

Small but Mighty: **Combining smaller** assessments to enhance overall reliability

F5 Networks (**Certified**

- Heidi Schreifels and Ken Salchow, DBA

Alpine Testing Solutions Alpine

- Jeff Kelley Ph.D., ICE-CCP & Jill R. van den Heuvel, Ph.D., ICE-CCP

Changing from one longer to multiple shorter assessments – Psychometric Perspectives









EXG2022



The need to evolve

Course List



EXG2022 OCT

OCTOBER 17-20, 2022 Savannah, Georgia



EXG2022



Adaptive Learning



credentialingexcellence.org = #ICEEXCHANGE22 () O O

Agile development



OCTOBER 17-20, 2022 Savannah, Georgia

Psychometrics of changing from one longer to multiple shorter assessments

Pros

certified

EXG2022

- Less seat time/expense per appt
 - Easier to schedule
 - Better fit for remote proctoring
- Focused study and preparation
- Easier to re-develop content
- Additional opportunities for feedback and recognition
- Equal or greater credentialing accuracy

Cons

- Higher expense per candidate?
- Greater chance of exam-level errors
- Lower reliability
- More psychometric & development expense
 - Exam management and maintenance
 - More item development

Reliability vs. Decision Consistency vs. Decision Accuracy

- Alpha Reliability (Internal Consistency)
 - Correlation between scores on multiple administrations
 - Affected by: Number of items, Candidate variability
- Decision Consistency (Livingston-Lewis)
 - Proportion of consistent classification decisions on multiple administrations
 - Affected by: Reliability, percent of candidates near the cut score
- Decision Accuracy (Livingston-Lewis)
 - Proportion of correct classification decisions
 - Affected by: Reliability, percent of candidates near the cut score



Goals for Reliability, Decision Consistency, and Decision Accuracy

- Alpha Reliability (Internal Consistency)
 - DeVellis (2012)
 - 0.65–0.70 = minimally acceptable
 - 0.70–0.80 = respectable
 - 0.80–0.90 = very good
 - Fabrey & Hartigan (2009) specific to certification exams
 - 0.80 or higher is acceptable
 - 0.90 or higher is hoped for
- Decision Consistency (Livingston-Lewis)
 - Subkoviak (1988) at least 0.85
- Decision Accuracy (Livingston-Lewis)
 - No referenced goal



F5 101 Current Compensatory Exam Specs

Candidates have 4 attempts to pass (initial attempt plus 3 retakes)

Form A

- Candidate count = 924
- 70 items
- Pass rate = 62.2%
- Overall reliability = 0.910
- Overall decision consistency = 0.881
- Overall decision accuracy = 0.916

Form B

- Candidate count = 883
- 70 items
- Pass rate = 62.3%
- Overall reliability = 0.901



- Overall decision consistency = 0.878
- Overall decision accuracy = 0.914



certified



OCTOBER 17-20, 2022 Savannah, Georgia

F5 201 Current Compensatory Exam Specs

Candidates have 4 attempts to pass (initial attempt plus 3 retakes)

Form A

- Candidate count = 1,416
- 70 items
- Pass rate = 88.5%
- Overall reliability = 0.862
- Overall decision consistency = 0.927

OCTOBER 17-20, 2022 Savannah, Georgia

• Overall decision accuracy = 0.950

Form B

- Candidate count = 1,377
- 70 items
- Pass rate = 87.1%
- Overall reliability = 0.814
- \bigcirc
- Overall decision consistency = 0.917
- Overall decision accuracy = 0.943



certified





But what about the new multiple shorter exams?







OCTOBER 17-20, 2022 Savannah, Georgia

F5 101 Analysis based on overall exam's 4 sections

- 4 sections contain 14 to 20 items
- Conjunctive model
- Form A
 - Actual reliability ranges 0.615 to 0.785
 - Spearman Brown estimates range 0.668 to 0.742
 - 3 sections actual is higher; 1 section estimate is higher
- Form B
 - Actual reliability ranges 0.650 to 0.752
 - Spearman Brown estimates range 0.645 to 0.722
 - 3 sections actual is higher; 1 section estimate is higher



