CU-101.7 Intro to Beta Testing with Alan Mead

0:00
Welcome to Certiverse University and welcome to this video Introduction to beta testing.

0:04
My name is Alan Mead and I'm the chief psychometrician for Certiverse.

0:07
In this video, we'll cover what is beta testing, how to select A beta form, and how to plan and conduct a beta test.

0:14
Let's get started.

0:17
Despite the careful steps that you took during item writing, a percentage of your items just won't work well in practice.

0:24
Some textbooks suggest that you be prepared to write as much as double the amount of items you want to end up with to allow margin for items to fail and still have enough good items to compose forms.

0:35
I usually recommend a margin of 10 to 30% because items authored using best practices on the Certiverse platform have a median failure rate of just 8%.

0:45
Of course that means half our programs have a higher failure rate that if you used a lighter weight development process, your rate might also be higher.

0:52
Beta testing or also called pilot testing refers to item tryouts with real candidates to find these poor quality items.

0:59
So that means administering the items to a representative sample of candidates under operational conditions or as close as you can get to that.

1:07
Usually beta candidates will receive a pass fail decision but at a later time after say the end of the month.

1:16
This is called delayed reporting.

1:18
That delay gives your program time to analyze the candidate responses and this item analysis allows you to detect and eliminate poorly performing items.

1:28
You can then select from the good items to compose operational forms that are more reliable and more fair.

1:35
The candidates then get rescored out to these operational forms to receive their pass fail decisions.

1:41
So beta testing allows you to score all candidates on higher quality exams by eliminating problems before you begin issuing scores.

1:50
And for this reason, beta testing is considered one of the most critical steps in developing an exam.

1:56
I should also note that there's no one right way to do a beta test.

2:01
I'll describe a method that works well for a lot of clients, but your situation may be unique and you might need to consult a psychometrician.

2:11
Having just said that, a beta test is critical.

2:13
What happens if you skip the step?

2:15
Because sometimes it's not possible.

2:17
If I go to the DMV and take the driver's license exam, they can't say thanks.

2:21
We'll let you know in a month whether you've passed.

2:23
In fact, regulatory exams like the DMV driver's license exam often cannot conduct beta exams.

2:28
There are other circumstances where it's easy to find a sample inconvenient.

2:32
For example, if you're building a personality scale, you could use a sample of random Internet people.

2:37
But that strategy will not work for certification exams.

2:39
So in these cases you could do what I'll call an operational beta, which basically means ignoring the probability that some items are poor quality and just administering the exam to a sample of 50 or 100 or 200 real

2:51
but Guinea pig candidates who get scored and receive a pass fail decision based on those scores, even if they're imperfect.

2:57
Once you have your sample of 50, 100 or 200 candidates, you perform item analysis and quickly revise and republish the forms to remove the poorly performing items.

3:09
Whether it's because you can't or won't perform a beta, there are some consequences of not performing a beta.

3:15
A percentage of your carefully constructed items just won't work and may confuse candidates, cause complaints, or even tarnish the reputation of your exam.

3:24
The pass fail decisions for some candidates will be less proper than they otherwise might have been.

3:31
You will have some issues in creating forms that are both reliable and valid.

3:35
And then finally, you have some future limitations which might worry a psychometrician more than you.

3:41
But these limitations affect the ability of psychometricians to improve your exam.

3:47
So here are the steps in planning a beta test.

3:49
First, you create one exam roughly adhering to the blueprint and roughly doubling. For example for exams with 50 scored items

3:56
we've had good luck with 120 item beta exams.

3:59
If your exam has 100 or more items, then you're probably looking at an exam length

4:04
more like 150 items.

4:06
You'll need additional administration time, but probably not double.

4:10
I strongly recommend randomizing item position for each examinee so that any fatigue effects are distributed across items.

4:17
Then you recruit 50 or more candidates representative of the target population. "Or more"

4:24
could be up to 200 and having 200 would be great, but it's often not feasible.

