

# Introduction to Assessment Centers

## (Assessment Centers 101)

A Workshop Conducted at the  
IPMAAC Conference on Public Personnel Assessment  
Charleston, South Carolina  
June 26, 1994

---

Bill Waldron   
Tampa Electric Company

Rich Joines  
Management & Personnel Systems



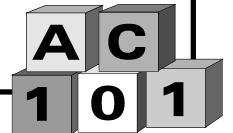
# Today's Topics



- ▶ What is—and isn't—an assessment center
- ▶ A brief history of assessment centers
- ▶ Basics of assessment center design and operation
- ▶ Validity evidence and research issues
- ▶ Recent developments in assessment methodology
- ▶ Public sector considerations
- ▶ Q & A



*What is an  
Assessment  
Center?*



# Assessment Center Defined

An assessment center consists of a standardized evaluation of behavior based on multiple inputs. Multiple trained observers and techniques are used. Judgments about behaviors are made, in major part, from specifically developed assessment simulations. These judgments are pooled in a meeting among the assessors or by a statistical integration process.

Source: *Guidelines and Ethical Considerations for Assessment Center Operations*. Task Force on Assessment Center Guidelines; endorsed by the 17th International Congress on the Assessment Center Method, May 1989.



# Essential Features of an Assessment Center

- ▶ Job analysis of relevant behaviors
- ▶ Measurement techniques selected based on job analysis
- ▶ Multiple measurement techniques used, including simulation exercises
- ▶ Assessors' behavioral observations classified into meaningful and relevant categories (dimensions, KSAOs)
- ▶ Multiple observations made for each dimension
- ▶ Multiple assessors used for each candidate
- ▶ Assessors trained to a performance standard

*Continues . . .*



# Essential Features of an Assessment Center

- ▶ Systematic methods of recording behavior
- ▶ Assessors prepare behavior reports in preparation for integration
- ▶ Integration of behaviors through:
  - Pooling of information from assessors and techniques; “consensus” discussion
  - Statistical integration process



## These *aren't* Assessment Centers

- ▶ Multiple-interview processes (panel or sequential)
- ▶ Paper-and-pencil test batteries
  - regardless of how scores are integrated
- ▶ Individual “clinical” assessments
- ▶ Single work sample tests
- ▶ Multiple measurement techniques without data integration

*nor is . . .*

- ▶ labeling a building the “Assessment Center”





*A Historical  
Perspective on  
Assessment  
Centers*

## Early Roots (“Pre-history”)

- ▶ Early psychometricians (Galton, Binet, Cattell) tended to use rather complex responses in their tests
- ▶ First tests for selection were of “performance” variety (manipulative, dexterity, mechanical, visual/perceptual)
  - Early “work samples”
- ▶ World War 1
  - *Efficiency* in screening/selection became critical
  - Group paper-and-pencil tests became dominant
- ▶ You know the testing story from there!



# German Officer Selection (1920's to 1942)

- ▶ Emphasized naturalistic and “Gestalt” measurement
  - Used complex job simulations as well as “tests”
  - Goal was to measure “leadership,” *not* separate abilities or skills
- ▶ 2-3 days long
- ▶ Assessment staff (psychologists, physicians, officers) prepared advisory report for superior officer
- ▶ No research built-in
- ▶ Multiple measurement techniques, multiple observers, complex behaviors



# British War Office Selection Boards (1942 . . .)

- ▶ Sir Andrew Thorne had observed German programs
- ▶ WOSBs replaced a crude and ineffective selection system for WWII officers
- ▶ Clear conception of leadership characteristics to be evaluated
  - Level of function
  - Group cohesiveness
  - Stability
- ▶ Psychiatric interviews, tests, many realistic group and individual simulations
- ▶ Psychologists & psychiatrists on assessment team—in charge was a senior officer who made final suitability decision



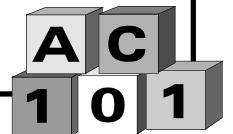
# British War Office Selection Boards (continued)

- ▶ Lots of research
  - criterion-related validity
  - incremental validity over previous method
- ▶ Spread to Australian & Canadian military with minor modifications



# British Civil Service Assessment (1945. . .)

- ▶ British Civil Service Commission—first non-military use
- ▶ Part of a multi-stage selection process
  - Screening tests and interviews
  - 2-3 days of assessment (CSSB)
  - Final Selection Board (FSB) interview and decision
- ▶ Criterion-related validity evidence
- ▶ See Anstey (1971) for follow-up



