

The background of the slide is a dark teal color with several lighter teal, wavy, horizontal bands that create a sense of movement and depth. The text is centered and uses a serif font.

Chipping Away at the Monument: A Critical Look at the Assessment Center Method

Symposium Presented at the
2000 IPMAAC Conference 2000
Arlington, Virginia

Background

- Assessment center (AC) method has taken on *monumental* status in human resource management
- Issues remain, however, regarding the quality of assessors' judgements
- Panelists will address plaguing issues which have implications for the continued viability of the AC method

Panelists

Lisa Donahue, M.S.
George Mason University

Kenneth N. Wexley, Ph.D.
Wexley Consulting & Human
Resource Decisions

Donna Denning, Ph.D.
City of Los Angeles

Deborah L. Whetzel, Ph.D.
U.S. Postal Service

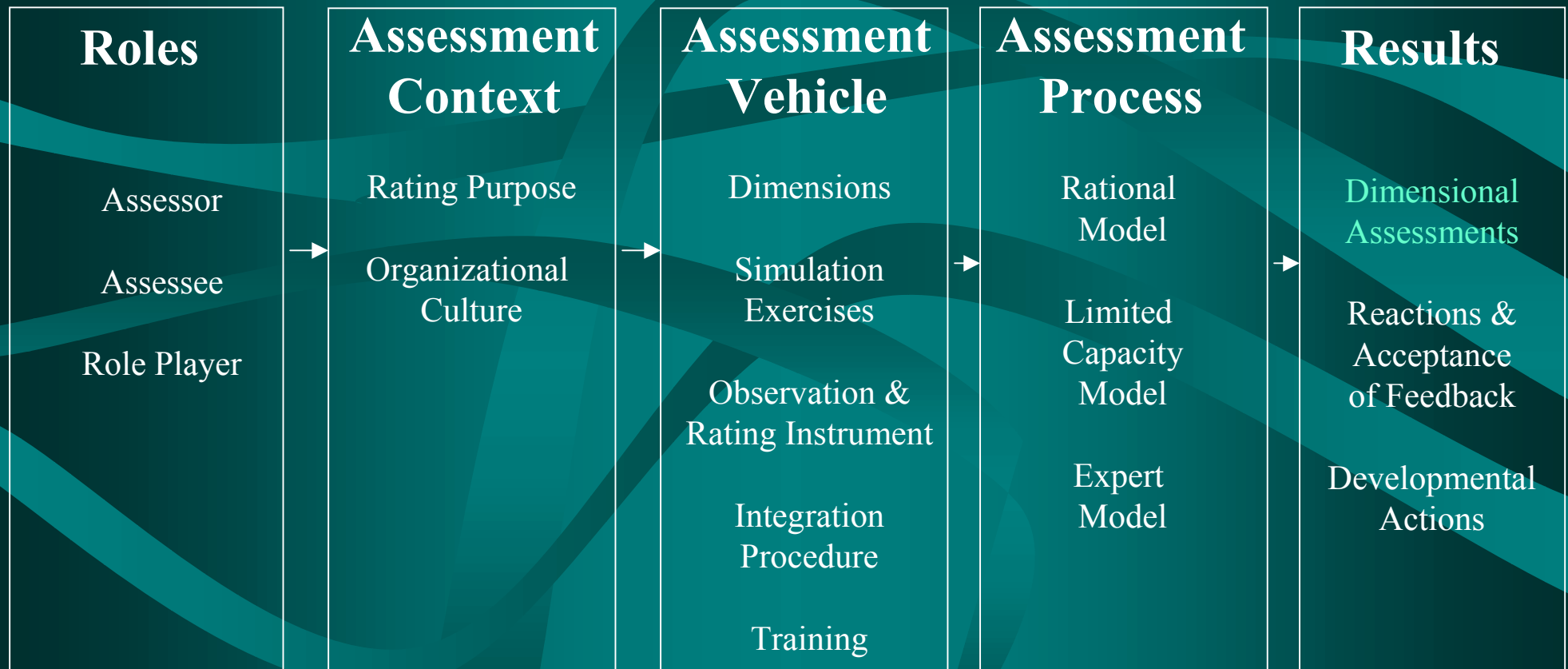
Richard J. Klimoski, Ph.D.
George Mason University

