

Evaluation of a Question Bank-Only Generic Fundamentals Exam (GFE)

Dr. Stephanie Morrow, Human Factors Engineer
Sean Peters, Branch Chief

Human Factors and Reliability Branch
Office of Nuclear Regulatory Research
Nuclear Regulatory Commission



GFE Background



- Intended to test basic knowledge of nuclear power plant operations
 - Components
 - Reactor Theory
 - Thermodynamics
- Each exam - 50 multiple choice questions
 - 40 questions from GFE question bank
 - 5 questions from modified questions
 - 5 newly developed questions
- GFE question banks include ~ 2000 items
- GFE question banks are publicly available

Research Question

- Can an examination with all of the test items taken directly from the NRC's GFE question bank maintain test validity, reliability, and integrity?
 - **If so**, how many items should the GFE question bank contain to ensure examination integrity?

Research Approach

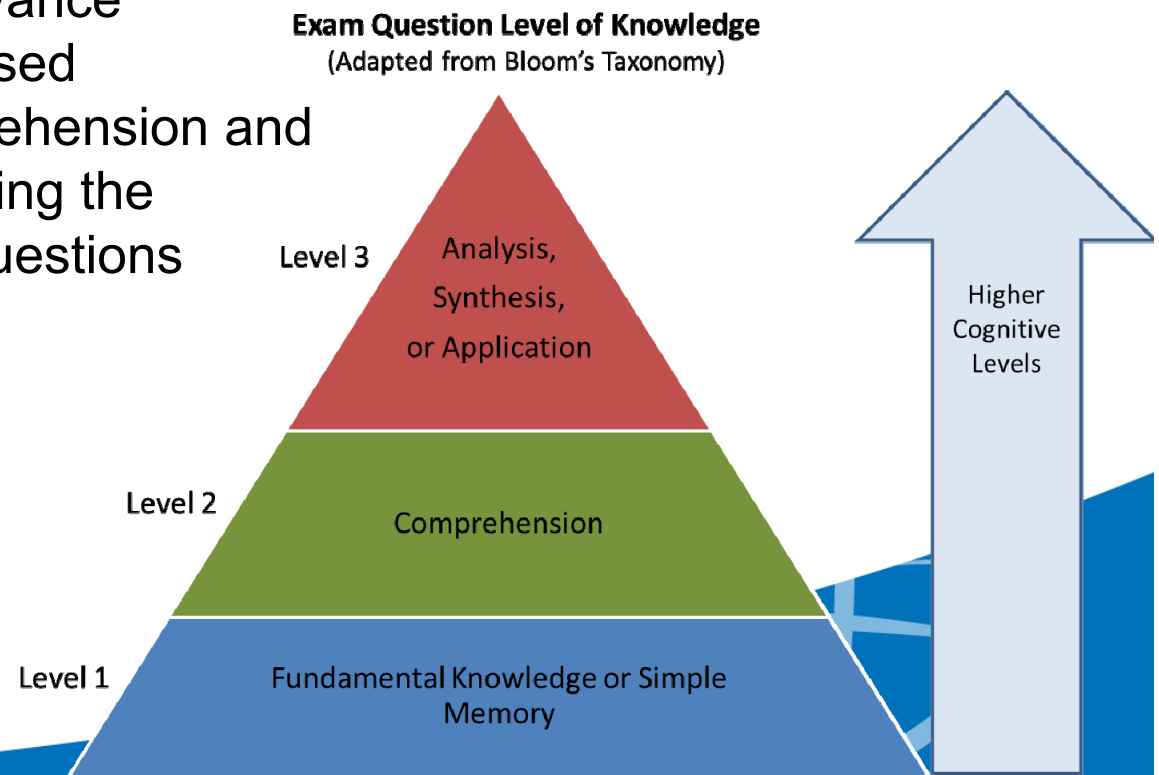
- Reviewed research and guidance on psychological testing
 - Standards for Educational and Psychological Testing
 - Articles, books, and technical reports
- Benchmarked with similar high-stakes testing program
 - Federal Aviation Administration (FAA) Airman Knowledge Test
- Analyzed GFE performance data
 - Trends examinee performance on new, modified, and question bank items
 - Comparisons examinee performance on new versus reused question bank items