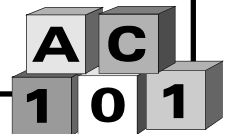
# Harvard Psychological Clinic Study (1938)

- ▶ Henry Murray's theory of personality
  - Sought understanding of life history, person "in totality"
- ▶ Studied 50 college-age subjects
- ▶ Methods relevant to later AC developments
  - Multiple measurement procedures
  - Grounded in observed behavior
  - Observations across different tasks and conditions
  - Multiple observers (5 judges)
  - Discussion to reduce rating errors of any single judge



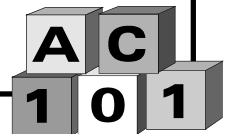
# OSS Program: World War II

- ▶ Goal: improved intelligence agent selection
- ▶ Extremely varied target jobs—and candidates
- ▶ Key players: Murray, McKinnon, Gardner
- ▶ Needed a practical program, quick!
  - Best attempts made at analysis of job requirements
  - Simulations developed as rough work samples
  - No time for pretesting
  - Changes made as experience gained
- ▶ 3 day program, candidates required to live a cover story throughout



## OSS Program (cont.)

- ▶ Used interviews, tests, situational exercises
  - Brook exercise
  - Wall exercise
  - Construction exercise (“behind the barn”)
  - Obstacle exercise
  - Group discussions
  - Map test
  - Stress interview
  
- ▶ A number of personality/behavioral variables were assessed
  
- ▶ Professional staff used (many leading psychologists)



## OSS Program (cont.)

- ▶ Rating process:
  - Staff made ratings on each candidate after each exercise
  - Periodic reviews and discussions during assessment process
  - Interviewer developed and presented preliminary report
  - Discussion to modify report
  - Rating of suitability, placement recommendations
  
- ▶ Spread quickly...over 7,000 candidates assessed
  
- ▶ *Assessment of Men* published in 1948 by OSS Assessment Staff
  - Some evidence of validity (later re-analyses showed a more positive picture)
  - Numerous suggestions for improvement



# AT&T Management Progress Study (1956 . . .)

- ▶ Designed and directed by Douglas Bray
  - Longitudinal study of manager development
  - Results not used operationally (!)
- ▶ Sample of new managers (all male)
  - 274 recent college graduates
  - 148 non-college graduates who had moved up from non-management jobs
- ▶ 25 characteristics of successful managers selected for study
  - Based upon research literature and staff judgments, not a formal job analysis
- ▶ Professionals as assessors (I/O and clinical psychologists)



# AT&T Management Progress Study

## Assessment Techniques

- ▶ Interview
- ▶ In-basket exercise
- ▶ Business game
- ▶ Leaderless group discussion (assigned role)
- ▶ Projective tests (TAT)
- ▶ Paper and pencil tests (cognitive and personality)
- ▶ Personal history questionnaire
- ▶ Autobiographical sketch



# AT&T Management Progress Study

## Evaluation of Participants

- ▶ Written reports/ratings after each exercise or test
- ▶ Multiple observers for LGD and business game
- ▶ Specialization of assessors by technique
- ▶ Peer ratings and rankings after group exercises
- ▶ Extensive consideration of each candidate
  - Presentation and discussion of all data
  - Independent ratings on each of the 25 characteristics
  - Discussion, with opportunity for rating adjustments
  - Rating profile of average scores
  - Two overall ratings: *would* and/or *should* make middle management within 10 years



# Michigan Bell Personnel Assessment Program (1958)

- ▶ First industrial application: Select 1st-line supervisors from craft population
- ▶ Staffed by internal company managers, not psychologists
  - Extensive training
  - Removed motivational and personality tests (kept cognitive)
  - Behavioral simulations played even larger role
- ▶ Dimensions based upon job analysis
- ▶ Focus upon behavior predicting behavior
- ▶ Standardized rating and consensus process
- ▶ Model spread rapidly throughout the Bell System



## Use expands slowly in the '60's

- ▶ Informal sharing of methods and results by AT&T staff
- ▶ Use spread to a small number of large organizations
  - IBM
  - Sears
  - Standard Oil (Ohio)
  - General Electric
  - J.C. Penney
- ▶ Bray & Grant (1966) article in *Psychological Monographs*
- ▶ By 1969, 12 organizations using assessment center method
  - Closely modeled on AT&T
  - Included research component
- ▶ 1969: two key conferences held



