Developing and Conducting Effective, Fair, and Valid Interviews

Jim Kuthy, Ph.D.
CritiCall Pre-Employment Testing Software

Heather Patchell, M.A.
Biddle Consulting Group, Inc.
Employment Interviews

- Can be a very reliable and valid method for helping to choose the best people for a job
- If an interview is used to make employment decisions (such as hiring or promotions), it is subject to the requirements of the federal Uniform Guidelines on Employee Selection Procedures (1978), similar to any other type of test or selection procedure
  - www.uniformguidelines.com
Theories Underlying the Use of Interviews

- **Goal Setting**: Intentions are the immediate precursor of a person’s actions

- **Behavioral Consistency**: The past is the best predictor of the future


What should be measured during an interview?

- Identify what you want to measure during an interview based upon the findings of a **job analysis**...
  - *Job analyses typically are based on information from those who currently hold the target job and/or from observations of those performing the job*
  - *Job analyses often include information based on critical incidents*
- It is best to limit yourself to **several key areas** identified during the job analysis as important and/or critical
  - *Do not try to measure too much during a single interview*
What should be measured during an interview?

- Job knowledge, job experience, and ability to perform job-related behaviors
- Measure appropriate, observable behaviors (and/or knowledge that results in observable behaviors)
  - Not traits or characteristics that only appear over time (such as creativity, dependability, and honesty)
  - Not behaviors that would be more appropriately measured using other methods
It would be more appropriate to use a test of computer skills than an interview to reliably learn if this person is proficient at using a computer.

In the interview he said he had hundreds of hours on the computer.
Use Interviews Wisely

• Use interviews in combination with other selection tools
  – Do not use only an interview in an attempt to measure everything
  – Only give as much weight to the interview as is justified by the job analysis
Three Types of Interview Questions

• Psychological
  – Good, but typically administered by a trained clinical/counseling psychologist

• Experienced-Based
  – Better for many jobs, best for a few

• Situational
  – Best for the vast majority of jobs
Types of Interview Questions

- **Experienced-Based Questions**
  - Focuses on evaluation of reactions to actual situations from the candidate’s past that are relevant to the target position
    - “Tell me about a time…”
    - “Describe a time you…”
  - Good for many jobs, best for a few
  - Not typically appropriate for many types of entry-level jobs or where many of the applicants do not have previous experience in the areas being measured
Types of Interview Questions

• **Situational Questions**
  – Focuses on evaluations of intended reactions to hypothetical job situations
    • “What would you do in this situation?”
    • “Describe how you would handle this situation”
    • “List the steps you would take…”
  – Be sure NOT to ask about knowledge, skills, or abilities they would learn “on the job” or during training after being hired
Which type of Question is better?

• In general, there is a relatively strong correlation between the results of situational and behavioral questions (i.e. the results are similar)
  – \( r = .85 \) (Conway & Peneo, 1999)
  – \( r = .76 \) (Campion, Campion, & Hudson, 1994)

• Use the type of questions that are most appropriate, considering both the job and the applicant pool


Potential wording of a “sticky question” if you must ask...

• **NOT** - “How many sick days off did you take on your last job?”

• **YES** –
  - Describe your attendance requirements and ask if he or she can meet them.
  - “How many days were you absent last year” or, if you wish to determine potential time-off abuse, “How many Mondays and Fridays were you absent last year?”
Google Admits Its Crazy Interview Questions Were 'A Complete Waste Of Time'

If you have ever experienced a job

What's Google's hiring tactic now? Bock says his team uses behavioral interviews and asks for examples of how a person acted in a particular situation. It no longer asks hypothetical questions. Google also uses structured behavior interviews with a consistent rubric for assessing each person.

charge to wash all the windows in Seattle?" and "Why are manholes round?" It now admits that those questions were useless.

"We found that brainteasers are a complete waste of time," Laszlo Bock, Google's Senior VP of
Using Interviews
Use Interviews Wisely

- It is a good idea to use interviews in combination with other selection tools
  - Using only an interview in an attempt to measure everything about a job candidate may be inappropriate
  - Only give as much weight to the interview as is justified by the job analysis
Use Formal, Structured Interviews

Ask the same questions in the same way... every time

(write it out in advance!)

