

CU-101.5

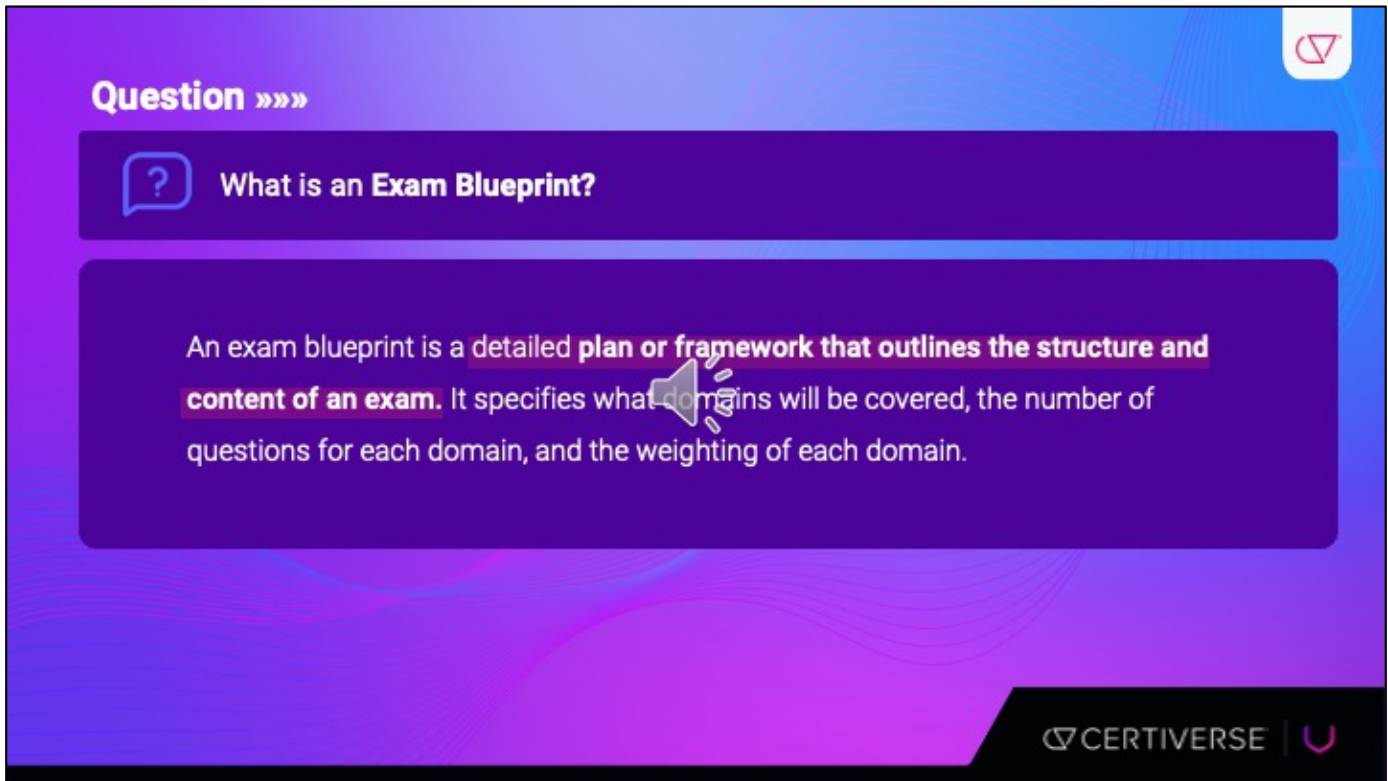
Introduction to

Exam Blueprinting



CERTIVERSE | U

Welcome to Certiverse University and welcome to this video, "Introduction to Exam Blueprinting." My name is Alan Mead, and I am the Chief Psychometrician for Certiverse. In this video, we're going to dive into the essential components of creating an effective exam blueprint. By the end of this video, you'll understand what an exam blueprint is, why it's important, where it fits into the exam development process, and how to create one. Let's get started.



The image is a screenshot of a video player interface. The background is a gradient of purple and blue with abstract wave patterns. In the top right corner, there is a small white icon of a play button. The main content is displayed in two dark purple rounded rectangular boxes. The first box contains a question icon (a speech bubble with a question mark) and the text "What is an Exam Blueprint?". The second box contains the answer text: "An exam blueprint is a detailed **plan or framework that outlines the structure and content of an exam.** It specifies what domains will be covered, the number of questions for each domain, and the weighting of each domain." A small speaker icon is overlaid on the answer text. In the bottom right corner, the logo "CERTIVERSE" is visible next to a small white icon of a play button.

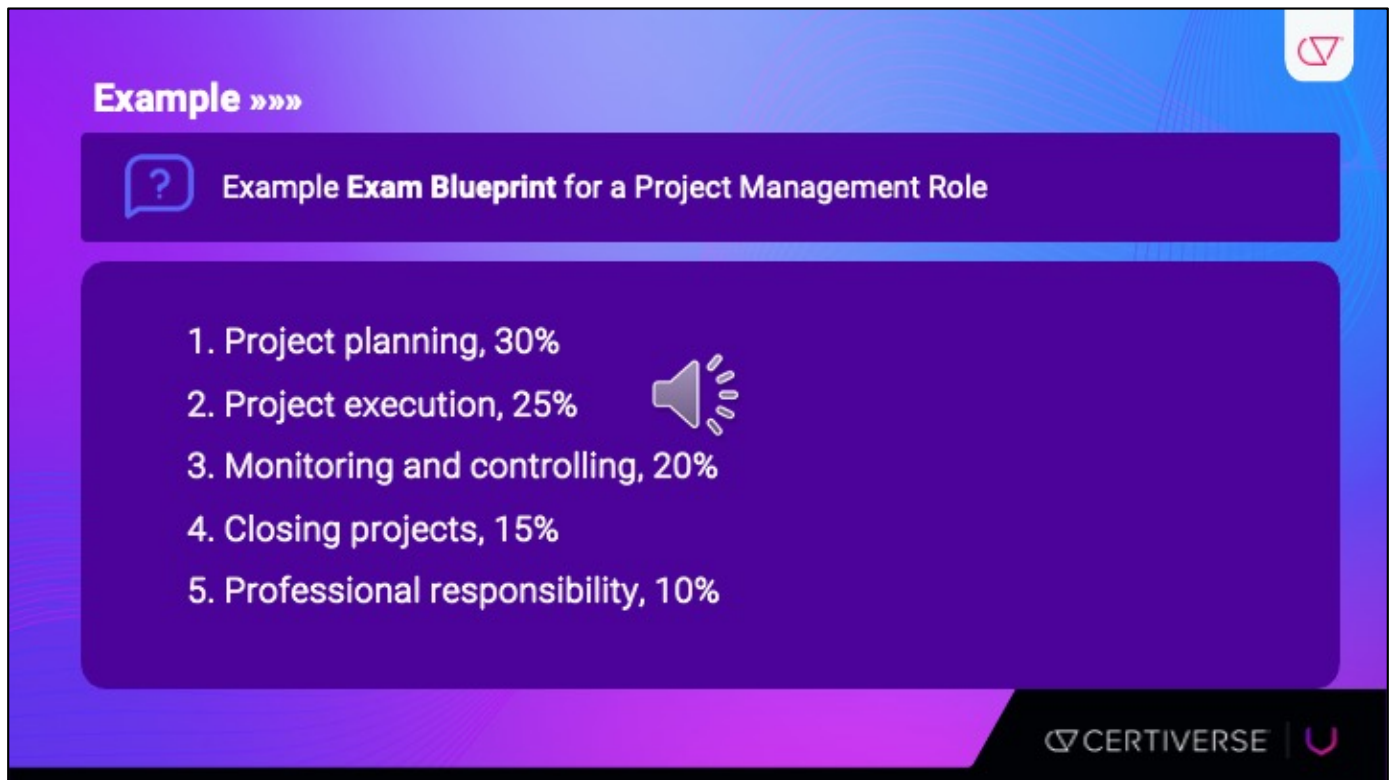
Question »»»

What is an **Exam Blueprint**?

