. Analytical ability.”

Wiesen (2024) IPAC Conference

9

Decision Example 1

- Defense said this item measured empathy
- The court critiqued this item as only measuring knowledge about empathy but not the ability to be empathetic or foster empathy in subordinates

Decision Example 2

“...the exams did not test many important job qualifications.”
“...not measure ability to **apply knowledge practically** and to **exercise judgment** on that topic **in specific situations**”

Decision Example 3

“Most of the questions on the exams at issue in this case tested topics that were important to the job of sergeant. That does not mean that HRD's format was reasonably job related. It was not.”

Related Police Promotion Cases

- Two closely related impact cases in MA
- *Lopez v Lawrence*, 2014 (trial in 2009)
 - Same claim, but in federal court
 - Court ruled exams were “**minimally valid**”
- *Smith v Boston*, 2015
 - Same claim in federal court but for lieut. exam
 - Court ruled 2 Boston exams were **invalid**

Wiesen (2024) IPAC Conference

13

Smith 2015

“... the Court has found that ... too many skills and abilities were missing from the 2008 test outline.”

Wiesen (2024) IPAC Conference

14

Issues and Proposals

- Some major issues/flaws in testing
- Some little discussed
- Based on *Tatum* and other exams I reviewed
- Approach of the rest of this presentation:
 - Describe issue
 - Offer a proposed solution

Wiesen (2024) IPAC Conference

15

Topics of Issues

- Items (7 issues)
- Job analysis (6 issues)
- Test outline and misc. (7 issues)
- New approach to promotional exams

Wiesen (2024) IPAC Conference

16

1. Issue: Allowing Item Appeals

- Some civil service rules allow item appeals
- Appeals typically heard by a lay body
- Item upheld if it closely reflects the source
- So, items quote sources
- Items measure recall of wording in source
- Little measurement of practical application

Wiesen (2024) IPAC Conference

17

Proposal: Allowing Item Appeals

- New law/rule to grant authority to SMEs
 - e.g.: Post-test agreement of 3 SMEs presumed to be adequate support for an item (SMEs who did not write the item)
 - Involve police academy, municipal attorney
- Concern: Candidates will say they cannot study for SME questions

Wiesen (2024) IPAC Conference

18

2. Issue: Replicate Legal Cases

- Law questions are common on police exams
- Questions typically completely replicate **all** the facts of an actual court case
- This avoids appeals
- Does not replicate job duties of sergeant
- Officers must respond to incidents that do not fully replicate past court cases.

Proposal: Replicate Legal Cases

- Use items that do not fully replicate past cases
- Require judgment in applying precedents
- Involve local attorney in item development
 - City attorney, etc.

3. Issue: Definition Items

- Easy to write & defend definition questions
- Knowing a definition does not mean person can use the concept

Proposal: Definition Items

- Use only a small proportion of definition items for any given KSAP
- Use definition item only if the definition is important to know in order to do the job

Wiesen (2024) IPAC Conference

22

4. Issue: Items on Procedures

- Easy to ask procedural step order or names
- Knowing the correct order of steps or names does not mean person can execute the steps
- Q: In the SARA problem-solving model, what should be done in the analysis step?
 - key: Collect information from a variety of public and private sources
 - Does not test if able to collect information

Wiesen (2024) IPAC Conference

23

Proposal: Items on Procedures

- Test the ability to implement the steps
- Test the name or sequential order of steps only if these are important (e.g., step is likely to be done out of sequence)
- Use only a small proportion of step name or sequence items for any given KSAP

Wiesen (2024) IPAC Conference

24

5. Issue: Academic Items

- Question on desired leadership style
 - Key: balances concern for people and task
- A correct answer does not mean the person can do either well

Wiesen (2024) IPAC Conference

25

Proposal: Academic Items

- Use academic items only if application clear
- Items on important, applied topics
- More situational questions
 - Video stimuli
 - Constructed responses
- May require item writers to have more skills
 - Work with actors and video content creators

Wiesen (2024) IPAC Conference

26

6. Issue: Limited Item Review

- SME item review questions, 2008 exam:
- Suitability for rank?
 - No definition of suitability or suitable
- Estimated difficulty “for the persons taking the examination”
- Estimated readability

Wiesen (2024) IPAC Conference

27

Proposal: Limited Item Review

- Gather better information from the SMEs
- Improve the item rating form
- Clarify the review topics and rating levels
 - Is this K important to do the job?
 - Is this the best way to measure this knowledge?
- Talk with SMEs about each item.
 - How is this knowledge used on the job?