F5 101 Analysis based on overall exam's 4 sections

Form A – Conjunctive Model – must pass each section/component

- Probability of a False Negative (should pass but fail; assuming qualified)
 - Full single longer exam (compensatory): 0.002%
 - Average of components: 0.020% (greater chance of failing any single component)
 - With 4 exams/components and 4 attempts: 0.079%
- Probability of a False Positive (should fail but pass; assuming unqualified)
 - Full single longer exam: 36.98%
 - Overall components: 14.78%
- How many people will fail when they should have passed?
 - Using multiple components 1 out of over 1,200 people
 - Using the longer overall exam 1 out of over 44,000 people



XG2022

F5 101 Analysis based on overall exam's 4 sections

Form B – Conjunctive Model – must pass each section/component

- Probability of a False Negative (should pass but fail; assuming qualified)
 - Full single longer exam (compensatory): 0.003%
 - Average of components: 0.032% (greater chance of failing any single component)
 - With 4 exams/components and 4 attempts: 0.127%
- Probability of a False Positive (should fail but pass; assuming unqualified)
 - Full single longer exam: 36.98%
 - Over all components: 15.82%
- How many people will fail when they should have passed?
 - Using multiple components 1 out of nearly 800 people
 - Using the longer overall exam 1 out of over 35,000 people



XG2022

F5 201 Analysis based on overall exam's 5 sections

- 5 sections contain 8 to 20 items
- Conjunctive model
- Form A
 - Actual reliability ranges 0.406 to 0.665
 - Spearman Brown estimates range 0.416 to 0.640
 - 3 sections actual is higher; 2 sections estimate is higher
- Form B
 - Actual reliability ranges 0.338 to 0.582
 - Spearman Brown estimates range 0.334 to 0.556
 - 3 sections actual is higher; 2 sections estimate is higher



F5 201 Analysis based on overall exam's 5 sections

Form A – Conjunctive Model – must pass each section/component

- Probability of a False Negative (should pass but fail; assuming qualified)
 - Full single longer exam (compensatory): 0.0001%
 - Average of components: 0.013% (greater chance of failing any single component)
 - With 5 exams/components and 4 attempts: 0.067%
- Probability of a False Positive (should fail but pass; assuming unqualified)
 - Full single longer exam: 59.65%
 - Over all components: 8.63%
- How many people will fail when they should have passed?
 - Using multiple components 1 out of nearly 1,500 people
 - Using the longer overall exam 1 out of over 1 million people



XG2022

F5 201 Analysis based on overall exam's 5 sections

Form B – Conjunctive Model – must pass each section/component

- Probability of a False Negative (should pass but fail; assuming qualified)
 - Full single longer exam (compensatory): 0.0002%
 - Average of components: 0.250% (greater chance of failing any single component)
 - With 5 exams/components and 4 attempts: 0.994%
- Probability of a False Positive (should fail but pass; assuming unqualified)
 - Full single longer exam: 59.24%
 - Over all components: 1.13%
- How many people will fail when they should have passed?
 - Using multiple components 1 out of approximately 100 people
 - Using the longer overall exam 1 out of nearly 480,000



XG2022

Psychometric & Programmatic Conclusions

- In a conjunctive model psychometricians don't need to worry about shorter assessments
- The conjunctive model creates more "hurdles"
 - Candidates are unable to compensate for a lack of knowledge in one component by having exceptional knowledge in another
 - Candidates are likely to be upset with the additional hurdles and higher risk for themselves.
- The probability of a false positive across all components is significantly lower
 - The probability of a false positive on any single component is higher than for the longer compensatory exam
- The certification should be based on the passing of all components
 - Lower stakes credentials could be based on the passing of a single component



Credits and Next Steps

- Brett Foley, Ph.D. conducted simulation analysis with this same concept and presented at CLEAR in 2021.
 - Foley, B. P. (2021, July). Increasing the accuracy of licensure decisions: On the benefits of more, shorter exams. Nicholasville, KY: Council on Licensure, Enforcement & Regulation. <u>https://www.clearhq.org/event-4433137</u>
- This analysis extends that work by using actual data from the current F5 Certified! certification program.
- Future work will evaluate the new/revised F5 Certified! certification program utilizing multiple smaller assessments.







OCTOBER 17-20, 2022 🔳 Savannah, Georgia