We’re sick of asking the same old questions to each applicant, so we’ve decided to be flexible and creative... What’s your favorite football team?
Allows for an Apple to Apple Comparison
1. **Reliability is a pre-condition of validity**

2. **Standardizing interviews increases their reliability**

3. **Thus, standardizing your interviews will almost always increase their validity**

4. **Strong evidence of validity helps insure that your interview is job related and consistent with business necessity**
Example of a “Clean” Structured Interview Process

• Briefly introduce the interviewers
  – Consider having nameplates in front of each interviewer

• Explain the purpose and structure of interview
  – Set the ground rules, such as time limits

• Ask same/similar situational and/or experience-based questions

• Thank the interviewee and excuse them
Use Interview Time Wisely

• If you have limited time, there are a number of areas that may be better measured at other stages of the process and/or using other methods, such as…
  – Asking why they want the job
  – About their minimum qualifications
  – Asking if they have anything to add
Some Types of Follow-up and/or Customized Questions Can be Okay

• It is okay to ask follow-up questions so the interviewee can clarify their response or provide more detailed information
  – However, all interviewees should be asked the same types of follow-up questions if they present similar information
  – You might consider “prompts” such as: “Can you tell us more about that…” or “Can you please describe that in another way for us…”

• Customized questions regarding specific aspects of the candidate might be acceptable (such as specifics about information contained in resume), as long as all interviewees who present the same information are asked similar customized questions
  – Write these out in advance of the interview to help insure consistency
Train your Interviewers/Raters

- Provide structured training to standardize process and minimize inconsistency across raters
  - Interview process
  - “Frame of Reference” training – Rating dimensions
  - Rating scales (including how to use anchors when rating)
  - Potential rating errors
- Provide examples
- Practice/role-playing
- Provide feedback and then practice again
- Test Raters
- Monitor the interview process as it progresses
- Repeat/update training if sufficient time has passed between the last training session and an interview

*Hint: Document your training!*
Anchor Your Rating Scales

- Anchors are benchmarks that provide examples of poor, acceptable, and highly-acceptable responses.
- These remind raters of different levels of competence.
- At least one study found that when ratings of naive raters were compared to job experts, the use of anchored rating scales resulted in similar reliability and accuracy for both groups (Maurer, 2002).

In addition to evaluating verbal responses you can also train interviewers how to...

- ...rate other types of observable behaviors, such as speaking and/or decision-making, if supported by a job analysis
  - We generally suggest that these be scored after all of the questions have been asked
Dimension H: SPEAKING ABILITY (All Questions)

Definition

The extent to which the applicant clearly, concisely, and effectively verbally conveys ideas, thoughts, information, etc., which can be easily understood or acted upon by others.

Characteristics

a. **Volume of Speech:** Does applicant speak at a decent volume (neither too loud nor too soft) to be easily heard? If volume of speech is a problem, does applicant make appropriate adjustments when requested?

b. **Rate of Speech:** Does applicant’s rate of speech (e.g., too fast or too slow) interfere with comprehension of what is being said? If rate of speech is a problem, does applicant make appropriate adjustments when requested?

c. **Pronunciation:** Is applicant's pronunciation of words clear and distinct?

d. **Word Usage:** Does applicant's word usage and speaking vocabulary coincide with what is reasonably required in the performance of the job in question?

e. **Organization:** Are applicant's responses organized and of sufficient detail? Does applicant tend to ramble in an unorganized fashion or insert irrelevant information or details?

f. **Thoroughness:** Are applicant's responses to questions complete and to the point? Is it necessary to repeatedly ask the applicant to elaborate on his/her responses?

g. **Listening Skills:** What is the extent to which applicant is attentive and properly interprets and/or responds to questions?
## Evaluation Scale - SPEAKING ABILITY – All Questions

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unacceptable</td>
<td>Acceptable</td>
<td>High Acceptable/Outstanding</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Scale Value Descriptions

#### Unacceptable

1. Responses are incomplete, sketchy, and tend to ramble. Speech volume is low and responses are difficult to hear. Rate of speech and pronunciation also contribute to the problem and applicant is not able to make adjustments when requested. Vocabulary skills are poor. Applicant is inattentive to questions when stated, to the point that consistent reminders are necessary.

2. Responses are vague and tend to be unclear. Speech volume and rate tend to fluctuate, but the majority of responses are difficult to hear and understand. Has difficulty organizing responses to effectively deal with problems stated. Requires frequent prompting. Vocabulary skills are limited.