An exam blueprint is a **detailed plan or framework that outlines the structure and content of an exam.** It specifies what domains will be covered, the number of questions for each domain, and the weighting of each domain.

CERTIVERSE | U

What is an exam blueprint (also called a test specification)?



The slide features a purple and blue gradient background. At the top left, the word 'Example' is followed by three right-pointing chevrons. Below this, a dark purple bar contains a question mark icon and the title 'Example Exam Blueprint for a Project Management Role'. The main content is a list of five items, each with a weight percentage. A speaker icon is positioned to the right of the second and third items. The bottom right corner of the slide displays the 'CERTIVERSE' logo and a small red 'U' icon.

Example »»»

Example Exam Blueprint for a Project Management Role

1. Project planning, 30%
2. Project execution, 25%
3. Monitoring and controlling, 20%
4. Closing projects, 15%
5. Professional responsibility, 10%

CERTIVERSE | U

Here's an extremely hypothetical example of a content outline for a certification exam for a project management role.

There are five domains, and each domain has a weight.

For example, Domain 1 is "Project planning" and it has a weight of 30%. If the exam has 100 items, 30 of them would be about project planning.

Domain 1, "Project planning", is about three times more important than Domain 5, "Professional responsibility"



Example »»»



Example Subdomains for a Project Management Role

1. Project Planning (30%)

1.1 Scope and Objectives: Defining project goals, deliverables, and requirements (15%)

1.2 Scheduling and Budgeting: Creating timelines and budget plans (15%)

2. Project Execution (25%)

2.1 Team Management: Leading the project team and managing resources (15%)

2.2 Task Coordination: Executing project tasks according to the plan (10%)

3. Monitoring and Controlling (20%)

3.1 Performance Tracking: Monitoring progress and performance metrics (12%)

3.2 Change Management: Managing changes to project scope, schedule, and costs (8%)

4. Closing Projects (15%)

4.1 Final Deliverables: Completing and handing over project deliverables (7%)

4.2 Documentation and Review: Finalizing project documentation and conducting post-project reviews (8%)

5. Professional Responsibility (10%)

5.1 Ethical Practices: Adhering to professional and ethical standards (5%)

5.2 Compliance: Ensuring legal and regulatory compliance (5%)



»»»

Blueprint's Role in Exam Development

- If a job task analysis (JTA) is conducted, the blueprint is the main output of the JTA
- Regardless, a blueprint must be finalized in order to start item-writing

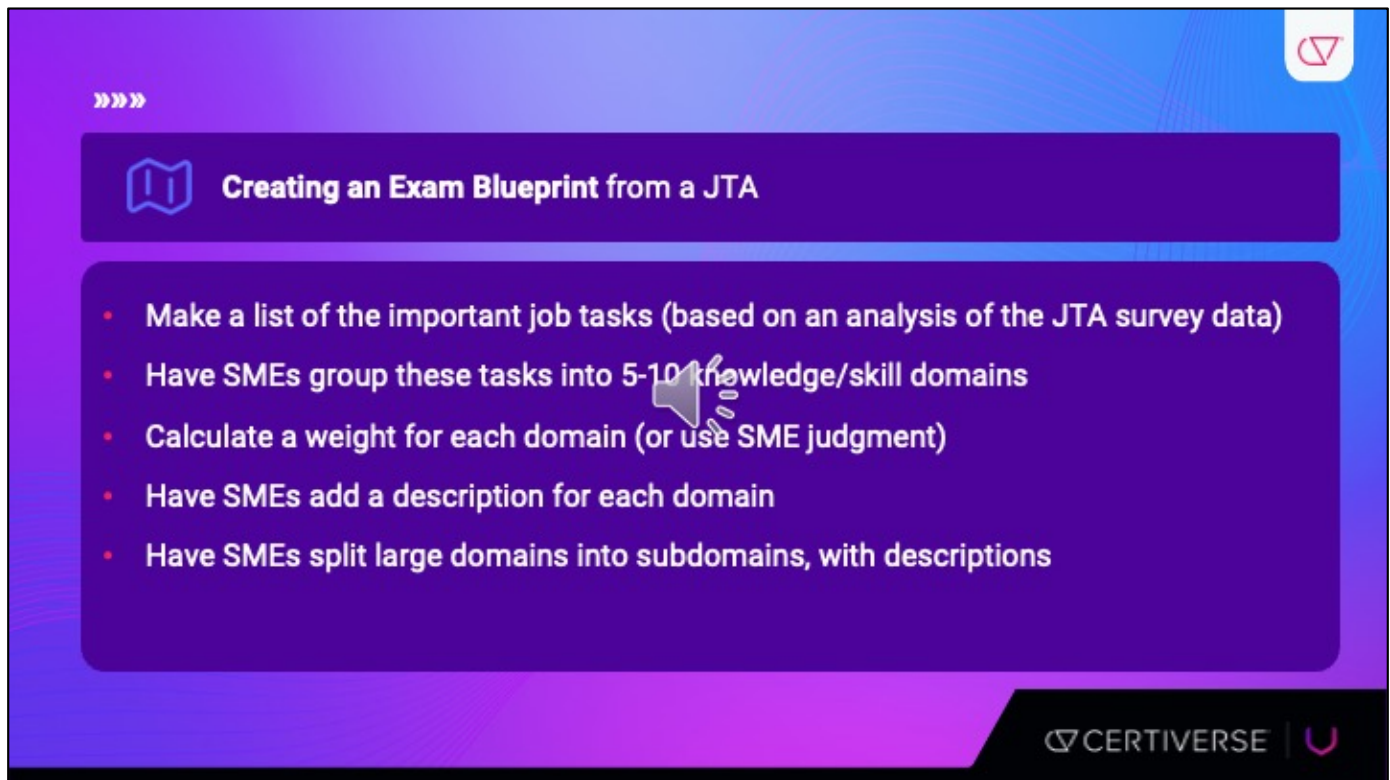
Job Task Analysis → Exam Blueprint → Item Writing → Beta Testing & Analysis → Standard Setting → Creating Forms & Scoring → Exam Delivery

CERTIVERSE

A blueprint is the primary output of a job analysis or JTA.

And just like a house's blueprint guides the building of a house...

...an exam blueprint guides the development of the exam content, structure, administration, and scoring.



The image is a screenshot of a presentation slide with a purple and blue gradient background. At the top right, there is a small white square containing a red logo. Below this, on the left, are three white arrows pointing right. The main title of the slide is 'Creating an Exam Blueprint from a JTA', preceded by a white icon of an open book. Below the title is a list of five bullet points, each starting with a red dot. A white mouse cursor is hovering over the second bullet point. At the bottom right of the slide, the text 'CERTIVERSE' is displayed next to a red logo.

»»»

Creating an Exam Blueprint from a JTA

- Make a list of the important job tasks (based on an analysis of the JTA survey data)
- Have SMEs group these tasks into 5-10 knowledge/skill domains
- Calculate a weight for each domain (or use SME judgment)
- Have SMEs add a description for each domain
- Have SMEs split large domains into subdomains, with descriptions

CERTIVERSE

Creating a blueprint from a JTA starts with the job task statements from the JTA survey. Analysis of the survey results tell you which are the important tasks.