## Explosion in the '70's

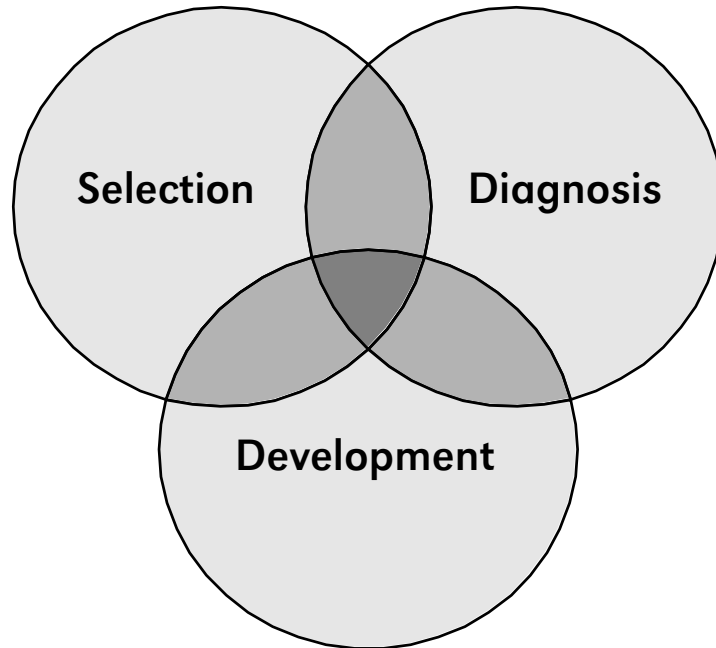
- ▶ 1970: Byham article in *Harvard Business Review*
- ▶ 1973: 1st International Congress on the Assessment Center Method
- ▶ Consulting firms established (DDI, ADI, etc.)
- ▶ 1975: first set of guidelines & ethical standards published
- ▶ By end of decade, over 1,000 organizations established AC programs
- ▶ Expansion of use:
  - Early identification of potential
  - Other job levels (mid- and upper management) and types (sales)
  - U.S. model spreads internationally





# *Assessment Center Design & Operation*

# Common Uses of Assessment Centers



- ▶ *Selection* and Promotion
  - Supervisors & managers
  - Self-directed team members
  - Sales
  
- ▶ *Diagnosis*
  - Training & development needs
  - Placement
  
- ▶ *Development*
  - Skill enhancement through simulations

# AC Design Depends upon Purpose!

	Selection	Diagnosis	Development
Participants	High potential employees or applicants	All interested employees	All interested employees
Target Position	Position opening to be filled	Current or future position	Current or future position
Dimensions	Few, global, traits	Many, specific, developable, distinct	Few, developable, very specific
Exercises	Few (3-5), generic	Many (6-8), moderate similarity to job	Many, work samples
Length	One half to one day	1.5 - 2 days	1.5 - 3 days
Key outcome	Overall rating	Dimension profile	Behavioral suggestions
Feedback to	Participant, Mgr up 2 levels	Participant, immediate Mgr.	Participant
Feedback type	Short, descriptive	Specific, diagnostic	Immediate verbal report, specific

Adapted from Thornton (1992)



# A Typical Assessment Center



Candidates participate in a series of exercises that simulate on-the-job situations



Trained assessors carefully observe and document the behaviors displayed by the participants. Each assessor observes each participant at least once



Assessors individually write evaluation reports, documenting their observations of each participant's performance



Assessors integrate the data through a consensus discussion process, led by the center administrator, who documents the ratings and decisions



Each participant receives objective performance information from the administrator or one of the assessors



# Assessor Tasks

- ▶ *Observe* participant behavior in simulation exercises
- ▶ *Record* observed behavior on prepared forms
- ▶ *Classify* observed behaviors into appropriate dimensions
- ▶ *Rate* dimensions based upon behavioral evidence
- ▶ *Share* ratings and behavioral evidence in the consensus meeting



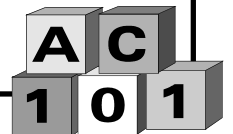
# Behavior!

