Developing gender-responsive assessments: recognizing difference, avoiding bias

Alexander M. Holsinger, ph.d.

Professor Criminal Justice & Criminology

University of Missouri-Kansas City

Assessment at the pre-sentence stage

- Historically assessment done via Pre-Sentence Investigation
  - Qualitative; Interview-based; Production of narrative
  - Useful, frames case notes, provides recommendations, foundation for future case management
  - Often unstructured, "weightiness" of info missed or misconstrued
  - Lacks objectivity

- What was often missing: An actuarial measure of risk propensity

Using actuarial risk assessment as part of pre-sent.

- Advantages when done well using effective tool:

  - Offers "group membership" that's tied to statistical probability of failure
  - Supports decision-making related to system-imposed sanctions
  - Facilitates case processing – supervision, other recommendations
  - Depending on content, can facilitate case management/planning
  - If dynamic can facilitate measurement of change over time

Using actuarial risk assessment as part of pre-sent. (cont.)

- Some challenges (to using them well):

  - Using the right tool (e.g., gender responsive set of tools)
  - Staff training and buy-in
    - often taken for granted
    - Underestimation re: what's necessary
  - Insuring tool(s) have validity (face, measurement, *statistical*)
  - Developing policy that maximizes use of, and support of, tools
  - Periodic, ongoing testing, maintaining an *assessment workgroup*
    - *They own/oversee the tool(s) administratively*

Developing gender-responsive assessments:
Some principles

- Recognizing the "set of predictors" for antisocial behavior may differ between men and women
  - Even if they do not differ regarding their substance, they may differ regarding form
    - "Antisocial peers" may come in different forms

- Recognizing gender/sex intersects with race, which intersects with justice systems in different ways
  - Even gender-responsive tools (any tools) should be examined for racial bias

What assessment development involves, broadly

- Collecting ***standardized*** dynamic (current) and historical information from 'everyone,' as part of a study
  - Each piece of info gathered, each item, is a potential '***predictor***'
- Observe a standardized follow-up period for each person in the study (e.g., 12, 18, 24 mos. from when info was gathered)
- Determine whether one or more failure criteria of interest occurred during follow-up period (recidivism); apply same methods of data collection for all

What assessment development involves, broadly (cont.)

- Analyze data – test for relationships
  - Relationships between all predictors (pairwise analysis)
  - Relationships between predictors and outcome(s)

- Conduct scaling tests for items revealing relationship w/outcome
  - Which items 'belong together' and form a coherent valid scale
  - Are they items that are reliable – easy to collect?

- Construct multiple scales (e.g., different by sex); analyses will reveal different scales that have different strengths/challenges

What assessment development involves, broadly (cont.)

- Determine how well scales differentiate between a "successful" case and an "unsuccessful" case
- Test scales prospectively
  - Conduct necessary training as necessary
  - Put them in place for a pilot period
  - Use them prospectively on a defined sample going forward
  - Subject sample to same follow-up/recidivism period
  - Conduct same analyses as those in development – see if scales hold up
  - Develop cutoff scoring to fit purposes of decision making
- Assuming things work out install them permanently
- Plan for ongoing tests of validity (~2 to 3 years)

Assuming different scales emerge by sex

- Conduct tests for racial or ethnic bias within each sex
  - Examine distribution of cases by risk category – make race/ethnicity comparison
  - Calculate and compare "false positive" and "false negative" rates
  - Determine if some groups are "over classified" (i.e., more likely to be incorrectly classified into a high risk category)