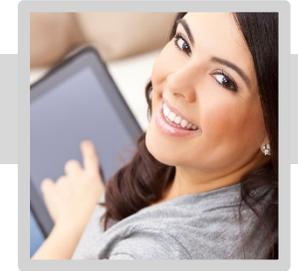


How Do You Know It's a Good Test?

Mr. Joe Fish
Questionmark



How do you know it's a good test?

Joe Fish

Product Director, Questionmark

Questionmark OnDemand for Government

Cloud-based Assessment Management System for US Government/Defense



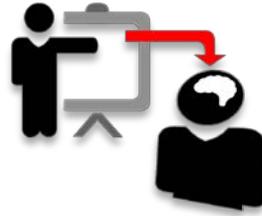
Security package now available

www.questionmark.com/trust/FedRAMP

What can we learn from assessments?



Did they
like it?



Did they
learn it?



Can they
do it?

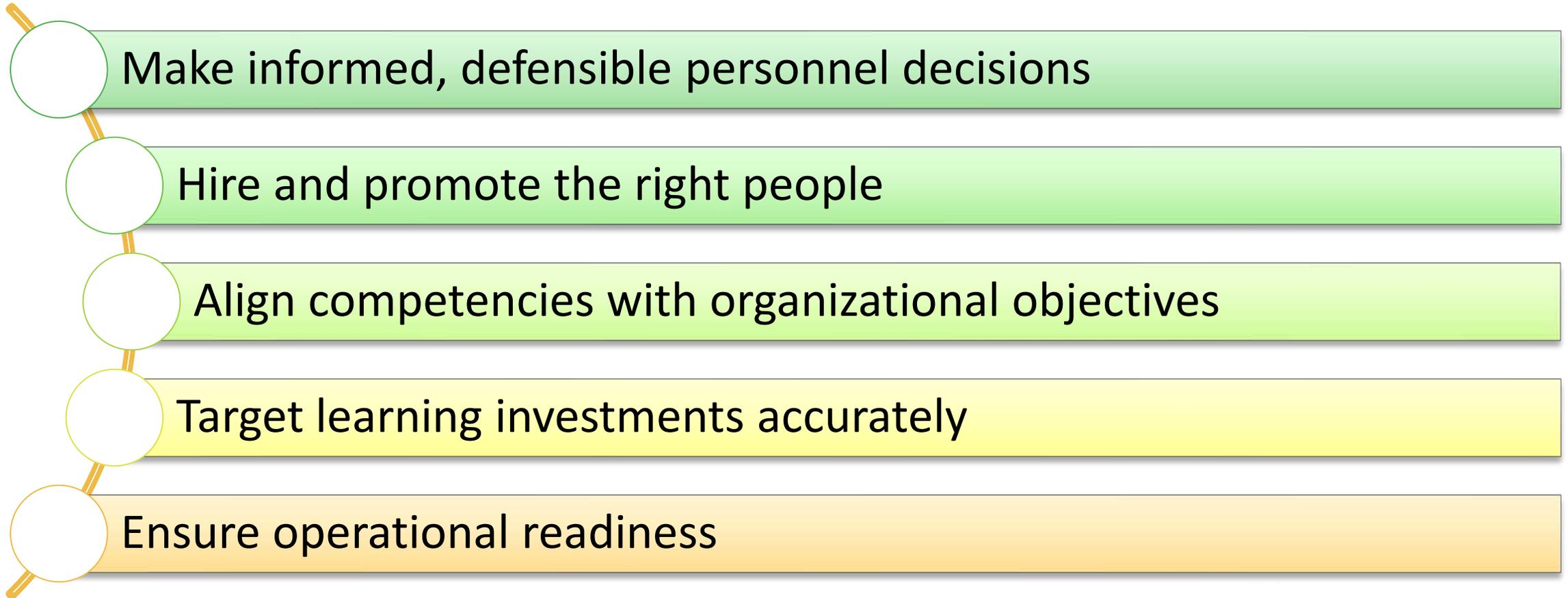


Can we
prove it?