- ▶ What a person actually says or does
- ▶ Observable and verifiable by others
- ▶ Behavior is *not*:
  - Judgmental conclusions
  - Feelings, opinions, or inferences
  - Vague generalizations
  - Statements of future actions
- ▶ A statement is not behavioral if one has to ask “How did he/she do that?”, “How do you know?”, or “What specifically did he/she say?” in order to understand what actually took place



## Why focus on behavior?

- ▶ Assessors rely on each others' observations to develop final evaluations
- ▶ Assessors must give clear descriptions of participant's actions
- ▶ Avoids judgmental statements
- ▶ Avoids misinterpretation
- ▶ Answers questions:
  - “How did participant do that?”
  - “Why do you say that?”
  - “What evidence do you have to support that conclusion?”



# Dimension

- ▶ Definition: A category of behavior associated with success or failure in a job, under which specific examples of behavior can be logically grouped and reliably classified
- ▶ identified through job analysis
- ▶ level of specificity must fit assessment purpose



# A Typical Dimension

***Planning and Organizing:*** Efficiently establishing a course of action for oneself and/or others in order to efficiently accomplish a specific goal. Properly assigning routine work and making appropriate use of resources.

Correctly sets priorities

- Coordinates the work of all involved parties
- Plans work in a logical and orderly manner
- Organizes and plans own actions and those of others
- Properly assigns routine work to subordinates
- Plans follow-up of routinely assigned items
- Sets specific dates for meetings, replies, actions, etc.
- Requests to be kept informed
- Uses calendar, develops “to-do” lists or tickler files in order to accomplish goals



## Sample scale for rating dimensions

- 5: *Much more than acceptable*: Significantly above criteria required for successful job performance
- 4: *More than acceptable*: Generally exceeds criteria relative to quality and quantity of behavior required for successful job performance
- 3: *Acceptable*: Meets criteria relative to quality and quantity of behavior required for successful job performance
- 2: *Less than acceptable*: Generally does not meet criteria relative to quality and quantity of behavior required for successful job performance
- 1: *Much less than acceptable*: Significantly below criteria required for successful job performance



# Types of simulation exercises

- ▶ In-basket
- ▶ Analysis
- ▶ Fact-finding
- ▶ Interaction
  - Subordinate
  - Peer
  - Customer
- ▶ Oral presentation
- ▶ Leaderless group discussion
  - Assigned roles or not
  - Competitive vs. cooperative
- ▶ Scheduling
- ▶ Sales call
- ▶ Production exercise



# Sample Assessor Role Matrix

Assessee	Assessor		
	A	B	C
1	LGD 1 Scheduling	LGD 2 In-basket	Interaction Fact-finding
2	LGD 1 Scheduling	LGD 2 In-basket	Interaction Fact-finding
3	Interaction Fact-finding	LGD 1 Scheduling	LGD 2 In-basket
4	Interaction Fact-finding	LGD 1 Scheduling	LGD 2 In-basket
5	LGD 2 In-basket	Interaction Fact-finding	LGD 1 Scheduling
6	LGD 2 In-basket	Interaction Fact-finding	LGD 1 Scheduling



# Sample Summary Matrix

Dimension	Simulation						
	LGD 1	LGD 2	Sched.	Fact-Finding	Interact.	In-basket	Summary
Oral Comm							
Written Comm							
Leadership							
Sensitivity							
Analysis							
Judgment							
Plng & Org							
Delegation							
Flexibility							
Initiative							



# Data Integration Options

- ▶ Group Discussion
  - Administrator role is critical
  - Leads to higher-quality assessor evidence—peer pressure
  - Beware of process losses!
  
- ▶ Statistical/Mechanical
  - May be more or less acceptable to organizational decision makers, depending on particular circumstances
  - Can be more effective than “clinical” model
  - Requires research base to develop formula
  
- ▶ Combination of both
  - Example: consensus on dimension profile, statistical rule to determine overall assessment rating



# Organizational Policy Issues

- ▶ Candidate nomination and/or prescreening
- ▶ Participant orientation
- ▶ Security of data
  - Who receives feedback?
- ▶ Combining assessment and other data (tests, interviews, performance appraisal, etc.) for decision-making
  - Serial (stagewise methods)
  - Parallel
- ▶ How long is data retained/considered “valid”?
  - Re-assessment policy
- ▶ Assessor/administrator selection & training



# Skill Practice!

- ▶ Behavior identification and classification
- ▶ Employee Interaction exercise
- ▶ LGD