Wiesen (2024) IPAC Conference

28

7. Issue: Job Analysis Accuracy

- Task and KSAP inventories with implausible results
 - Tasks not done daily; KSAPs omitted
- Major disagreement among SMEs
- Illogical ratings:
 - Tasks of budgeting; read, interpret tables/graphs: but no math ability required

Wiesen (2024) IPAC Conference

29

Proposal: Job Analysis Accuracy

- Do not blindly rely on SME ratings
- Probe discrepant and suspect ratings
- Review the JA results for plausibility
- Conduct reviews of JA findings with SMEs
- Gather ratings on KSAPs from
 - Supervisors
 - Training academy staff

Wiesen (2024) IPAC Conference

30

8. Issue: KSAPs Not Well Defined

(A) K of principles of management
versus

(B) POSDCORB areas listed separately

- If (A), are all SMEs rating the same area?
- If (A), how much emphasis on each facet?
- Lack of clarity affects job analysis
- Lack of clarity affects test outline

Wiesen (2024) IPAC Conference

31

Issue: KSAPs Not Well Defined

- Consider these 3 K statements from a job analysis for police lieutenant. and captain:
- Principles of police administration
- Supervision, management, and leadership principles
- Community-policing and problem-solving principles
- What do these cover?

Wiesen (2024) IPAC Conference

32

Proposal: KSAPs Not Well Defined

- Ask if every SME will agree on KSAP scope
- Ask if the KSAP can be broken down into components that are not highly correlated
- Try using operational definitions of KSAPs

Wiesen (2024) IPAC Conference

33

9. Issue: Many Tasks/KSAPs

- Often there are many tasks and KSAPs
- Group tasks into categories loses detail
- Grouping KSAPs into broad competencies loses detail

Wiesen (2024) IPAC Conference

34

Proposal: Many Tasks/KSAPs

- Use tasks and KSAPs when writing items
- Do not rely on task/KSAP groupings

Wiesen (2024) IPAC Conference

35

10. Issue: KSAP Weight

- Now we rate KSAPs with **ordinal** scales
 - Mean importance guides test plan
- But no standard size of a KSAP
 - Some KSAPs encompass much information
 - Some sources are many pages
- Generating items easier for some sources

Wiesen (2024) IPAC Conference

36

Proposal: KSAP Weight

- Rate KSAPs with **ratio** scale
- How much of successful job performance depends on this KSAP?
- Allot 100 or 1,000 points among the KSAPs
 - Use Excel to ease math burden
- Frank Landy used this approach in his job analysis of police officer in Massachusetts

Wiesen (2024) IPAC Conference

37

11. Issue: Past Job Performance

- Calls to measure past job performance
 - Fields (2007) PTC Presidential Message
- Empirically keyed biodata has high validity
 - In top 3
 - Sackett, Zhang, Berry, and Lievens (2022)

Wiesen (2024) IPAC Conference

38

Proposal: Past Job Performance

- Body Cam Review
 - By outside raters
 - Videos provided by candidate and supervisor
- Accomplishment Record
 - Fields (2007) PTC Presidential Message
- Job Performance Evaluation
 - Many articles
 - Landy (1977) Police Foundation Report

Wiesen (2024) IPAC Conference

39

12. Issue: Critical Incident Usage

- List of tasks (briefly stated) is inadequate
- Short “ride-a-longs” are inadequate
- Often critical incidents are not collected

Wiesen (2024) IPAC Conference

40

Proposal: Critical Incident Usage

- Collect critical incidents from incumbents
- Collect critical incidents from supervisors
- Goal: Many hundreds of incidents
- More nuanced understanding of the job
- Incidents provide grist for item writing

Wiesen (2024) IPAC Conference

41

13. Issue: Outline Based on Tasks

- Typically test outlines are based on KSAPs
- KSAPs are a step removed from the job
- Sources for KSAPs can be quite academic
- How to tell if a test is representative of job?
 - Covering KSAPs ≠ covering tasks
 - KSAP importance may not map to task criticality

Wiesen (2024) IPAC Conference

42

Proposal: Outline Based on Tasks

- Develop test outline based on tasks
- Can have 2-way outline
 - sources and tasks
 - KSAPs (or KSAP groupings) and tasks
- Easier to show test is representative of job
- Ask SMEs/candidates about missing topics

Wiesen (2024) IPAC Conference

43

14. Issue: Weighting by # Items

- Important topics get more questions
- Score on exam is typically # correct
- Easy and hard items have same weight
- Some topics tested with few items
 - **No reliable** measure of such topics

Wiesen (2024) IPAC Conference

44

Proposal: Weighting by # Items

- Validity capped by square root of reliability
- Valid topic w/ few items ➔ invalid measure
- Enough items to reliably measure a topic
 - Minimum of 10 items per area
- Weight topic scores by importance

Wiesen (2024) IPAC Conference

45

15. Issue: Item Weight

- We now weight all items equally.

Wiesen (2024) IPAC Conference

46

Proposal: Item Weight

- Weight item by increase in job performance
- Weight item by consequence of error

Wiesen (2024) IPAC Conference

47

16. Issue: Setting Passing Point

- Angoff rating is compensatory
 - An essential area might have all easy questions
- Illusive “minimally qualified incumbent”
 - This is indirect rating of passing point
 - Assumes that because a minimally qualified person knows the topic, the topic is required to do the job!