Types of Assessments



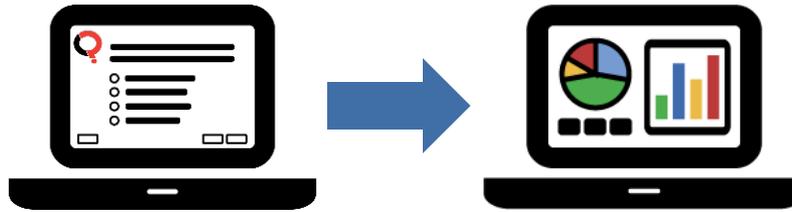
Assessment results help us in many ways...

- 
- Make informed, defensible personnel decisions
 - Hire and promote the right people
 - Align competencies with organizational objectives
 - Target learning investments accurately
 - Ensure operational readiness

When the stakes are high, assessment results must be **trustworthy** and **defensible**



Keys to getting trustworthy results....



Assessments need to be

- Valid
- Reliable
- Secure

Sound practices for
developing, delivering
and analysis

Technologies to support
sound assessment
practices

Reliability & Validity

Reliable:

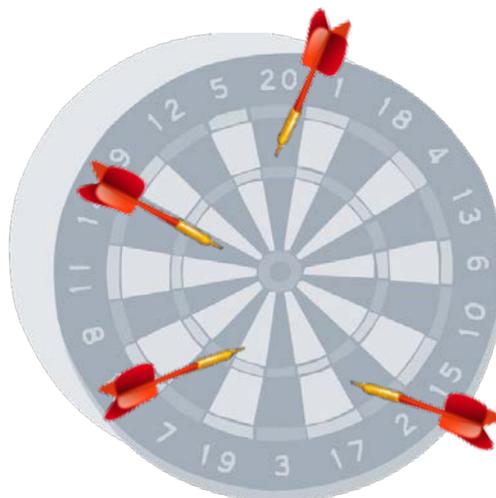
- Dependable, repeatable, consistent

Valid:

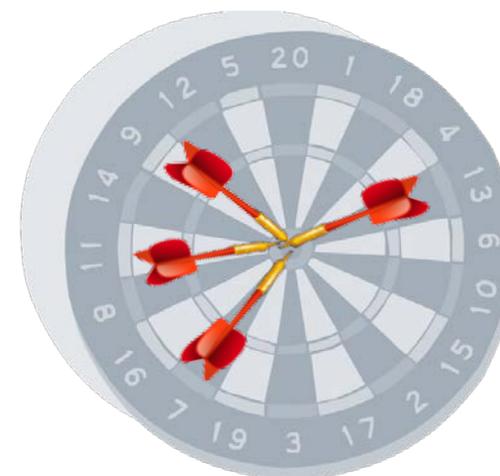
- Measures appropriate knowledge and skills