*Validity  
Evidence and  
Research Issues*

# Validation of Assessment Centers

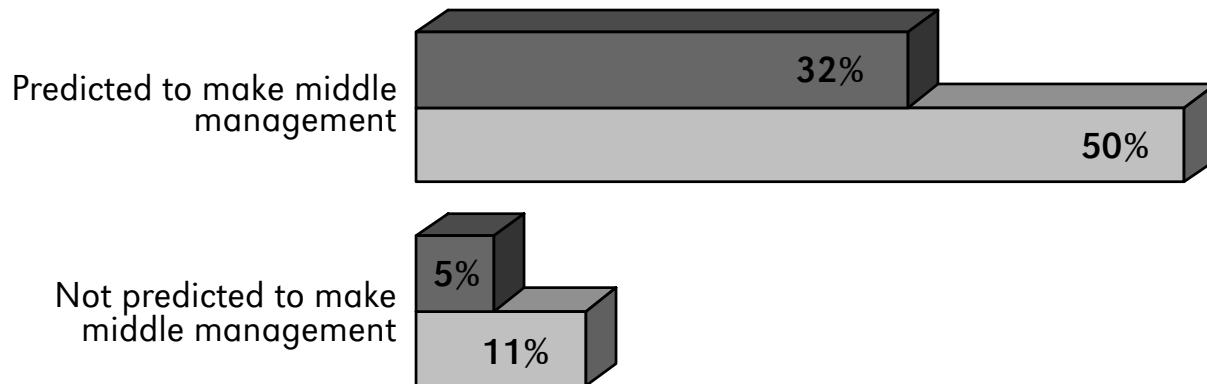
- ▶ Content-oriented
  - Especially appropriate for diagnostic/developmental centers
  - Importance of job analysis
- ▶ Criterion-oriented
  - Management Progress Study
  - Meta-analyses
- ▶ Construct issues
  - Why do assessment centers work?
  - What do they measure?



# AT&T Management Progress Study

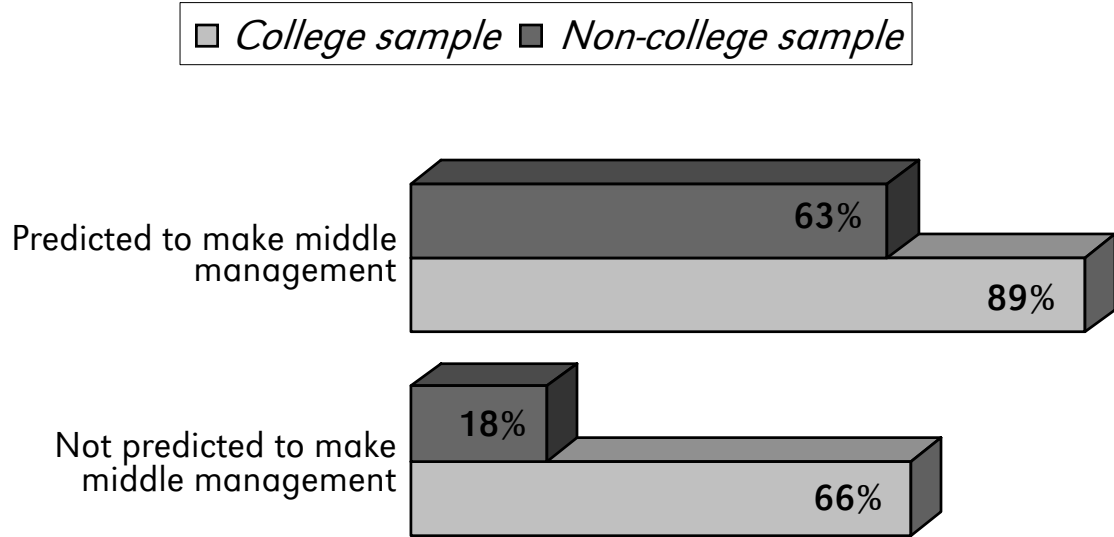
## Percentage attaining middle management in 1965

□ *College sample* ■ *Non-college sample*



# AT&T Management Progress Study

Percentage attaining middle management in 16 yrs.