#### Acceptable

3. Ideas and thoughts are adequately conveyed in an understandable, usually organized, manner. Responses occasionally require clarification. Vocabulary skills are sufficient to perform job. Applicant listens attentively and only rarely requires questions to be repeated.

#### High Acceptable/Outstanding

4. Vocabulary skills are ample. Responses are well-organized and very seldom require clarification. Answers are normally clear and concise. Requests for repetition of questions are very infrequent. Rate of speech, volume, and pronunciation are acceptable and responses are easy to understand.

5. Responses are always complete, to the point, and require no clarification. Information provided is clear and concise. Thoughts are always well formulated and organized. No repetition of questions is required. Vocabulary skills are high.
Frame-of-Reference Training

- Designed to provide a common evaluation standard of interviewees’ responses
- Target rating dimensions are presented and explained using behavioral examples at different performance levels
- Discussion of rating dimensions follows
- Practice actual evaluations (role playing)
  - One alternative to actual role playing is to provide examples of poor, acceptable, and highly acceptable responses using video or in writing and have the raters assign each to the appropriate performance level
- Provide feedback with explanation
- Continue cycle until raters share a similar frame of reference

Which is Better? Frame-of-Reference Training or Anchored Rating Scales?

- **Both** substantially improved rating accuracy and inter-rater agreement in a recent study
- Using **both** led to greater improvements than using either one alone

Potential wording of a “sticky question” if you must ask

- **NO** - “How many sick days off did you take on your last job?”

- **YES** –
  - Describe your attendance requirements and ask if he or she can meet them.
  - “How many days were you absent last year” or, if you wish to determine potential time-off abuse, “How many Mondays and Fridays were you absent last year?”
Also...Train About Potential Rating Errors

Rating errors tend to make the interview process less effective for selecting the best people for the job.
• **First impressions** (studies show early impressions can crystallize in only FOUR minutes!)

• **Self-fulfilling prophesies** (Be careful what pre-interview information is provided to the interviewer)
Other Potential Rating Errors

- Raters should be familiar with:
  - Halo errors
  - Leniency/Harshness errors
  - Primacy and Recency effects
  - Contrast effects - *comparing one applicant to another, rather than to a “standard”*
  - Stereotyping interviewees
  - “Similar to me”/”Different than me” biases

*Reinforce learning by performing “mock” interviews and providing feedback/insight*
Panel vs. Single Interviewers

**Panel Interview**
- More expensive (-)
- Typically more reliable, thus more valid (+)
- More information is used when making selection decision (+)
- Allows decisions to be reviewed by others before decision is made (+)

**Single Interviewer**
- Less expensive (+)
- Lower reliability & validity (-)
- Typically does not allow decision to reviewed by others (-)
- Individual interviewers do not get better without formal feedback (-)
You might Increase Fairness with Panel Interviews

• Mixed race/gender panels may help reduce the “similar-to-me”/“different from me” biases that individual interviewers might introduce
Take notes during the interview

Without notes: Less than half of managers could accurately report on the information produced during a 20 minute interview

Do **NOT** rely on...

- **Non-verbal behaviors that are not directly related to the job**
  - TV shows that suggest people can positively identify honest responses are not really true.
  - Studies consistently show that humans, even trained ones, make lousy lie detectors.
  - Decisions should be made on actual behaviors, such as the person’s responses or the clarity of their responses.
  - Relying on non-verbal behaviors that are not job-related is **not** defensible in the event of litigation.

- **Body type**
  - Relationships between body type and personality or behaviors are generally based on false stereotypes.
  - Do not rely on this information when making decisions.
Get Some Sleep!

- Research finds that suppressing stereotypes and prejudices requires self-control over one’s thoughts.
- Sleep is necessary for exerting self control.
- A recent study suggests that lack of sleep (or when raters are sleepy) can negatively influence the suppression of prejudice.
  - Even relatively small amounts of lost sleep can have effects on prejudice.

