# Use Assessment & Evaluation Strategies

Laura Schwarz, DNP, RN, CNE

Minnesota State University, Mankato

# Category 3
# Use Assessment & Evaluation Strategies

- 17% of Exam
- Areas A-L (12 areas)

# Assessment & Evaluation Strategies- Ways to Maximize Effectiveness

- Use literature to develop EBP A&E strategies
- Use a variety of approaches for A& E
- Implement EBP A&E *appropriate to learners & goals*
- Use A & E data to enhance T & L process
- "Timely, constructive and thoughtful feedback"
- Skill in design and use of A & E tools

Baumlein, 2015, p. 49

# A. Provide Input for the Development of Nursing Program Standards and Policies
# B. Enforce Nursing Program Standards

## Three Areas

- Admission
- Progression
- Graduation

# Admission-Typical Policies

- Clearly Defined and support program goals
- Reliable and valid with goals: prevent attrition & graduate those qualified to sit for licensure exams
- Graduated from program approved or accredited: foundation
- Minimum GPA
- Minimum standardized test score
- Can read/write English (ELL/ESL students)
- Official transcripts
- Completed relevant prerequisites

Baumlein, 2015, p. 49; Ellis, 2016

# Progression Policies

- Regulate progression
- Regulate fails or withdraws
- Timeline/limit
- Should be based on data
- Should be fair, justifiable, support program goals, consistent with institutional standard
- Appeals process for learners
- Determine reasons for attrition (& progression)

Baumlein, 2015; Ellis, 2016

# Graduation Policies

- Met program SLOs
- Completed  all coursework
- Minimum GPA
- Met financial obligations
- May implement high-stakes testing (use caution here)
  - See "The Fair Testing Imperative in Nursing Education" (Published by NLN in2012) https://mn.gov/boards/nursing/education/nln-fair-testing-imperative/

Baumlein, 2015, p. 53

# C. Use a Variety of Strategies to Assess and Evaluate Learning in These Domains:

**Blooms Taxonomy**-3 domains of learning; hierarchy

- **Cognitive:** Knowledge acquisition-least to most complex
- **Psychomotor:** Performance of manual or physical skills-lowest to highest
- **Affective:** Emotions or feelings-range from receiving to internalizing

Baumlein, 2015, p. 54- 55

# Cognitive

- Creating
- Evaluating
- Analyzing
- Applying
- Understanding
- Remembering

Baumlein, 2015, p. 55

# Psychomotor

- Naturalization
- Articulation
- Precision
- Manipulation
- Imitation

Baumlein, 2015, p. 55

# Affective

- Internalizing values
- Understanding the concept
- Conceptualizing and organizing
- Valuing
- Responding
- Receiving

Baumlein, 2015, p. 55

# D. Incorporate Current Research in Assessment & Evaluation Practices

# J. Use Evaluation Strategies that are Appropriate to the Learner & Learning Outcomes

- Responsible for using evidence-based A & E methods in classroom and clinical
- Maintain quality
- Need assessment in all 3 domains