Reliable but not Valid

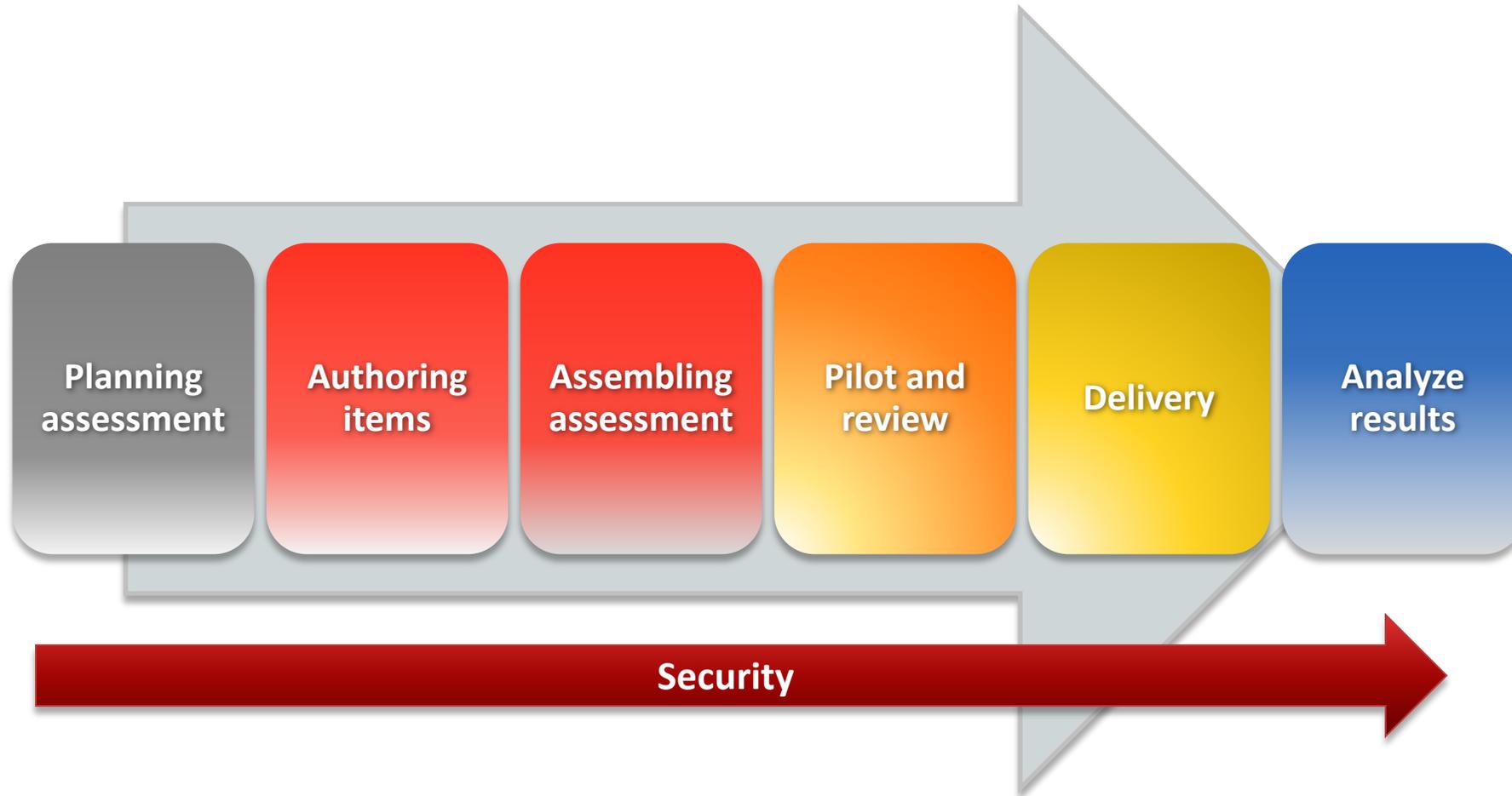


Not Reliable, not Valid



Reliable and Valid

Steps to getting trustworthy assessment results



Validate assessment content using Job Task Analysis

2 of 3

What is your role in the organization?

3 of 3

Answer questions about nursing.

	Applicability			Difficulty				
	Do Task	Supervise Task	N/A	Very Easy	Easy	Neither Easy or Difficult	Difficult	D
Administering medication	<input type="radio"/>	<input type="radio"/>						
Assessing patients	<input type="radio"/>	<input type="radio"/>						
Assisting patient	<input type="radio"/>	<input type="radio"/>						
Communicating with family members	<input type="radio"/>	<input type="radio"/>						
Cleaning surgical area	<input type="radio"/>	<input type="radio"/>						
Checking smother	<input type="radio"/>	<input type="radio"/>						

**Survey SMEs about
key tasks:**

How Difficult?
How Important?
How Frequent?
How Critical?

Analyze JTA results to “blue print” your test

Electrical Skills Test - Blueprint	% of test	# items	Skill level	Criticality
Section 1: AC Motors	45%	34		
1.a - Lorem ipsum dolor		21	3	3
1.b - Duis aute irure		8	3	2
1.c - Excepteur sint occaecat		5	2	1
Section 2: High-voltage switchgear	25%	19		
2.a - Neque porro quisquam		12	2	3
2.b - Modi tempora incidunt		7	3	2
Section 3: Synchronous Motors	30%	23		
3.a - Totam rem aperiam		11	2	2
3.b - Eaque ipsa quae		12	2	2

How many questions
for each topic?