## Meta-analysis results: AC validity

<b>Criterion</b>	<b>Estimated True Validity</b>
Performance: dimension ratings	.33
Performance: overall rating	.36
Potential rating	.53
Training performance	.35
Advancement/progression	.36

Source: Gaugler, et. al (1987)



## Research issues *in brief*

- ▶ Alternative explanations for validity:
  - Criterion contamination
  - Shared biases
  - Self-fulfilling prophecy
  
- ▶ What are we measuring?
  - Task versus exercise ratings
  - Relationships to other measures (incremental validity/utility)
    - § Background data
    - § Personality
    - § Cognitive ability
  
- ▶ Adverse impact (somewhat favorable, limited data)
  
- ▶ Little evidence for developmental “payoff” (very few research studies)





# *Developments in Assessment Methodology*

# Developments in Assessment Methodology

## “Drivers”

- ▶ Downsizing, flattening, restructuring, reengineering
  - Relentless emphasis on efficiency
  - Fewer promotional opportunities
  - Fewer internal assessment specialists
  - TQM, teams, empowerment
  - Jobs more complex at all levels
  
- ▶ Changing workforce demographics
  - Baby boomers
  - Diversity management
  - Values and expectations
  
- ▶ Continuing research in assessment
  
- ▶ New technological capabilities



# Developments in Assessment Methodology

## Assessment goals and participants

- ▶ Expansion to non-management populations
  - Example: “high-involvement” plant startups
- ▶ Shift in focus from selection to diagnosis/development
  - Individual employee career planning . . . person/job fit
  - Succession planning, job rotation
  - “Baselining” for organizational change efforts
- ▶ More “faddishness”?



# Developments in Assessment Methodology

## Assessment center design

- ▶ Shorter, more focused exercises
- ▶ “Low-fidelity” simulations
- ▶ Multiple-choice in-baskets
- ▶ “Immersion” techniques: “Day in the life” assessment centers
- ▶ Increased use of video technology
  - As an exercise “stimulus”
  - Video-based low-fidelity simulations
  - Capturing assessee behavior
- ▶ 360 degree feedback incorporated



# Developments in Assessment Methodology

## Assessment center design (cont.)

- ▶ Taking the “center” out of assessment centers
- ▶ Remote assessment
- ▶ Increasing structure for assessors
  - Checklists
  - Behaviorally-anchored scales
  - Expert systems
- ▶ Increasing automation
  - Electronic mail systems simulated
  - Exercise scoring and reporting
  - Statistical data integration
  - Computer generated assessment reports
  - Job analysis systems



# Developments in Assessment Methodology

## Increasing integration of HR systems

- ▶ Increasing use of AC concepts and methods:
  - Structured interviewing
  - Performance management
- ▶ Integrated selection systems
- ▶ HR systems incorporated around dimensions





# *Public Sector Considerations*

# Exercise selection in the public sector

- ▶ Greater caution required
  - 👍 Simulations are OK
  - 👍 Interviews are OK
  - 👎 Personality, projective and/or cognitive tests are likely to lead to appeals and/or litigation
  
- ▶ Candidate acceptance issues
  - Role of face validity
  - Role of perceived content validity
  - Degree of candidate sophistication re: testing issues
  - Special candidate groups/cautionary notes



# Data integration & scoring in the public sector

- ▶ Issues in use of traditional private sector approach
  - Clinical model . . . apparent inconsistencies/appeals
- ▶ Integration and scoring options
- ▶ Exercise level
- ▶ Dimensions across exercises
- ▶ Mechanical overall or partial clinical model
- ▶ Exercises *versus* dimensions



# Open Discussion: Q & A



# Basic Resources

- Anstey, E. (1971). The Civil Service Administrative Class: A follow-up of post-war entrants. *Occupational Psychology, 45*, 27-43.
- Bray, D. W., & Grant, D. L. (1966). The assessment center in the measurement of potential for business management. *Psychological Monographs, 80* (whole no. 625).
- Gaugler, B. B., Rosenthal, D. B., Thornton, G.C., & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology, 72*, 493-511.
- Cascio, W. F., & Silbey, V. (1979). Utility of the assessment center as a selection device. *Journal of Applied Psychology, 64*, 107-118.
- Howard, A., & Bray, D. W. (1988). *Managerial lives in transition: Advancing age and changing times*. New York: Guilford Press.
- Thornton, G. C. (1992). *Assessment centers in human resource management*. Reading, MA: Addison-Wesley.
- Thornton, G. C., & Byham, W. C. (1982). *Assessment centers and managerial performance*. New York: Academic Press.
- ◆ Also useful might be attendance at the *International Congress on the Assessment Center Method*. Held annually in the Spring; sponsored by Development Dimensions International. Call Cathy Nelson at DDI (412) 257-2277 for details.