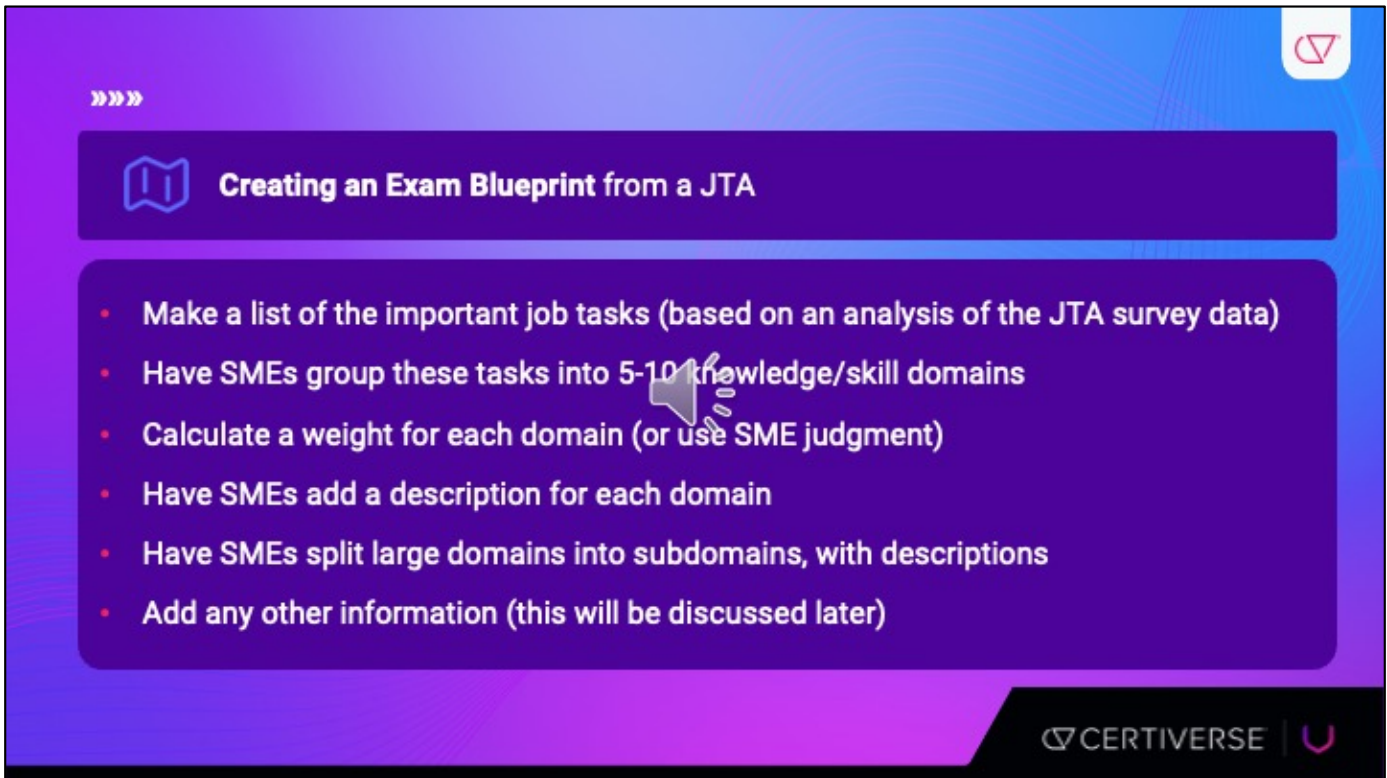
Ask your SMEs to group these important tasks into 5-10 knowledge/skill domains. On the Certiverse platform, we ask SMEs to do this when they author tasks, so this isn't an extra step.

Next, you must weight these domains. You can calculate a weight from the JTA results, and I'll show you an example. Or you can rely on SME judgment. In practice, it's usually mostly calculated with a little judgment.

You also need to add descriptions for each domain that tell the item-writers what items to write.

But then you come to another judgment call. Broad domains

may need to be broken up into more specific subdomains. Each subdomain needs a description and a weight.



»»»

Creating an Exam Blueprint from a JTA

- Make a list of the important job tasks (based on an analysis of the JTA survey data)
- Have SMEs group these tasks into 5-10 knowledge/skill domains
- Calculate a weight for each domain (or use SME judgment)
- Have SMEs add a description for each domain
- Have SMEs split large domains into subdomains, with descriptions
- Add any other information (this will be discussed later)

CERTIVERSE | U

Finally, there is some additional information that you need to specify so that you can get started with next steps. We'll discuss those items later because they need a little explanation.

Example »»»

Calculating a Weight from JTA data

- Assume that there are 124 tasks
- And 37 of them are related to **Domain 1: Project Planning**
- For these 37 tasks, the sum of the mean importance rating is 167.5
- For all 124 tasks, the sum of the mean importance rating is 545.6
- So, $167.5/545.6 = 0.307 = 30.7\%$, which rounds to 31%
- We then rounded 31% to a "simpler" value of 30%

1. Project Planning (30%)

- 1.1 Scope and Objectives: Defining project goals, deliverables, and requirements (15%)
- 1.2 Scheduling and Budgeting: Creating timelines and budget plans (15%)

2. Project Execution (25%)

- 2.1 Team Management: Leading the project team and managing resources (15%)
- 2.2 Task Coordination: Executing project tasks according to the plan (10%)

3. Monitoring and Controlling (20%)

- 3.1 Performance Tracking: Monitoring progress and performance metrics (12%)
- 3.2 Change Management: Managing changes to project scope, schedule, and costs (8%)

4. Closing Projects (15%)

- 4.1 Final Deliverables: Completing and handing over project deliverables (7%)
- 4.2 Documentation and Review: Finalizing project documentation and conducting post-project reviews (8%)

5. Professional Responsibility (10%)

- 5.1 Ethical Practices: Adhering to professional and ethical standards (5%)
- 5.2 Compliance: Ensuring legal and regulatory compliance (5%)

CERTIVERSE

Here's an example of how the 30% weight for Domain 1 Project Planning may have occurred.

Assume that there are 124 tasks

And 37 of them are related to Domain 1: Project Planning

And for these 37 tasks, the sum of the mean importance rating is 167.5. So, if you add up the importance rating for each of these 37 tasks, you get 167.5.