Enable SMEs write test items collaboratively, securely

People ▾ Authoring ▾ Administration ▾ Reporting Analytics ▾ Results

Home » Role & Permission Management » Add a role

Users		Roles	
<input type="checkbox"/>	400467	<input type="checkbox"/>	Assessment Author
<input checked="" type="checkbox"/>	bjm5757	<input checked="" type="checkbox"/>	Item Author
<input checked="" type="checkbox"/>	bjm-test2@outlook.com		
<input type="checkbox"/>	Brian McNamara		
<input type="checkbox"/>	Brian.McNamara		
<input type="checkbox"/>	brian@questionmark.com		

1 - 9 of 9 items

1 - 2 of 2 items

Topics include all sub-items

- Automotive
- Chemistry
- Course Evaluation
- demo for export
- Demonstration
- Electrical Skills
 - AC Motors
 - High Voltage Switchgear
 - Synchronous Motors
- Maps - Political

Set role-based access to items



Blueprint guides authors on item selection

Select questions to include

- All questions from topic
 - Include sub-topics
 - As single questions
- Randomly select from topic
 - Include sub-topics
 - Avoid previously delivered
- Single question

Select a topic

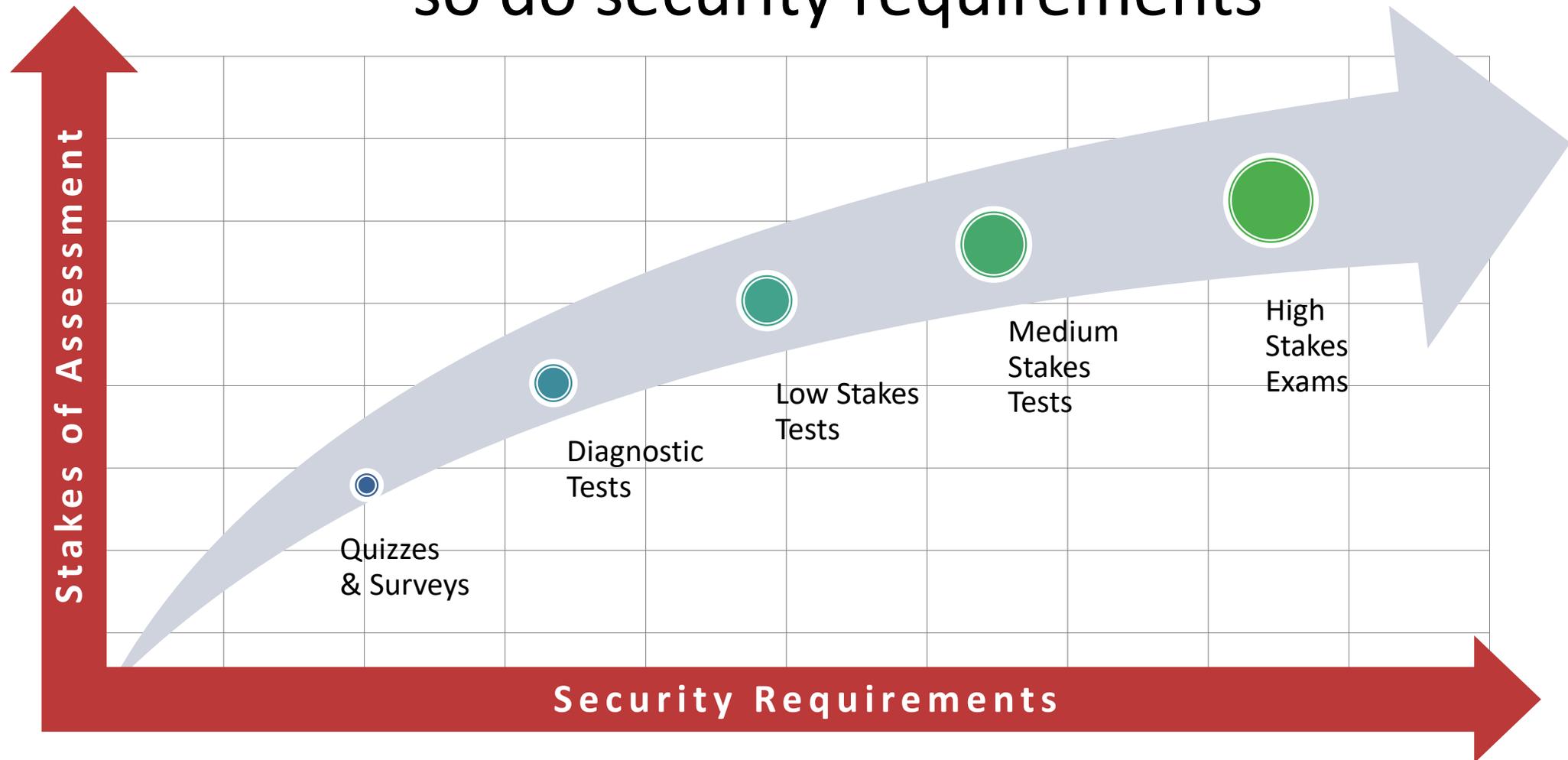
- San Antonio
- Skills Test
 - AC Motors
 - High Voltage Switchgear
 - Synchronous Motors
- Test set of questions
- The FDA Process for Approving Generic Drugs