Key Findings

- Using only questions from the publicly available GFE question bank is likely to challenge the validity and reliability of the exam.
- The FAA changed their testing procedures to limit availability of active question banks due to concerns about rote memorization and recall undermining the validity of the exam.
- GFE performance data indicates that examinees are significantly more likely to answer question bank items correctly, as compared to new or modified items – also suggesting effects due to memorization and recall.
- Question bank items may be systematically different from new or modified items:
 - Lower cognitive level
 - More predictable
 - Less difficult
 - Less discriminating

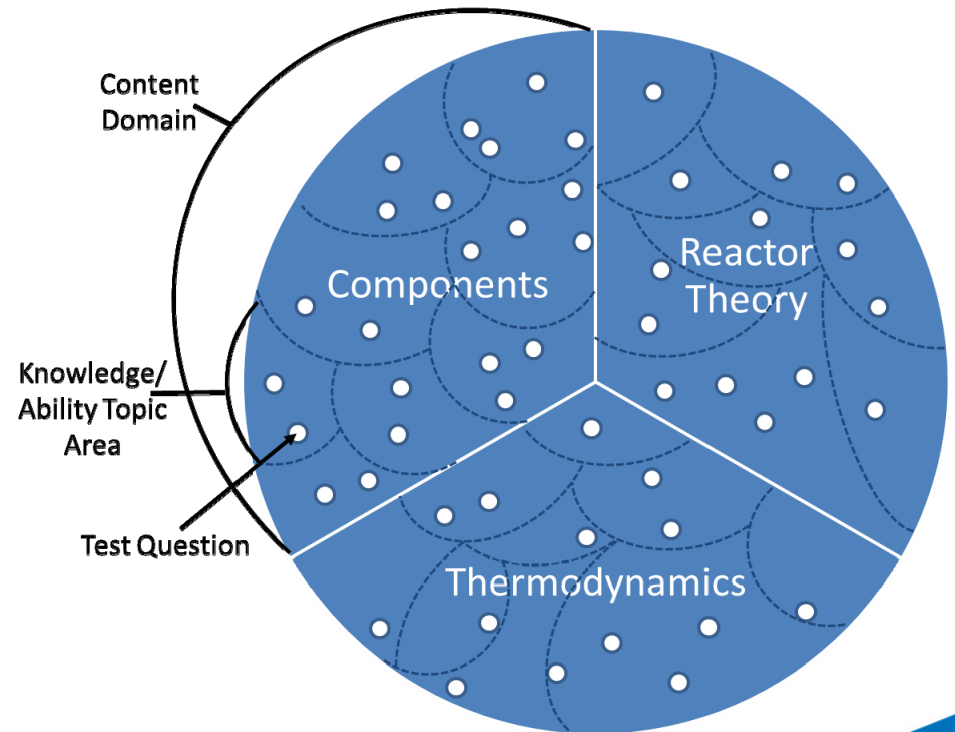
Lower Cognitive Level

- Level of knowledge – the range of mental demands required to answer a question
- Higher cognitive levels - require analysis, synthesis, or application of information to predict an outcome or solve a problem
- For questions known in advance examinees may answer based on recall rather than comprehension and analysis – effectively reducing the cognitive level of the test questions