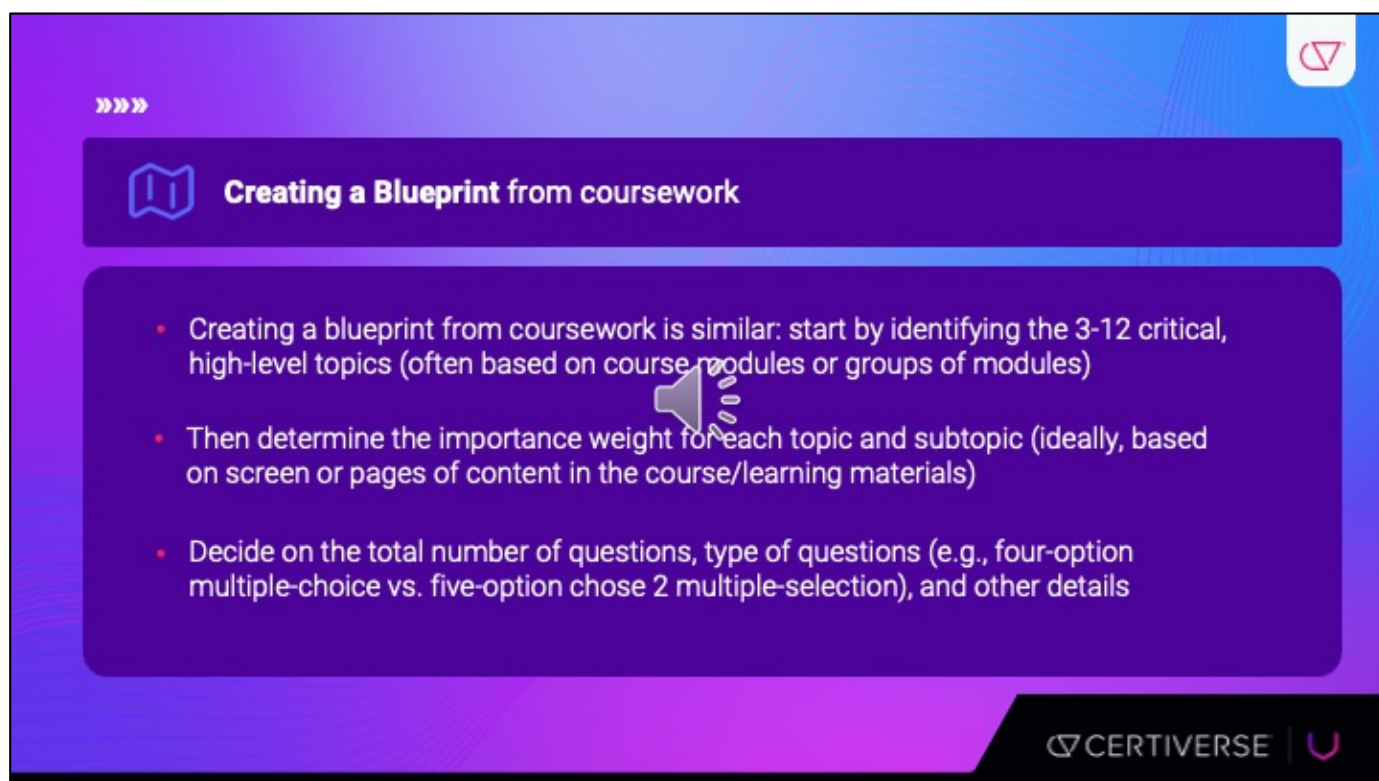
And for all 124 tasks, the sum of the mean importance rating is 545.6.

So, taking the ratio $167.5/545.6$ gives you a proportion of 0.307, which is 30.7%, which rounds to 31%


We then rounded 31% to a "simpler" value of 30%. This is what I

meant when I said that it's mostly calculation and a little judgment.


As I said earlier, the 15% weights for subdomains 1.1 and 1.2 are based on SME judgment that those two topics are about equal weights out of a total of 30%.



»»»

 **Creating a Blueprint from coursework**

- Creating a blueprint from coursework is similar: start by identifying the 3-12 critical, high-level topics (often based on course modules or groups of modules)
- Then determine the importance weight for each topic and subtopic (ideally, based on screen or pages of content in the course/learning materials)
- Decide on the total number of questions, type of questions (e.g., four-option multiple-choice vs. five-option chose 2 multiple-selection), and other details

CERTIVERSE | 

If you are building an exam to certify learning the material in a specific course, then you should treat the learning materials like a JTA.

Have subject matter experts identify the critical high-level clusters of knowledge and skills which will be your domains and decide on relevant subdomains, as needed.

Bear in mind that course material typically needs to be compressed to fit on an exam. You might have 100 hours of instruction, but only a single hour of assessment. So clearly the exam outline (i.e., the blueprint) will be at a higher level than the course outline. And the exam blueprint must emphasize the most critical knowledge and skills.

Once you have all the domains and subdomains, you need to assign weights. I like to start with weights based on some objective evidence. If a 200-page textbook book has 6 pages on

a particular topic, the default is that should be 3% of the exam because 6/200 is 3%. This may then be adjusted by subject matter experts.

And then finally, you need to decide the details that will guide item-writing, in exactly the same manner as if you had used a JTA to produce the Blueprint.

One note about defensibility: It depends upon the linkage to the course materials. You should be able to show that the exam blueprint covers the course material. However, you are relying on the course materials covering the critical elements of the job or role. The defensibility of the exam ultimately relies on having a curriculum that is job-related.

»»»

Organization of a blueprint

- There is no one "right" way to organize a blueprint (but probably lots of wrong ways). The number one issue is: Is this outline clear?
 Is there a logical start, middle, end to the content? Or externally accepted outline? Or other natural organization?
- Probably only 5-10 main domains and fewer than 25 subdomains. Avoid blueprints so detailed that forms cannot ever fully reflect the weights.
- Prefer weights as percent of total or number of items

CERTIVERSE

One question that arises is how you should organize your content?

There is no "right way" although there are probably wrong ways.

The fundamental question is: "Is this outline of the exam content clear to all the stakeholders?" You want the answer to be "Yes"

I've seen blueprints organized according to a typical workflow. For example: gathering requirements, building prototypes, coding, beta testing.

Or if you were building an exam to cover a standard, you'd probably organize the content according to the standard

Or else, arrange content in a logical way with simpler or more basic topics before more advanced or complex topics

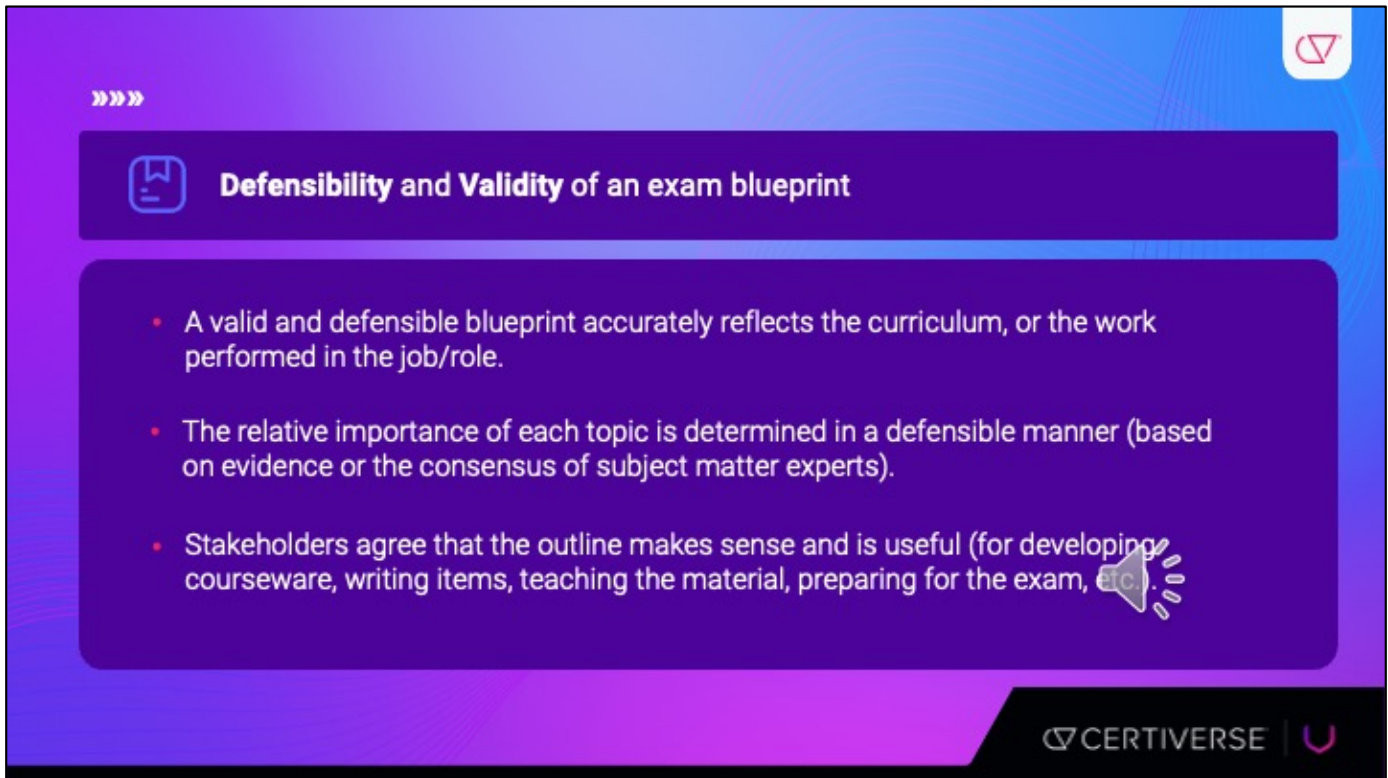
--

Blueprint complexity varies dramatically, and my ideal blueprint has 5-10 domains and fewer than 25 subdomains. There are blueprints with so much detail that no exam form can actually exactly meet the blueprint, and that's probably excessive for most exams.