Wiesen (2024) IPAC Conference

48

Proposal: Setting Passing Point

- Consider alternatives to Angoff question
- Rate exam as an entity
- How many items answered correctly indicate a person can do the job?
- Will test screen in people w ability to do the job and screen out unqualified people?
- Do this by test area

Wiesen (2024) IPAC Conference

49

17. Issue: Single Pass Point

- Items weighted equally
- Grading compensatory
- Can pass exam with zero on a KSAP
- If only one person passes exam ➡ promote

Wiesen (2024) IPAC Conference

50

Proposal: Single Pass Point

- Identify essential KSAPs
- Set passing point for each essential KSAP

Wiesen (2024) IPAC Conference

51

18. Issue: Test Outline Secrecy

- Often testing groups do not reveal outlines
- Often sources are voluminous
- Candidates do not know what to study most
- Identifying material to study is not job related

Wiesen (2024) IPAC Conference

52

Proposal: Test Outline Secrecy

- Tell candidates the # items per source
 - Allows candidates to apportion study time
- Give guidance on what will **not** be tested
 - History older than 10 years
 - Chapters x, y, z

Wiesen (2024) IPAC Conference

53

19. Issue: Professional Secrecy

- There are no compendia of:
 - Test outlines
 - BARS scales
 - Practical exercises

Wiesen (2024) IPAC Conference

54

Proposal: Professional Secrecy

- Perhaps IPAC could publish compendia

20. Issue: BARS Reliability

- Within board reliability: 0.97, 0.91
 - Between board reliability: **0.44**
 - But $.97 \times .91 = .88$ ($n > 100$)
 - Within board reliability: 0.97, 0.94
 - Between board reliability: **0.66**
 - But $.97 \times .94 = .91$ ($n > 100$)
- (Note: data not from MA)

Proposal: BARS Reliability

- Use duplicate rating boards
- Research into reasons for disagreement
 - Within a board
 - Across boards

New Approach to Exams

- For police promotional exams
- Several goals of the new approach
- Better job performance of sergeants
- Reduced adverse impact
- Better acceptance of promotion process
 - Candidates
 - Management

Wiesen (2024) IPAC Conference

58

Issue: New Approach Needed

- Candidates prepare themselves for promo.
- Hard to learn supervision, management, leadership, strategy, tactics, etc. from books
- Exam grades show candidates lack KSAPs
- Many high paying occupations have training programs
 - \$200,000+ average gross pay, Boston Sergeant

Wiesen (2024) IPAC Conference

59

Issue: New Approach Needed

- Currently, many/most exam scores are low
 - A low-scoring candidate may be promoted
- A promotion exam with 100+ takers:
 - Lowest closed book score was 67%
 - Of promotees, over half < 80% on closed book
- Candidates weak in essential KSAPs
- No training for newly promoted sergeants

Wiesen (2024) IPAC Conference

60

Proposal: New Approach Needed

- Establish a thorough training program for promotional candidates
 - Strategy and tactics for incidents
 - Planning and resource allocation
 - Interpersonal aspects of policing
- This is a major undertaking
 - Fund course development
 - Fund training time for candidates

Wiesen (2024) IPAC Conference

61

In Closing

- We can improve police testing practices
 - Better content valid knowledge tests
- Both candidates and management benefit
- Avoid or win more testing court cases
- **Improve job performance of sergeants**

Wiesen (2024) IPAC Conference

62

Contact Information

- Session URL: **<http://jpwphd.com/ipac2024>**
- Related 2024 SIOP Master Tutorial
- Email: **jw@jpwphd.com**
- Telephone: **(617) 244-8859 (land/no text)**
- **Email and telephone calls welcome!**
- **Q&A's**

Wiesen (2024) IPAC Conference

63

Q&A

Main Cases Cited

- *Lopez v. Lawrence*
- *Lopez v. City of Lawrence*, Civil Action No. 07-11693-GAO (D. Mass. Sep. 5, 2014)
- <https://casetext.com/case/lopez-v-city-of-lawrence-1?q=07-11693-GAO&sort=relevance&p=1&type=case>

Main Cases Cited

- *Smith v. Boston*
- *Smith v. City of Boston*, 144 F. Supp. 3d 177 (D. Mass. 2015)
- <https://casetext.com/case/smith-v-city-of-bos-1?q=12-10291-WGY&sort=relevance&p=1&type=case>

Main Cases Cited

- *Tatum v. Massachusetts*
- Tatum v. Mass., C.A. 0984CV00576 (Sup. Court 2022)
- <https://www.mass.gov/doc/tatum-et-al-v-human-resources-division-related-superior-court-decision-102722/download>

Wiesen (2024) IPAC Conference

67

Presentation Citation

Wiesen, J.P. (2024). Possible Solutions to Content Validity Challenges in Court; An Insider's Analysis and Insights. *International Personnel Assessment Council*, New Orleans, LA, United States.

Wiesen (2024) IPAC Conference

68