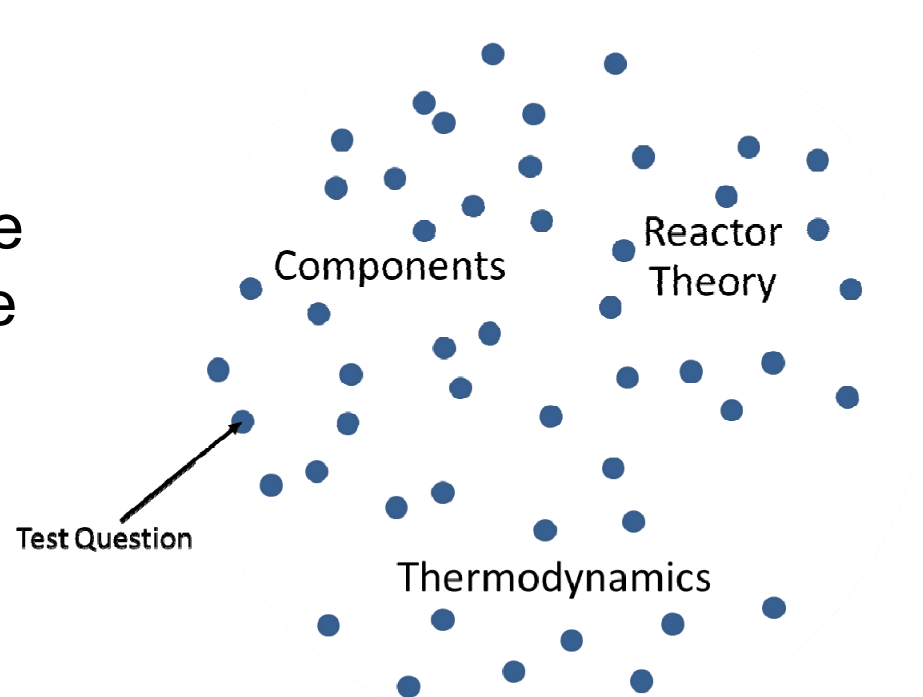
More Predictable

- GFEs are a sample of test questions from knowledge/ability topic areas within three content domains
- The standardized sampling plan allows us to infer that the test is valid – the performance on the test indicates mastery of the subject matter