--

Your weights should naturally translate into number of items per topic, so I would avoid saying "Domain 1 is 10% and subdomain 1.1 is 10% **of that** and subdomain 1.1.1 is 10% **of that** because that's .0.1% of the exam which isn't feasible.


Instead, **all weights** should be percent of the total items or just the number of items, and percent is more flexible for a new exam where the total number of items might be a guessimate. Percents are also more precise in some cases.




»»»

 **Defensibility and Validity of an exam blueprint**

- A valid and defensible blueprint accurately reflects the curriculum, or the work performed in the job/role.
- The relative importance of each topic is determined in a defensible manner (based on evidence or the consensus of subject matter experts).
- Stakeholders agree that the outline makes sense and is useful (for developing courseware, writing items, teaching the material, preparing for the exam, etc.).



CERTIVERSE | 

As we've discussed, the evidence that your exam blueprint is valid is based on having a clear linkage to the coursework or to a JTA. This enables you to assert that exams created from this blueprint will either cover the curriculum of the course or cover the knowledge and skills required by the job/role.

A good blueprint is well-organized with clear weights, and a good blueprint serves many purposes:

Item authors know how to write appropriate items

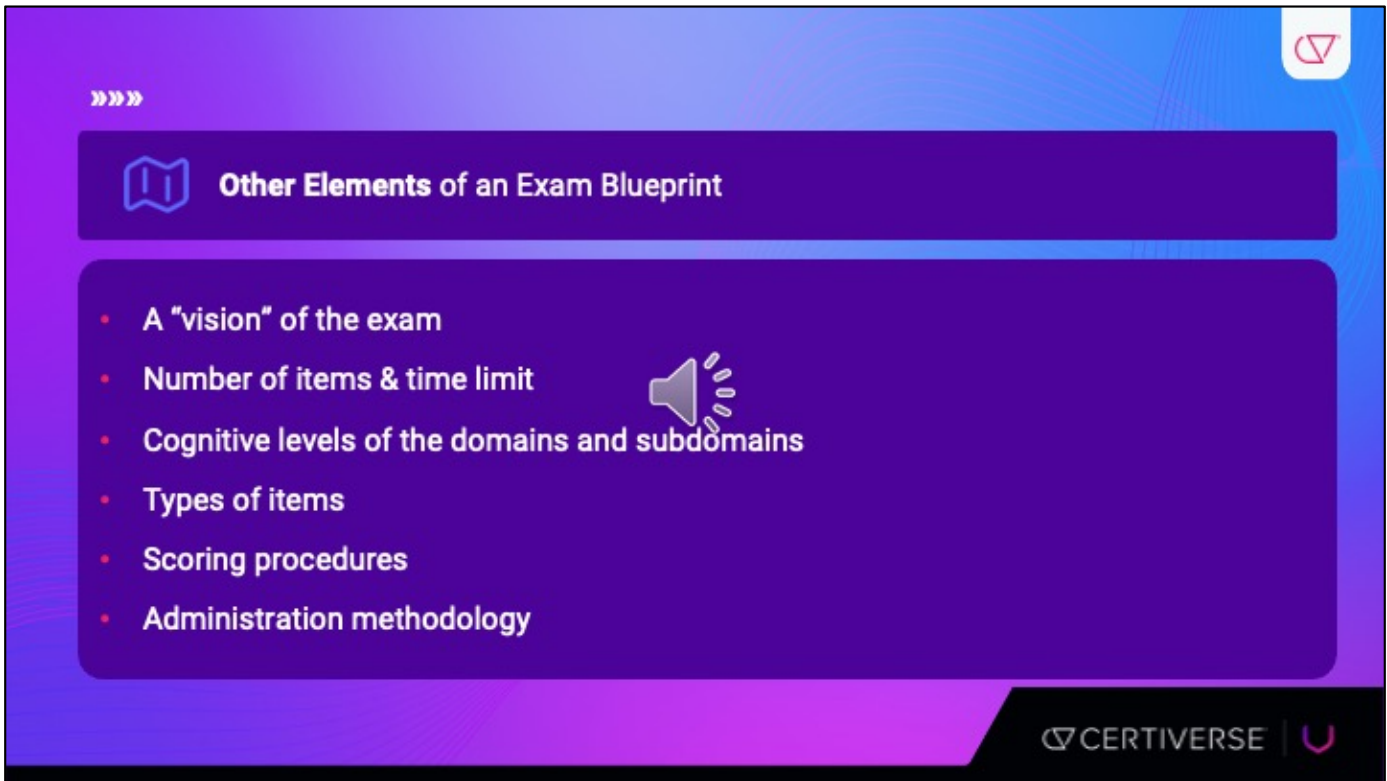
Item reviewers can determine whether an item is on topic, or not

Exam developers know how many items to include on a content-valid form

Candidates know how to study for the exam, and

Instructors know what content to include in each lesson

This is what I mean when I say that "a valid blueprint makes sense" to all the stakeholders of an exam.



»»»

Other Elements of an Exam Blueprint


- A "vision" of the exam
- Number of items & time limit
- Cognitive levels of the domains and subdomains
- Types of items
- Scoring procedures
- Administration methodology

CERTIVERSE


We've been focused on the content outline part of an exam blueprint. There are some other elements that should be included. I'll discuss each one:

- Your "vision" of the exam should clearly define the exam's purpose, target population, and difficulty level.
- The number of items on the exam and the time limit (or that the exam is untimed).
- Cognitive levels of the domains and subdomains
- Types of items that should be used on the exam
- Scoring procedures (where applicable)
- The methodology for exam administration, proctoring, etc.

»»»

 Your "vision" for the Exam

- What is the purpose of your exam?
- Who is the intended candidate population?
- What difficulty level are you targeting?