Selected questions - drag to set question order

+ 5 random question(s) from topic 'Skills Test\High Voltage Switchgear' including subtopics	x
+ 7 random question(s) from topic 'Skills Test\AC Motors' including subtopics (Avoid previously delivered)	x
+ 6 random question(s) from topic 'Skills Test\Synchronous Motors' including subtopics	x

Clear

When assessment stakes increase, so do security requirements



Key threats to address



Cheating



Content Theft



Impersonation



Combating and Mitigating threats



Secure authentication &
strong password
management



Secure apps for locked-
down test delivery and
monitoring



Proctoring - Live Online,
On-site (in person),
Record-and-review

Analyzing reliability of assessment results



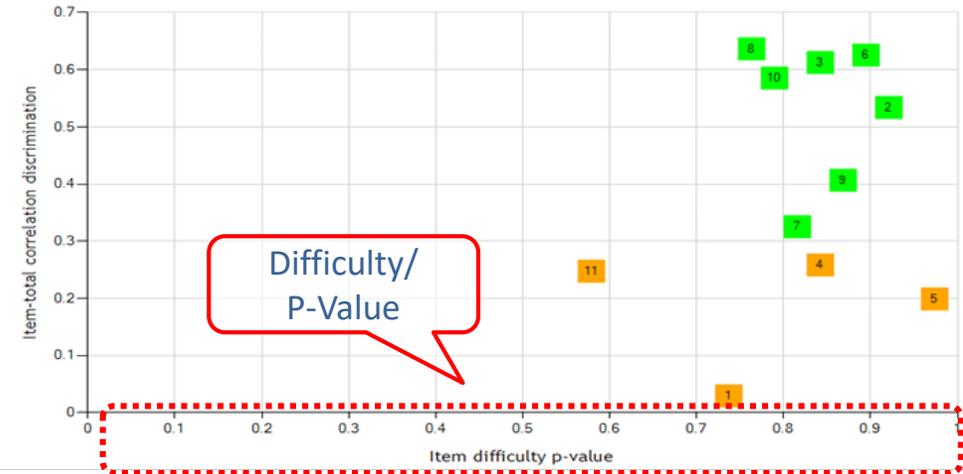
Are the questions too easy? Too hard?
(Should I care?)

Are the results from individual questions **consistent** with the overall test results?



Analyze Item Statistics: Difficulty Index

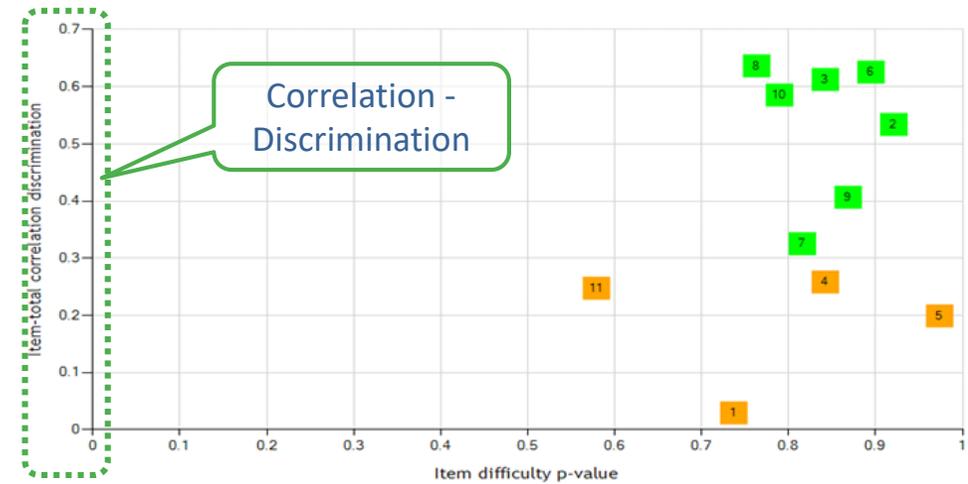
- ▶ **Measures % correct responses**
- ▶ **Expressed as number: 0.0 - 1.0**
 - Higher = Easier
 - Lower = Harder
- ▶ **Flags possible problems:**
 - Too high: Too easy!
 - Too low: Too difficult?
Confusing/ambiguous? Miscoded?