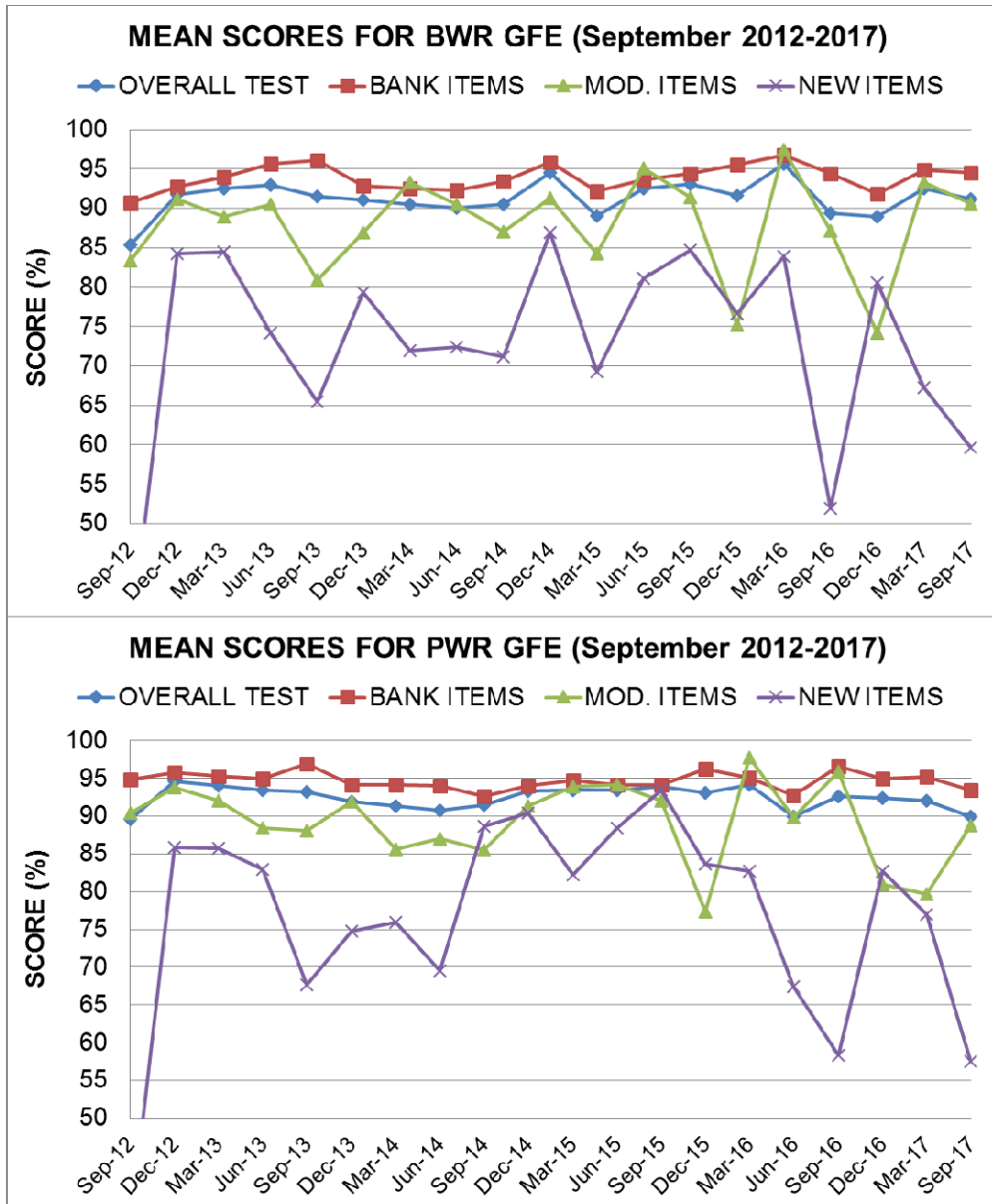


More Predictable

- Only using disclosed question bank items makes exams more predictable and encourages studying from previous tests rather than the larger body of knowledge the test is intended to sample
- Examinees only need to know the test questions, rather than the underlying content domains



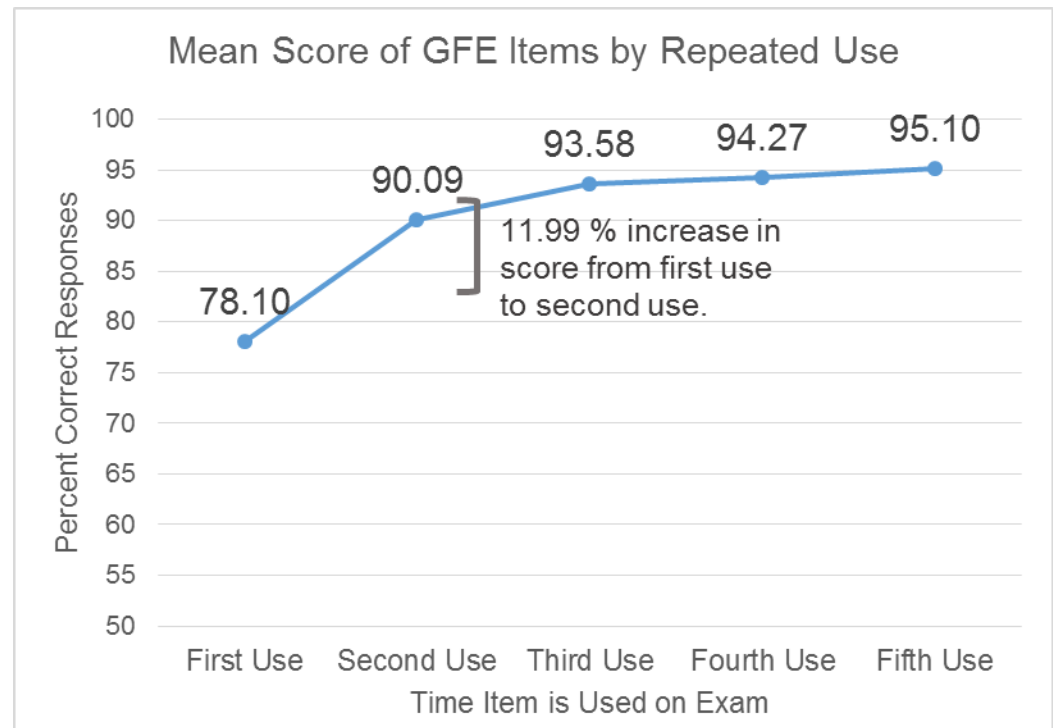
Less Difficult



- GFE performance trends show that question bank items are significantly less difficult than new or modified items
- Examinees are more likely to answer bank items correctly than new or modified items