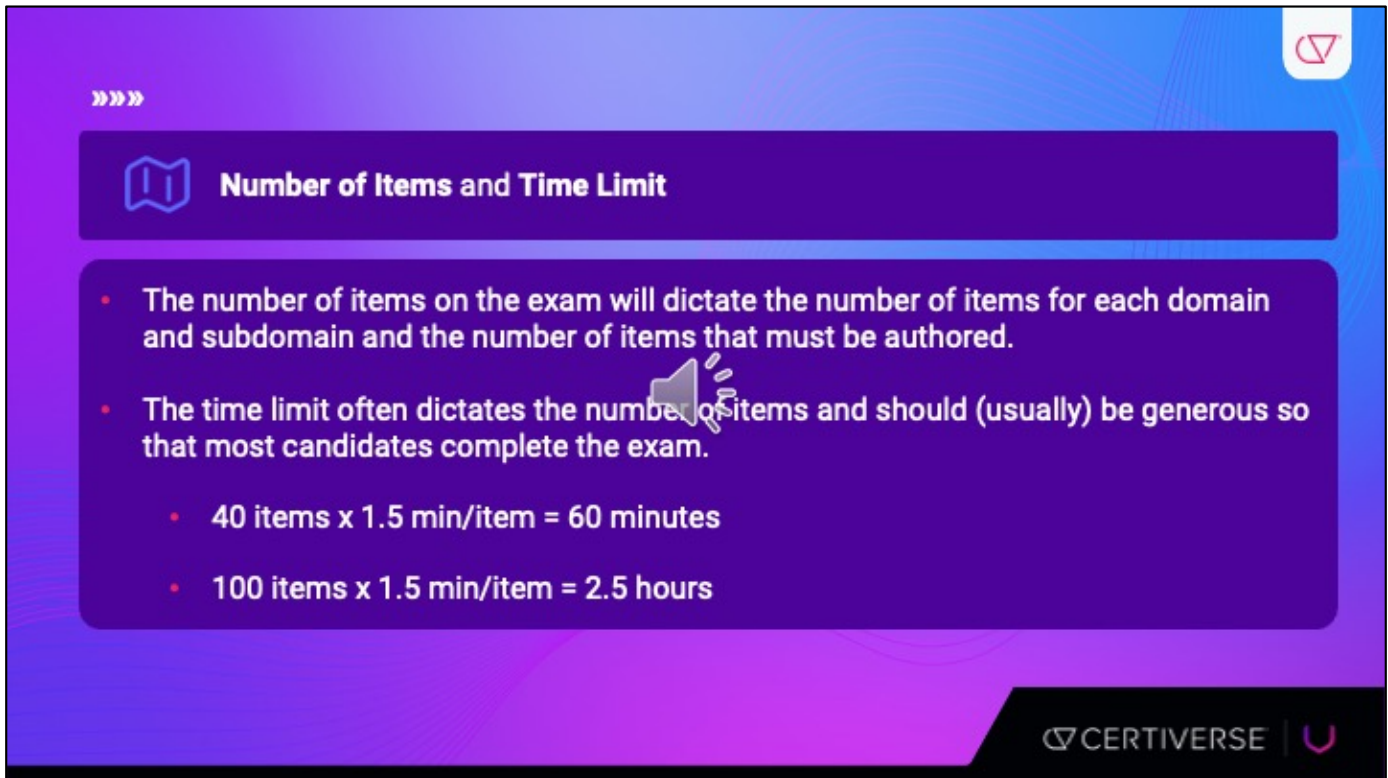
CERTIVERSE | 

Your exam blueprint should communicate to all stakeholders what kind of exam it is. That includes at least three topics.

What is the purpose of your exam? This usually includes the job or role and maybe additional information.

What is the intended candidate population? Is this exam for anyone? Or only people who have five years of experience?

What difficulty level are you targeting? Most exams target about 70% correct for the average item in the target population.



»»»

Number of Items and Time Limit

- The number of items on the exam will dictate the number of items for each domain and subdomain and the number of items that must be authored.
- The time limit often dictates the number of items and should (usually) be generous so that most candidates complete the exam.
 - 40 items x 1.5 min/item = 60 minutes
 - 100 items x 1.5 min/item = 2.5 hours

CERTIVERSE

Obviously, the number of items and administration time are important choices.

The total number of items on the exam will dictate the number of items for each domain and subdomain and the number of items that must be authored.

For example, if you have a 100-item exam and 15% should be for subdomain 1.1, then each exam form will have 15 items from this topic.

For most multiple-choice exams, a few dozen items are sufficient for reliable overall scoring. Sometimes more are needed to hit all the points on the content outline for the exam.

But often the time limit dictates the number of items. You want a reasonably generous time limit so that most candidates have the opportunity to complete the exam. The amount of time per

item varies across content areas and item types, but the industry average is about 1.5 minutes/item.

So, a one-hour exam can only have 40 items

And you may need over two hours for a 100-item exam.

»»»

Cognitive complexity

Eval.
Synthesis
Analysis
Application
Comprehension
Knowledge

- Bloom's taxonomy of learning objectives specifies levels complexity from simple recall, comprehension, application, analysis, evaluating, up to novel creation.
- It is important that exam items measure the intended level, which requires that each domain/subdomain on a blueprint has an intended level of complexity.
- As a simpler alternative, consider avoiding trivial recall. Ask yourself: Does a proficient candidate need to know this? (no "What does SQL stand for?" questions)

CERTIVERSE | U

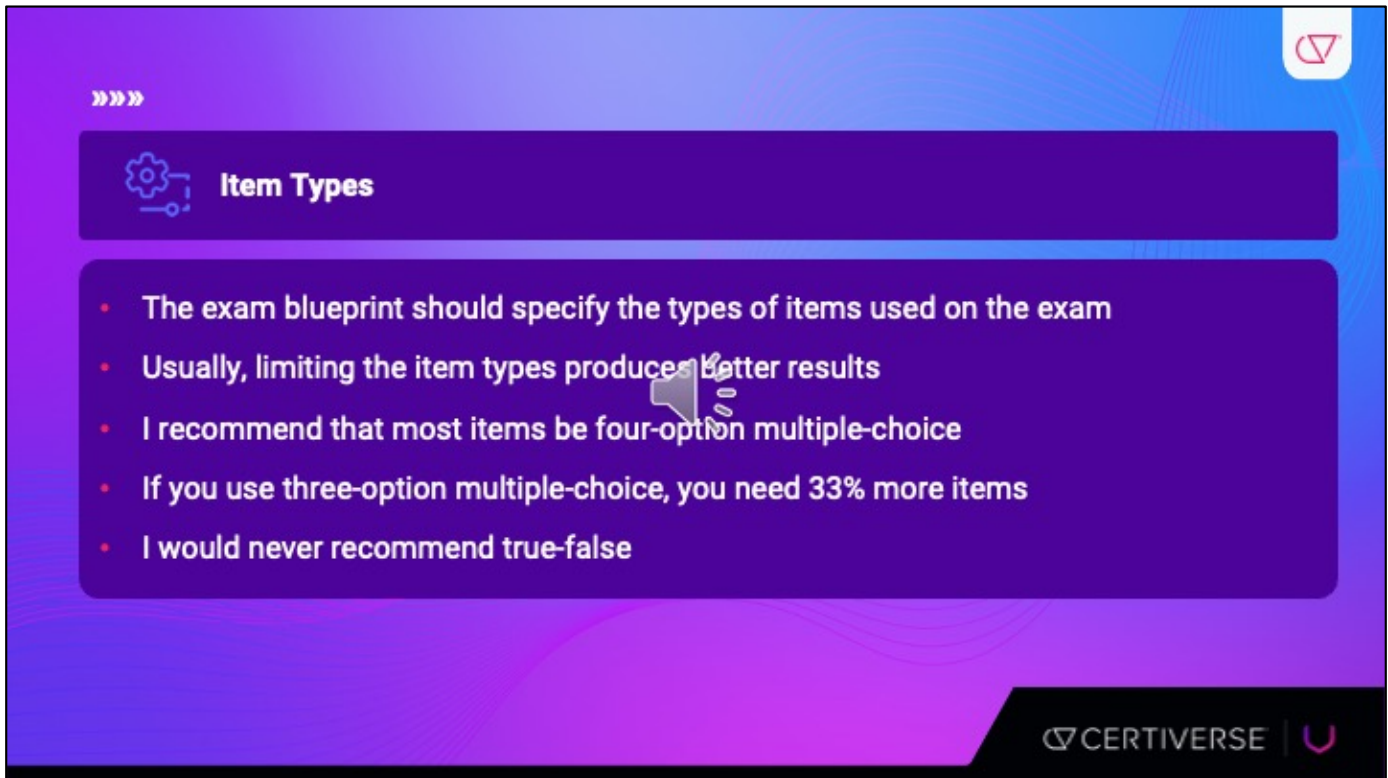
Another element that is commonly specified in the exam blueprint is the mix of cognitive complexity of the items.