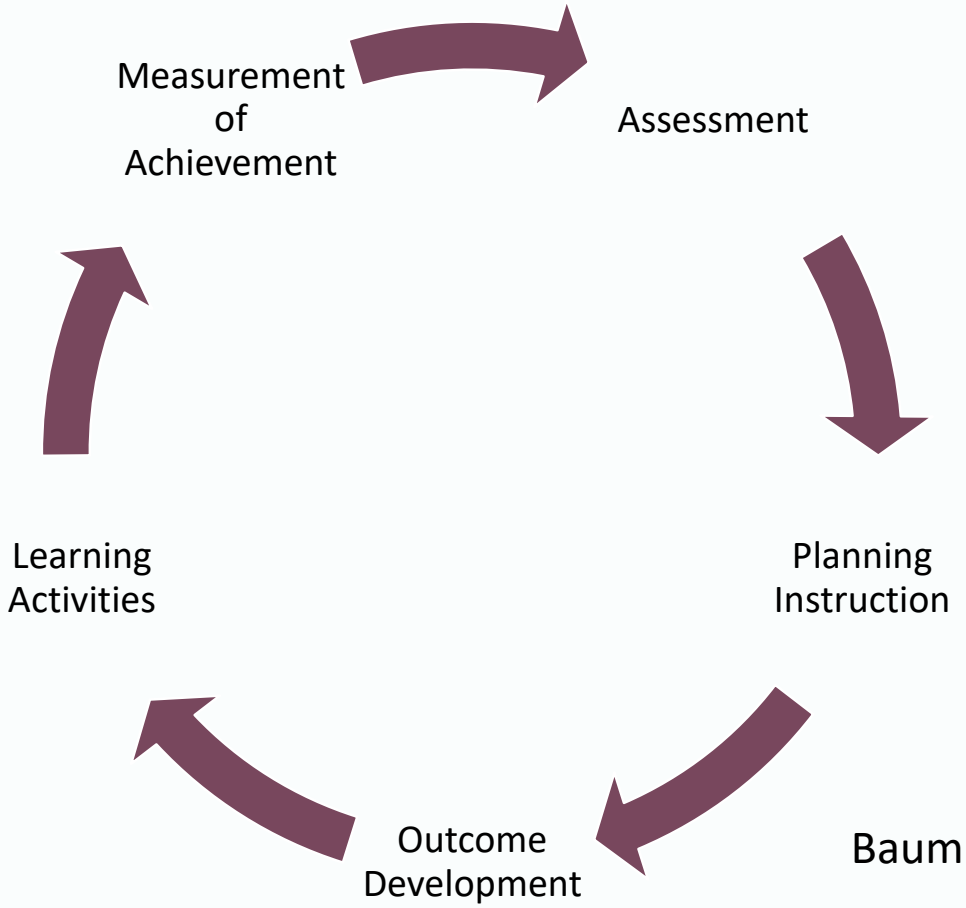
Baumlein, 2015

# Assessment

- "Measures provide information about students' abilities"
- Qualitative and quantitative data
- Ongoing throughout teaching-learning cycle
- Modify teaching based on results

Baumlein, 2015, p. 50

# Interaction of Assessment in Planning, Outcome Development, Learning Strategies, Measuring Achievement




Baumlein, 2015, p. 50, figure 3.1

# Measurement

- Not the same as assessment

"A process of assigning numbers to represent student performance or achievement" (Oermann & Gaverson, 2009 as cited in Baumlein, 2015)

- **Norm referencing:** How a student compares to others
- **Criterion referencing:** evaluation based on quality, AKA competency based measurement

Baumlein, 2015, p. 51

# Evaluation

- "Systematic appraisal of the quality of education"
- **Formative evaluation:** takes place throughout educational process, feedback about progress, goal to improve learning and clinical competency
- **Summative evaluation:** Takes place at the end of the educational process, "sums-up" outcomes

Baumlein, 2015, p.51

# F. Create Assessment Instruments to Evaluate Outcomes

- assessment  must have **alignment** with lesson/module and course SLO's
- assessment methods must be **valid;** measure what was learned
- Assessment methods must be **reliable/consistent;** produce comparable results whenever used
- Objectives must be **measurable**
- **SMART Objectives:** Specific, Measurable, Achievable, Relevant and Timebound

Baumlein, 2015

G. Use Assessment Instruments to Evaluate Outcomes

H. Implement Evaluation Strategies that are Appropriate to the Learner & Learning Outcomes

E. Analyze Available Resources for Learner Assessment & Evaluation

# Grading Rubrics

- Assess “performance of subjective assignments using specific, measurable criteria” Baumlein, 2015, p. 56
- Three necessary components: Reddy & Andrade (2010)
  - Evaluation criteria
  - Quality definitions
  - Scoring strategy

# Example

| Criteria | Met<br>5 points | Partially Met<br>3 Points | Unmet<br>0 points |
|---|---|---|---|
| Spelling, grammar, sentence format | Sentences are well organized, complete and free of spelling and grammar errors | Sentences are well organized and complete but some grammar and/or spelling errors | Sentences inadequate organization/structure, several grammar and/or spelling errors; run-on sentences |
| | | | |

Baumlein, 2015

# Clinical Evaluation

- More complex
- Use formative evaluation
- Clinical evaluation tools with specific, measurable criteria that speak to course-level SLOs
  - Observation
  - Oral communication
  - Written communication
  - Simulation
  - Self-evaluation-affective behaviors, self-reflection, self-evaluation

Baumlein, 2015

# Classroom Assessment

- Papers
- Debates
- A & V recordings
- Presentations
- Group projects
- Journals
- Simulation &Gaming
- Portfolios
- Reflection
- Role play
- Service learning
- Concept mapping

Baumlein, 2015, p. 58

# Developing Valid and Reliable Tests

- Skill, practice, time
- Decide the purpose
  - Readiness
  - Formative
  - summative

Baumlein, 2015

# Test Blueprints

- Connects content and outcomes to test items
- Develop before creating the exam
- Course/unit outcome & cognitive level (Blooms)
- Total number of items
- Weight/% in each area
- Level of difficulty should match learning level