Question description	Revision	Topic	Item difficulty p-value	Item-total correlation discrimination
A multiple choice question must have at least four choices.	2	Item Writing Best Practices	0.737	0.03
In test questions, it is good practice to include implausible, nonsense, or humorous, incorrect choices.	2	Item Writing Best Practices	0.921	0.534
Phrasing a question as, Which of the following is not ... is considered best practice.	3	Item Writing Best Practices	0.842	0.613

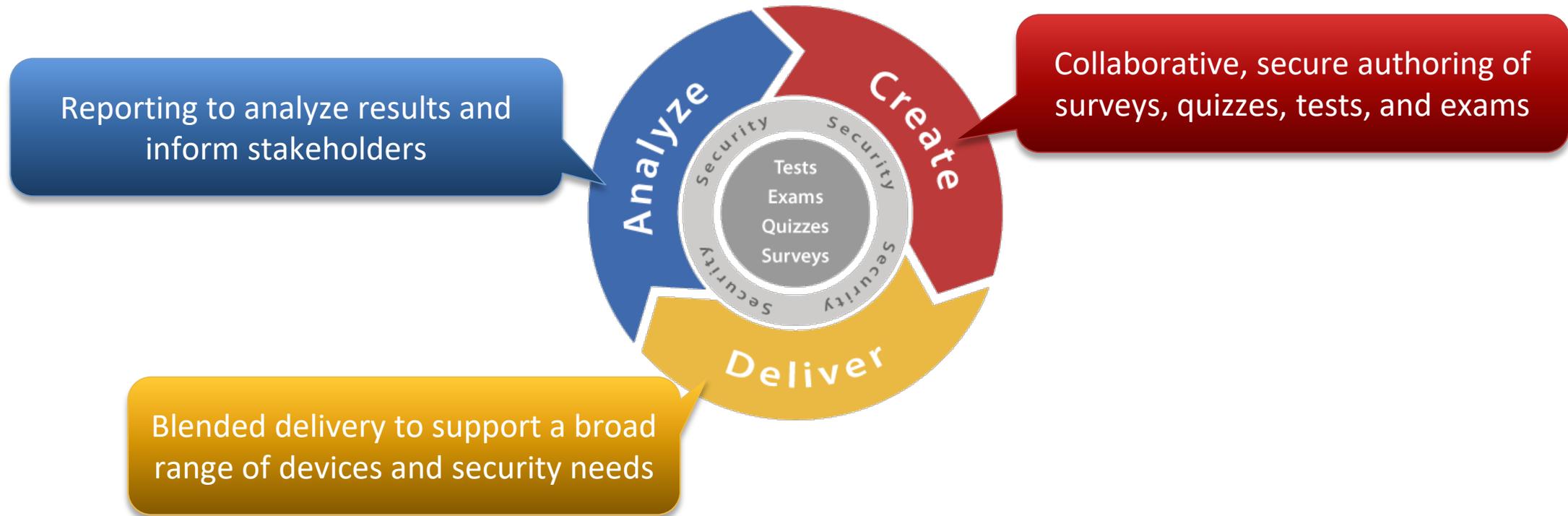
Analyze Item Statistics: Correlation Discrimination

- ▶ Measures correlation between **item score** and **test score**
- ▶ Are the item results consistent with assessment results?
- ▶ Is this item a good/poor predictor of overall assessment results?



Question description	Revision	Topic	Item difficulty p-value	Item-total correlation discrimination
A multiple choice question must have at least four choices.	2	Item Writing Best Practices	0.737	0.03
In test questions, it is good practice to include implausible, nonsense, or humorous, incorrect choices.	2	Item Writing Best Practices	0.921	0.534
Phrasing a question as, Which of the following is not ... is considered best practice.	3	Item Writing Best Practices	0.842	0.613

Assessment Management System



Enables and supports sound assessment practices



Learn more...

Come and visit us at our booth!



Follow us...