4:28
And representation effects some outcomes, but any reasonably representative sample adds value in this step.

4:35
Next, you administer this form under operational conditions.

4:38
Again, perfect fidelity is not absolutely essential.

4:42
Then you perform item analysis, and we'll discuss how to do this in the next video.

4:46
Then you select operational forms.

4:48
Usually you want 2 forms for retesting purposes.

4:51
There are some details and we'll talk about this in a future video.

4:55

Finally, when you have your operational forms, you're ready to rescore your beta candidates onto those operational forms and give them their pass Failed decisions To select A beta form, determine the total beta form length by doubling the total operational exam.

5:13
For example, a 60 item exam with 50 scored and 10 pretest item would imply A 120 item Beta exam.

5:21
Select items proportionate to the blueprint weights.

5:24
If domain 1.1 needs 15% of 120, that's 18 items.

5:28
So you would select 18 items from domain 1.1, have SMEs review the form and identify item enemies, which are items that give each other away or otherwise shouldn't be administered together, and try to eliminate item enemies by swapping out one of the enemy items.

5:45
This may not always be entirely possible with a long beta form.

5:50
Note I haven't said how you select the items.

5:52
At this stage, all your items have passed through several layers of review, but I don't think you have any basis for preferring one item over the other, so I'd select at random and let the SMEs adjust the forms as needed.

6:05
While random may sound bad, it ensures that the beta form is representative of the pool, and the SME review can swap out any items that they dislike.

6:13
As an alternative, you could let the SMEs pick each item, but I think this is likely to be much less efficient.

6:23
Recruiting candidates for your beta exam could be its own course, and it varies quite a bit depending on the exam program.

6:29
If you have a large community of professionals, recruiting may be fairly easy.

6:33
In other instances, it can be a difficult marketing challenge.

6:36

Beta candidates might be real candidates who are queued up to and waiting for the chance to certify, but often you recruit from sources like learners, experts, supporters, and early adopters.

6:49

It's ideal if you do not include SMEs who have pre knowledge of the items, like the item authors and reviewers, but practically I'm not sure that's a huge issue if you don't have a very big beta sample.

7:02

Representativeness is often hard to achieve in practice, so the good news is that beta testing will successfully identify flawed items even in unrepresentative samples.

7:12

But you may not have as good a handle on the difficulty of your exam if you have an unrepresentative sample.

7:17

So, for example, it's very hard to estimate the eventual operational pass rate from beta participants.

7:22

Often beta pass rates are low, and this doesn't always translate into low operational pass rates.

7:27

So do the best you can and take representativeness into account when interpreting the results.

7:33

You should ensure that the beta exam administration is as close as possible to operational conditions, including proctoring, and this is a good dress rehearsal for operational testing of the exam, and it's often an early milestone for marketing your exam.

7:51

Arguably the last step of beta testing is rescoring those beta candidates onto operational forms and awarding pass fail decisions.

7:58

But there are a few things that need to be done 1st and those will be the subject of the next few videos.

8:03

So we'll go into details in those videos.

8:06

But the next steps are performing an item analysis to identify bad items and possibly performing an option analysis to guide your SMEs in fixing the bad items.

8:17
Then selecting scored and unscored items for the operational forms.

8:21
You probably want 2 forms for retesting purposes.

8:24
Then setting the standard, gathering the data that will allow you to determine the passing score for your forms, configuring the scoring and reporting, and then finally rescoring the beta exam.

8:34
And of course, that's also at a point where you could launch your exam.

8:39
So I hope you've enjoyed this video on beta testing, another critical step in developing fair and reliable exams.

8:45
Hopefully you have a better idea how you'd perform beta testing.

8:48
In the next video, we'll address the next step, which is item analysis.

8:52
That's it for this video.

8:53
Thanks for watching.

8:54
We'll see you in the next.