Baumlein, 2015, p. 58

# Example

| Outcome/Content area | Percent of Exam in content area | Number of items at Knowledge Level | Number of items at comprehension level | Number of items at Application level |
|---|---|---|---|---|
| Respiratory | 25% | 5 | 10 | 10 |
| Cardiovascular | 25% | 5 | 10 | 10 |
| Neuro | 25% | 5 | 10 | 10 |
| GI | 15% | 3 | 6 | 6 |
| GU | 10% | 2 | 4 | 4 |
| Total | 100% | 20 | 40 | 40 |

# Test Construction & Item Writing

- Measure competency/mastery
- Many students, score quickly
- Difficult to write and take time
- Common types:  Multiple choice, T/F, matching, short answer, fill in the blank & ordered response

Baumlein, 2015

# Criteria for critical thinking test items

- Include rationales
- Are at the application or higher level
- Require high-level discrimination for selecting correct answer
- Require multilogical thinking (sequential reasoning) with more than one step in thinking to answer

Morrison, Nibert & Flick, 2006 as cited in Baumlein, 2015, p. 60

# I. Analyze Assessment & Evaluation Data

Three vital measures

– Difficulty level

– Item discrimination (of key and distractors)

– The reliability of the exam

Morrison et al, 2006 as cited in Baumlein, 2015

# Difficulty Level

- **Exam difficulty:** Review the mean, median & mode
- **Item difficulty (*p* value):** % of learners who correctly answered the item
  - Range reported as 0.00-1.00
  - A difficulty factor of .82 denotes that 82% of students correctly answered the item
  - Acceptable level item difficulty 30-90 %

(Morrison 2010 as cited in Baumlein, 2015)

# Item Discrimination

- Discrimination between learners who did and did not know the content
- High scorers correctly answer and low-scorers do not=discrimination (differentiates low & high scorers)
- Best indicator of test quality (Morrison 2010 as cited in Baumlein, 2015)

(Baumlein, 2015)

# Item Discrimination-PBCC

- **Point of Biserial Correlation (PBCC):** Statistic for item discrimination
- **Good discrimination:** PBCC will be highly positive for correct answer and negative for distractors
- .40 or greater=excellent discrimination, .30-.39 =good, .15-.29 satisfactory, <.15=low discrimination
- Maximized with the item difficulty is moderate (*P*=0.5)

Billings, 2016

# Example-Item Statistics

| Item Number | Difficulty level | Overall PBCC | Option | Response Proportion | PBCC |
|---|---|---|---|---|---|
| 1 | 0.69 | 0.42 | A | 0.03 | -0.46 |
| | | | B | 0.23 | -0.30 |
| | | | **C** | 0.69 | 0.42 |
| | | | D | 0.05 | -0.28 |
| 2 | 0.75 | 0.09 | **A** | 0.75 | 0.09 |
| | | | B | 0.08 | -0.28 |
| | | | C | 0.11 | 0.08 |
| | | | D | 0.06 | -0.26 |

Adapted from table 3.5, Baumlein, 2015, p. 62

# Reliability

- Consistency of exam results
- **Test-retest:** give same test to same person a second time-results are correlated
- **Parallel-form reliability:** two forms of same exam given to the same person-results are correlated
- **Internal consistency/reliability:** Kuder Richardson (KR-20) range -1.0-1.0; 1.0=perfect reliability, 0.0 lacks reliability, rarely see a negative KR-20. A KR-20 of .6-1.0 is acceptable

(Baumlein, 2015)

# K. Advise Learners Regarding Assessment & Evaluation Criteria

- Handbooks: Polices for class expectations, progression, testing, clinical expectations
- Syllabi: descriptions & expectations of assignments and evaluation methods
- Provide with blueprints, rubrics & expectations
- Clinical-provided objective, measurable performance criteria

(Baumlein, 2015, p. 64)

# L. Provide Timely, Constructive & Thoughtful Feedback to Learners

- Timely
- Specific
- Constructive
- Measurable
- Sensitive
- Balanced

# References


- Baumlein, G. B. (2015). Chapter 2 Use assessment and evaluation strategies. In L. Caputi (Ed). *Certified Nurse Educator review book: The official NLN guide to the CNE Exam* (49-67). Philadelphia: Wolters Kluwer.
- Billings. D. M. (2016) Chapter 24 Developing and using classroom tests: Multiple-choice and alternative format test items. In D. M. Billings & J. A. Halstead (Eds.) Teaching in nursing, 5th ed. (423-442). St. Louis, MO: Elsevier.
- Ellis, P. (2016). Chapter 26 Systematic program evaluation. In D. M. Billings & J. A. Halstead (Eds.) Teaching in nursing, 5th ed. (463-507). St. Louis, MO: Elsevier.
- Reddy, Y. M. & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education, 35*, 435-438