Some variation of Bloom's Taxonomy of cognitive complexity is most common, frequently a modified taxonomy formed by collapsing some of these levels.

If it's important to target specific levels of cognitive complexity, then the blueprint should specify the correct level or mix of levels for each topic.

However, it can be hard to know what levels to include, so I often use a simpler alternative: I instruct the item writers and reviewers to ask themselves, "Does a proficient candidate need to know the answer to this question?" The answer should either be: "Yes, absolutely a proficient candidate **MUST** know this." or else that "Getting this item correct increases my confident that a candidate is proficient."

My example for item writers is NOT to write “What-does-SQL-stand-for?” questions. An expert SQL user may not know this and knowing this does not necessarily increase my confidence that the candidate can do anything useful with SQL.



»»»

Item Types

- The exam blueprint should specify the types of items used on the exam
- Usually, limiting the item types produces better results
- I recommend that most items be four-option multiple-choice
- If you use three-option multiple-choice, you need 33% more items
- I would never recommend true-false

CERTIVERSE | U

Your exam blueprint should specify the types of items that will be used on the exam. This is particularly important if some items require more time. You would want to ensure that each form of the exam has the same mix so that the timing can be the same across forms.

Unless there are extenuating circumstances, I think it's a better candidate experience to have a single item type for all items. You don't want candidates to be confused about the type of item.

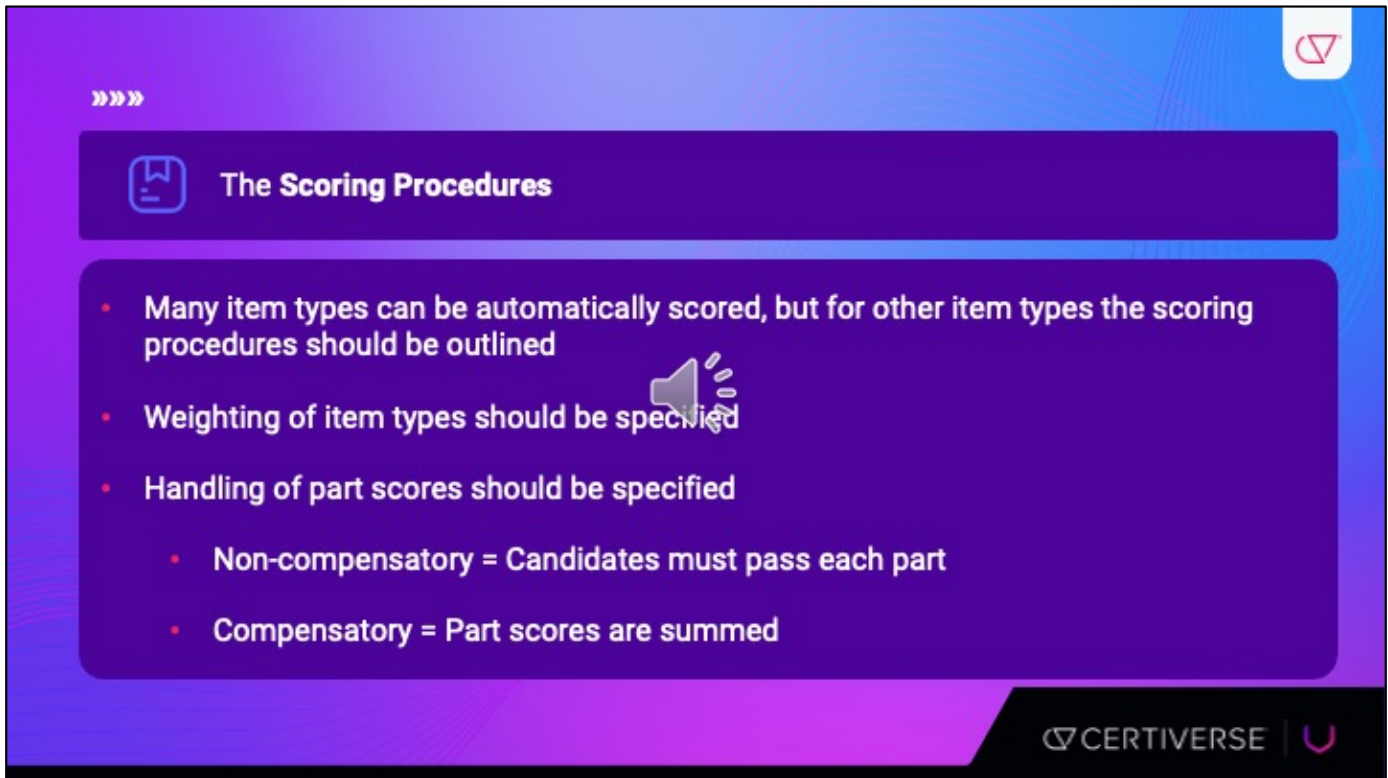
Again, unless there are extenuating circumstances, I think four-option multiple-choice with one correct answer and three distractors is the best format for most items.

Psychometricians have done quite a lot of research that shows that three-option multiple-choice are better PROVIDED that you


use more items. As a rough estimate, I'd say you need 33% more 3-option items.


If you simply replace four-option items with three-option items and DO NOT increase the number of items, then exam score reliability will be reduced.


I would never recommend using the true-false item type, because my dog has a 50% chance of getting the item correct. That means you need a lot of true-false items to get a reliable score.



»»»

 **The Scoring Procedures**

- Many item types can be automatically scored, but for other item types the scoring procedures should be outlined
- Weighting of item types should be specified 
- Handling of part scores should be specified
 - Non-compensatory = Candidates must pass each part
 - Compensatory = Part scores are summed

