**Broad Components of Test Design**

1. Content standards are defined
   1. Constructs are defined
   2. Content specifications
   3. Cognitive tasks
2. Tests are developed
   1. Security issues:
      1. multiple forms within administration
      2. multiple forms over time across administrations
   2. Item specifications
   3. Test specifications
   4. Growth: multiple forms across grades/ages; content shifts
   5. Competition: multiple competing measures, different construct definitions, short forms
3. Tests are administered, scored, & scaled
   1. Raw scores are typically ineffective to meet the purposes of testing
   2. Scaling can improve norm-referenced or criterion-referenced inferences
   3. Scaled scores provide a metric that supports score comparability across forms
   4. Test developers create raw-to-scale-score conversion tables
   5. One or more forms may be used to define the scale; original form(s) is called the base or reference form to which new forms or alternate measures may be linked
4. Performance Standards are set
   1. Performance Level Descriptors
   2. Standard setting methods are evaluated
5. Decisions are made about the reporting score scale
   1. Range of scores
   2. Scale and precision
   3. Whether to truncate scores: LOSS & HOSS
   4. Mean and SD
   5. Performance categories defined through standard setting process
6. Scores are equated

Definition of Equating (Livingston, 2004)

“A score on the new form and a score on the reference form are equivalent in a group of test-takers if they represent the same relative position in the group” (p. 12).

* The equating relation is defined for a particular group
* “relative position” must be defined