Discussion Focus: Quality AC Judgements

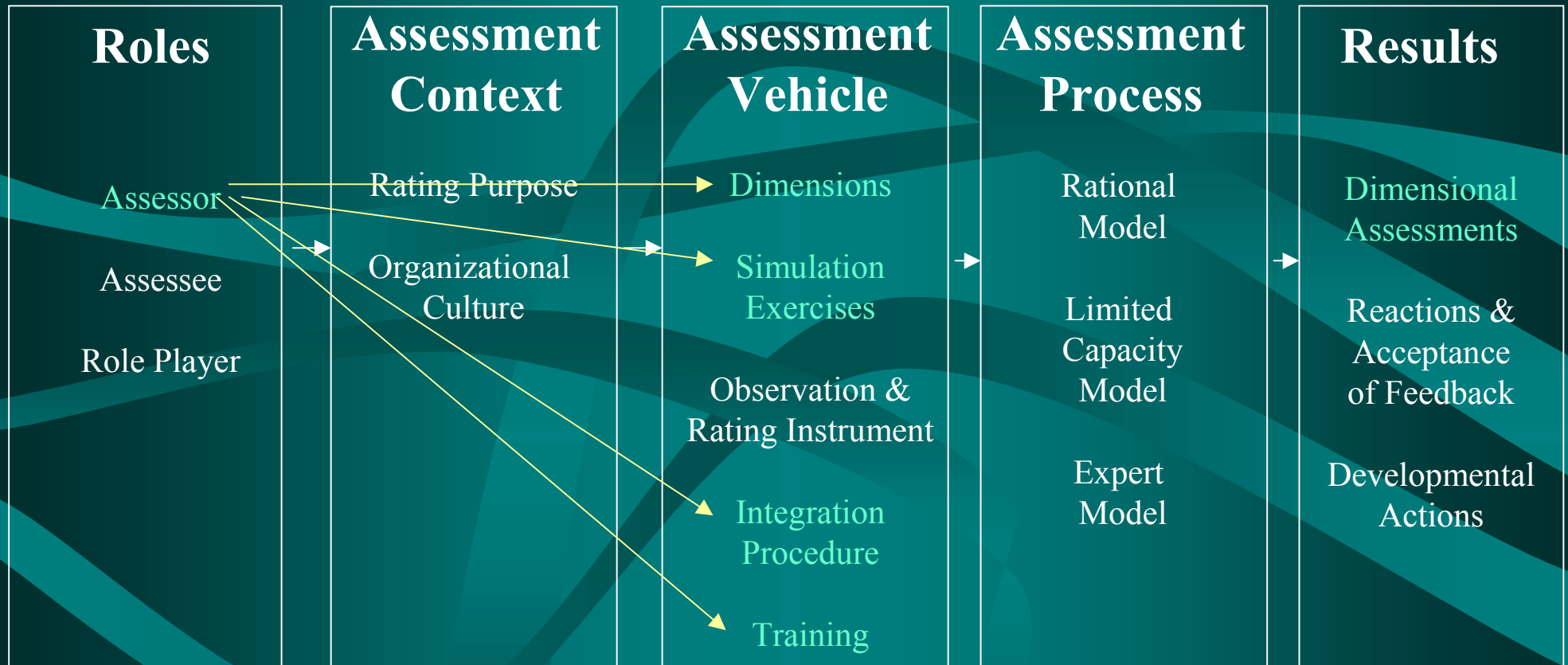
- Usefulness of AC method rests on its ability to produce high quality judgements
- Indices of quality include adequate levels of:
 - reliability
 - criterion-related
 - construct validity (convergent & discriminant)
 - acceptability by stakeholders

Framework for Analyzing Quality Judgements

Adapted from Lievens & Klimoski, in press



Panel Focus: Assessor-Components Link



Understanding the Assessor- Components Link

“...whereas most of the research has focused on what the components of the method ought to be, and how these components affect outcomes, there is little, if any research on how the components-results link is impacted by a particular group of persons—the assessors—and the dynamics of the situation in which they are functioning” (Zedeck, 1986, p. 260)

Understanding the Assessor- Components Link

- AC rating task is grounded not in information processing alone, but in social information processing
- Much can be contributed to our understanding of AC construct validity by considering the theories and findings within the domain of social cognition
- In adopting this perspective, the issues to be addressed by the panel represent *monumental ideas* in AC theory and practice

Overview of Discussion Questions

- *What type of training would enhance assessors' ability to make discriminant judgements?*

Kenneth Wexley

- *What is the best way to configure the assessors' role?*

Richard Klimoski

Overview of Discussion Questions (cont.)

- *What should AC designers assess?*

Primary: Donna Denning

- *How can AC exercises be designed to better elicit dimension-relevant behavior?*

Primary: Deborah Whetzel

Overview of Discussion Questions (cont.)

- *How can the integration process be improved?*
Primary: Richard Klimoski

Question #1:

What type of training would enhance assessors' ability to make discriminant judgements?

Background on Assessor Training

- Very little research exists on the effects of training on the quality of AC judgements
- Type of training, more than the length or amount, leads to more differential use of the AC dimensions (Lievens, 1998)
 - Frame-of-reference (FOR) training (Schleicher et al., 1999) is associated with greatest improvements in discriminant validity

Background on Assessor Training (cont.)