Best Practices

• Combine rating scores mechanically
  – Average or sum them
  – Consensus among raters is good goal, but different raters do not need to absolutely agree for ratings to be valid
    • You can average ratings if multiple raters do not agree
    • Some large interview panels automatically remove the highest and lowest rating to minimize the effects of extreme ratings

• **Listen…** do not talk too much

• Document everything
Strengthening your Interviewers

- Provide your interviewers with feedback
  - Without feedback, experienced interviewers agree with each other to no greater extent than do interviewers with differing experiences
**Possible Enhancements**

- Video- or Audio-tape interviews for later review
  - *Potentially save time and money*
  - *Eliminate discrepant raters*
  - *Record of interview retained*
- If appropriate under the job analysis, ask some interview questions via a telephone or other non-visual medium
Is There a Connection Between Interview Process and Success in Court?

- **Research...**
  - Study involving 84 disparate treatment and 46 disparate impact cases where interviews were litigated
  - 17 interview characteristics were evaluated (e.g., objective, subjective, standardized, etc.).
  - Study resulted in clear findings that revealed the three primary ingredients for successful interview validity defense

Is There a Connection Between Interview Process and Success in Court?

**The Three Primary Factors Are…**

- **Interview objectivity and job relatedness**, such as:
  - Objective and specified criteria
  - Trained interviewers
  - Validation evidence
- **Standardized administration**, including:
  - Scoring guidelines being followed
  - Minimal rater discretion
  - Common questions across candidates
  - Consistency
- **Multiple Interviewers**
  - Implies a shared decision making process
  - Increases rater reliability
Validation: Have a Job Expert team evaluate and rate interview questions using the following ratings:

- Is the question clear and understandable?
- Is the question at an appropriate difficulty level?
- Does the question measure a Knowledge, Skill, Ability that is needed on the first day on the job (i.e., before training)?
- Is the question job related?
  - Does it represent a situation that has occurred or is likely to occur on the job?
- What job-related Knowledge, Skill, Ability is measured by this question?
- Will applicants be able to provide a sufficient response without possessing job-specific (or employer-specific) knowledge
  - Job specific or employer specific knowledge prior to entry to that job must be justified by the Job Analysis
Interview Question

Danger Zone!

- Marital Circumstances
- Age
- Disabilities
- Gender & Physical Appearance
- Citizenship & National Origin

Use validation as your guide

-- BE CAREFUL OF “SMALL TALK” --
“Small Talk”

• If you wish to informally talk with candidates to “break the ice,” stick to a pre-planned script to minimize potential problems or misinterpretations
  – *Do this even if you know the candidate*

• To help to maintain the perception of fairness, try to keep the length of the “small talk” similar across interviewees

• Limit informal conversations with candidates outside of the formal interview area
Did you know?

- Applicants react negatively to recruiters or interviewers with poor interpersonal skills.
- Applicants also generally have a low impression of the organization if the interviewer lacks good interpersonal skill.
- So… make sure your interviewers are competent communicators.
Additional Hints

• Combine rating scores mechanically
  – Average or sum them
  – Consensus among raters is good goal, but different raters do not need to absolutely agree… you can average their scores if they do not agree

• Listen… do not talk too much

• Document everything
Prepare for **unusual behavior**

- Have a formal process (including documentation) in place for including information about unusual statements, behavior, or dress during the interview itself
- Have a formal process in place (including documentation) for including information obtained outside of the interview setting into the selection process
  - *For example, integrating into the selection system information about an applicant who makes a racially derogatory remark while talking on the cell phone in the hall while waiting to be interviewed*
Finally...

It is your responsibility, not the interview’s developer, to insure that the interview is fair and valid

- Be proactive
  
  • Periodically review the validity report and modify the interview as needed and/or as the job changes
  
  • Monitor the interview process for potential issues on a regular basis
    
    – Look out for biased raters
  
  • Retrain or retire interviewers when appropriate
  
  • Stay current in your research
SOMETIMES I LIE AWAKE AT NIGHT, AND I ASK “IS LIFE A MULTIPLE-CHOICE TEST OR IS IT A TRUE OR FALSE TEST?”
THEN A VOICE COMES TO ME OUT OF THE DARK AND SAYS, "WE HATE TO TELL YOU THIS, BUT LIFE IS AN UNSTRUCTURED INTERVIEW."
Thank you

Questions or Comments?

Under *interests* you put “cannibalism.” Can you tell me a little more about that Mr. Smith?
Used by over 500 Public Safety Agencies

Used by thousands of organizations

Hundreds of Equal Employment Opportunity/AAP Clients Nationwide