CERTIVERSE | 

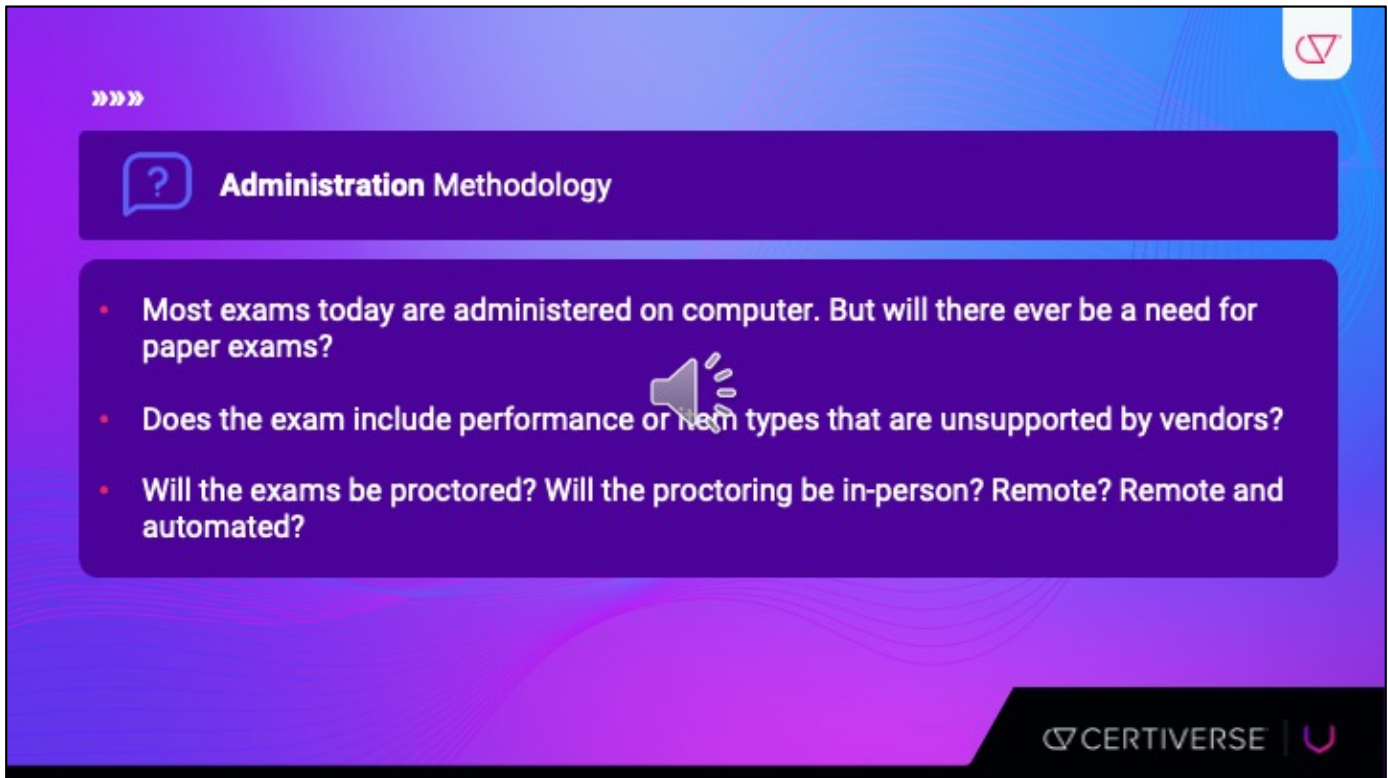
If you included item types that require more complex scoring, then the blueprint should specify the scoring mechanism, such as human graders, or a combination of a human grader and one machine grader.

Scoring also includes weighting of the items. Exams with all multiple-choice items usually give each question equal weight. But if there are other item types or sections, the weighting scheme should be specified.

Scoring procedures can also include how parts of an exam are handled. It's common, for example, for some regulatory exams to have general and state-specific parts and a candidate may be required to pass both parts in order to pass the overall exam. This is called a "noncompensatory" or "conjunctive" model. If you add up the score on both parts and it needs to meet a cutscore, that's a "compensatory" model because a person could compensate for a lower score on one part with a higher score

on another part.

Although noncompensatory scoring is often the “obvious” way to score parts, it's worth noting that compensatory is psychometrically better because it eliminates a subtle double-jeopardy issue that slightly damages the reliability of noncompensatory scores.



»»»

Administration Methodology

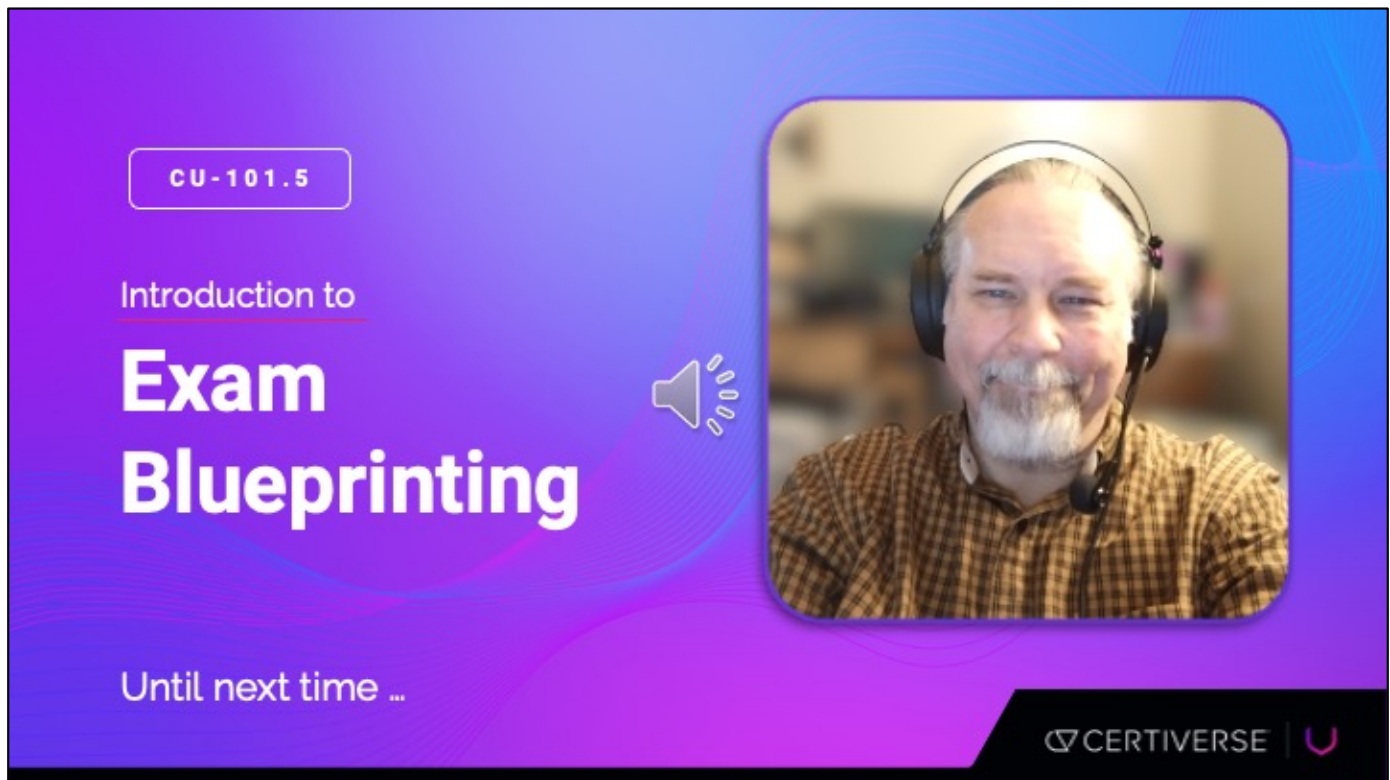
- Most exams today are administered on computer. But will there ever be a need for paper exams?
- Does the exam include performance or item types that are unsupported by vendors?
- Will the exams be proctored? Will the proctoring be in-person? Remote? Remote and automated?

CERTIVERSE

Today, the predominant method for administering exams is through computer-based systems. However, it's essential to consider the potential necessity for paper exams in certain scenarios such as testing in under-developed areas, as a fallback security measure, event testing, etc.

Additionally, assessing whether the exam includes performance tasks or item types not supported by vendors is crucial for determining the administration methodology.

And finally, Proctoring is a vital aspect ensuring integrity of the exam score; deciding whether it will be in-person, remote, or automated is a significant consideration in the administration process.



CU-101.5

Introduction to

Exam Blueprinting

Until next time ...

CERTIVERSE

The video player interface features a purple and blue gradient background. On the left, the course code 'CU-101.5' is in a white box, followed by the title 'Introduction to Exam Blueprinting' in white text. A speaker icon is positioned to the right of the title. On the right side, there is a rounded rectangular video frame showing a man with a grey beard and headphones. At the bottom right, the 'CERTIVERSE' logo and a purple 'U' icon are visible.

So, I hope you have enjoyed this video on Exam Blueprints, which are a foundation for developing fair, reliable, and valid exams! Hopefully, you can now see the path to creating a blueprint and you have a gasp of the steps involved and the judgment calls required.

In the next video we'll address the next step, item-writing.

That's it for this video! Thanks for watching ... we'll see you in the next!