- *What type of training would enhance assessors' ability to make discriminant judgements?*
- *Can concepts within the realm of social cognition be integrated into assessor training to increase assessors' ability to make discriminant judgements?*

Enhancing Assessors' Ability To Make Discriminant Judgments & To Make Accurate Judgments

◆ **Panel of SMEs - Job Analysis**

- **Systematic Identification of the Competencies (via Linking)**
- **Define Them**
- **Refine Them**

◆ **Assessors**

- **Discuss the Definitions**
- **Provide Examples of Effective & Ineffective Behaviors For Each Competency**
- **Differentiate**

◆ **Mock Applicants**

- **Feedback Regarding Proper Placement of Their Behavioral Observations to Competencies (PDSs)**
- **Continued Feedback At The Rating Session**
- **Recognition That Certain Behavioral Observations May Sometimes Be Placed In More Than One Competency**

◆ Minimize Judgments Errors Experientially & Generate Assessment Center Solutions

- Halo
- Similarity
- Dissimilarity
- Leniency
- Comparison
- Central Tendency
- First & Last Impressions
- Stereotyping

Question #2:

*What is the best way to configure
the assessors' role?*

Background on Assessor Role Configuration

- In most operational ACs, there is a “division of labor” whereby assessors do not rate each candidate in every exercise
- Typically, assessors are assigned to a particular exercise
 - Poor convergent validity may result from low inter-rater agreement (Adams & Osburn, 1998)

Background on Assessor Role Configuration

- Some research has investigated the effects of alternative rating procedures (e.g., within-dimension rating) (Silverman et al., 1986; Harris et al., 1993) and rotation schemes (see Lievens & Klimoski, in press)
- Few studies have examined the effects of assigning assessors to specific:
 - assessees
 - exercises
 - dimensions across exercises

Assessee A

Exercise A

Dimension A

OR

Assessee B

Exercise B

Dimension B



Assessor

Background on Assessor Role Configuration

- *What is the best way to configure the assessor's role? What are the benefits/limitations relative to the quality of the dimension ratings associated with assigning assessors to specific (a) assessees, (b) exercises, and (c) dimensions within exercises?*

Question #3:

What should AC designers assess?

Background on AC Dimensions

- Research has found that discriminant validity is improved when AC dimensions are small in number, conceptually distinct, and general in nature (Lievens, 1998; Lievens & Klimoski, in press)
- Main problem is that AC dimensions are an amalgamation of “stuff” (Howard, 1997)
 - traits
 - learned skills (e.g., planning)
 - readily demonstrable behaviors (e.g., oral communication)
 - basic abilities (mental ability)
 - motives (e.g., need for achievement)
 - knowledge
 - other attributes

Background on AC Dimensions

- Further complicating matters, the definition of the dimensions is unclear, and their interpretation often changes from one exercise to the other
- Evidence from the social cognition literature that we are infer general traits (e.g., “The Big Five”) with greater accuracy and consistency than others (Fiske, 1993)

Background on AC Dimensions

- *So, what should be assessed in order to improve the construct validity of ratings? Traits? Performance dimensions (e.g. competencies)?*

Question #4:

How can AC exercises be designed to better elicit dimension-relevant behavior?

Background on Exercise Design

- Content validation approach primarily used to construct exercises
- Wide variations exists across exercises in the opportunity to demonstrate dimension-relevant behavior (e.g., Donahue et al., 1997; Reilly et al., 1990)

Background on Exercise Design

- Some evidence that construct validity is stronger for more observable dimensions (Shore et al., 1992)
 - Consistent with the principle of aggregation (Epstein, 1979)
- Findings that improved convergent validity exists for exercises similar in form (Schneider & Schmitt, 1992) and perceived to be similar by assessors (Highhouse & Harris, 1993)

Background on Exercise Design

- *Are current content validation-based approaches to exercise design sufficient? Are there other approaches that may better elicit the behavior that would reveal the level of the specific dimensions of interest?*

Question #5:

*How can the integration process
be improved?*

Background on Integration Process

- Integration procedure is very popular among operational ACs (Spsychalski et al., 1997)
- Still, very little research has been conducted on the integration process (Lievens & Klimoski, in press)
- In particular, “Group dynamics seems to be totally ignored with the assessment center literature” (Zedeck, 1986, p. 290)

Background on Integration Process

- Interaction can have both positive (e.g., social validation) and negative (e.g., polarization) effects on the group's decision
- *How can AC designers facilitate the type of interaction that will lead to improved discriminant validity? How can assessor teams be constructed and trained in order to increase the quality of their judgements?*

Assessor Team A



Oral
Communications

Planning &
Organization

Interpersonal

Candidate A



Assessor Team B



Planning &
Organization

Candidate A



Candidate B



Candidate C



Improving The Integration Process

- **Facilitator Decides on the Order For Assessing the Assesses**
- **Facilitator Asks For The Number of E's & IN's**
- **Read Out**
- **Ready To Rate, Or Need For Discussion?**
- **Make Independent Ratings (Acceptable or Unacceptable?)**
- **Poll**
 - **If Consensus ⇒ Move Onto Next Competency**
 - **If Disagreement ⇒ Further Discussion & Re-rate Until Consensus Reached**

Improving The Integration Process

Beh. Competen.	Group Discussion	Counseling	Customer	Team Building	Training	Writing	Reading & Comp.
Leader.	5E 0IN	2E 0 IN		12E 0 IN			
Interper. Skills	0E 13IN	0E 15IN	0E 4IN	0E 5IN	0E 5IN		
Prob. Solv. Skills	13E 4IN	5E 2IN	2E 11IN	6E 9IN			
Team Skills	2E 3IN	6E 3IN	8E 4IN	5E 11IN			
Motivates Others	4E+ 3IN	6E 5IN		7E 4IN			