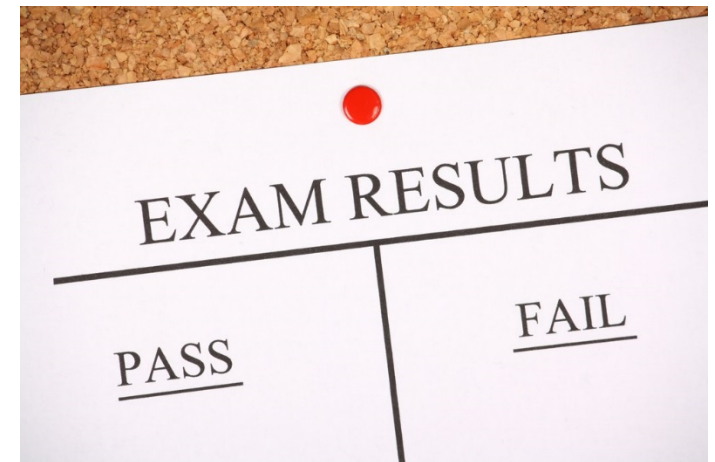
Less Difficult

- When a new item is used on the GFE about 78 percent of examinees answer it correctly
- When the same item is then added to the question bank and reused, about 90 percent of examinees answer it correctly
- This suggests that recall has a significant effect on exam performance
- A GFE composed of only question bank items would be less difficult



Less Discriminating

- An 80 percent score will pass the GFE
- We assume people who correctly answer 80 percent of items have demonstrated sufficient mastery of the knowledge being tested
- A less difficult and more predictable exam that does not engage higher cognitive levels is less likely to be able to discriminate between individuals who have and have not mastered the generic fundamentals of nuclear power operations



Conclusions

- A GFE derived solely from items in a disclosed question bank would decrease the validity and reliability of the exam and cannot assure examination integrity.
- Given that the GFE question banks are and have been disclosed to the public, the current size of the bank should not be used as a justification for deriving all exam items from the question bank.

For More Information

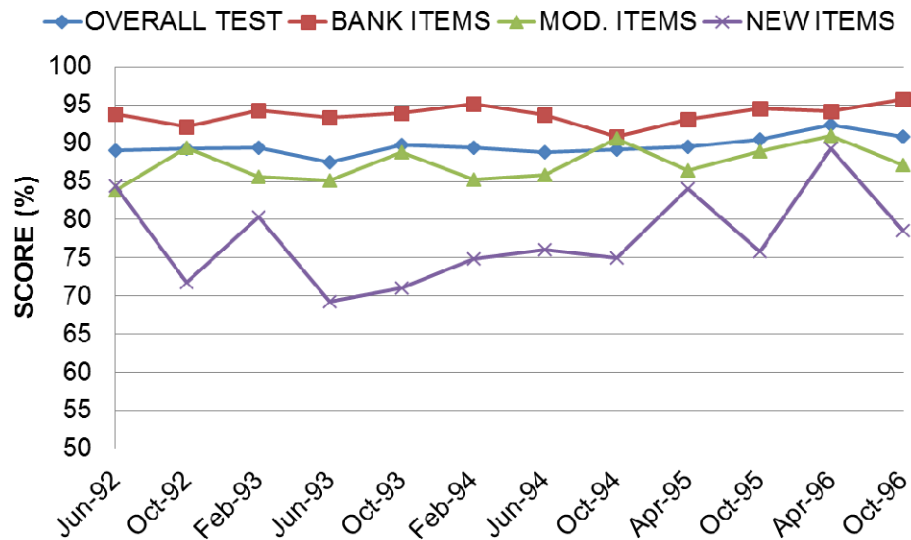


- The full technical report is available at:
 - ADAMS Accession Number ML
 - Link:

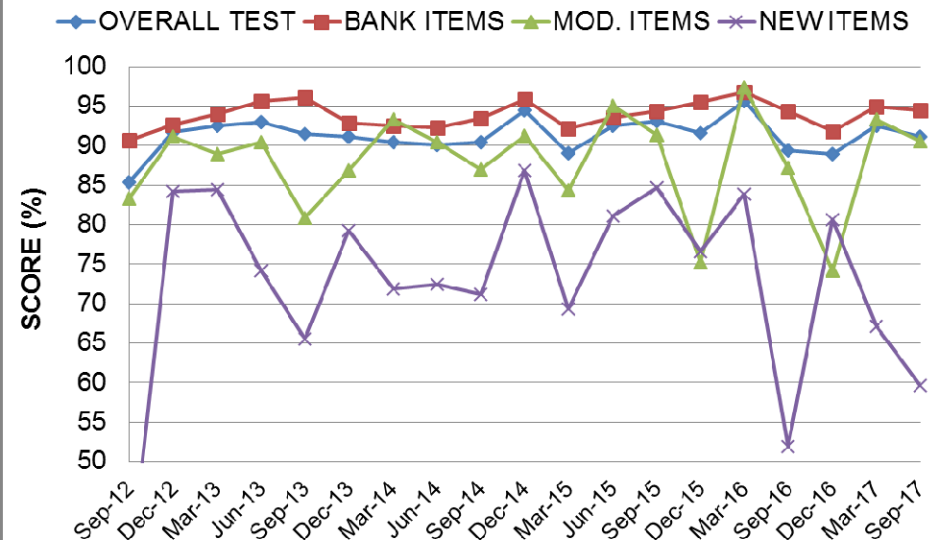
SUPPLEMENTAL SLIDES

Question banks doubled in size, but no change in trend from 1990's to 2010's

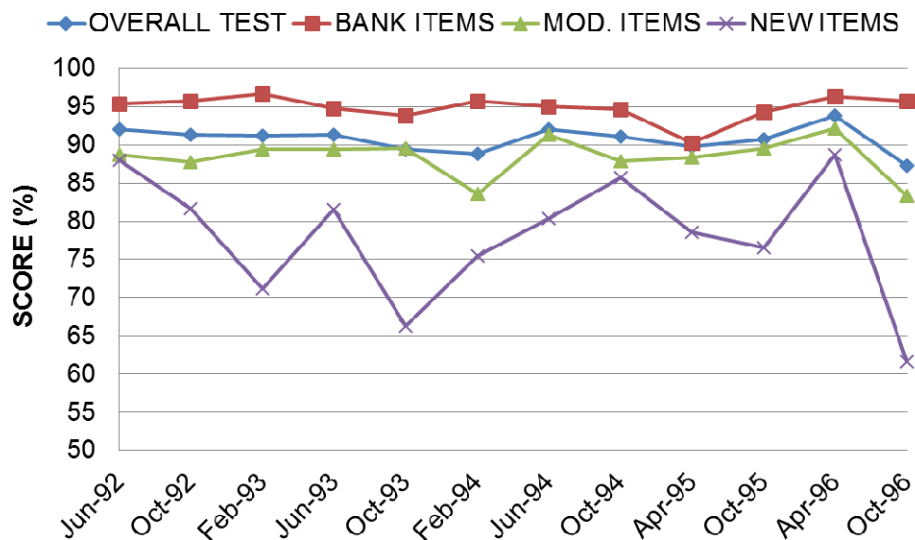
MEAN SCORES FOR BWR GFE (June 1992 - Oct. 1996)



MEAN SCORES FOR BWR GFE (September 2012-2017)



MEAN SCORES FOR PWR GFE (June 1992 - Oct. 1996)



MEAN SCORES FOR PWR GFE (September 2012-2017)

