

Performance Criteria Tips

Instructions:

Figuring out how you are going to be assessed is a really good way to determine what work needs to be done, as well as where your priorities should be. This guide provides tips on how to determine performance criteria which can be used to both assess the quality of your work and help you make decisions.

Tip: Talk the Talk

You should talk professionally about your performance criteria with your stakeholders and clients. By using the correct terminology, you will inspire confidence in your work.

- One of the most common terms is a **metric**, which is simply a formally defined way of measuring something.
- A quality of the something that you care to measure is called a **criterion** (plural criteria) or an **attribute**.
- The result of the metric used to measure criteria is referred to as the metric or criteria **score** (also known as a metric *result* or *value*).
- A group of metrics used to measure a collective group of criteria is accurately referred to as a **rubric** or can also be called a **decision matrix***
- The importance of any single criterion vs another criterion is referred to as that criterion's **weight, priority, or importance**. This is often shown as a percentage (i.e. "this criterion counts toward 25% of the total performance score).

*The combined scores from the group of metrics laid out in a matrix can be used to help you make a decision. However, you will hear in conversation people still referring to a group of metrics as a metric as well. This is fine so long as you know the difference yourself.

Tip: Watch out for subjective performance metrics, metrics that can only be measured at the end of your work, and metrics that are difficult or costly to measure.

Performance metrics should *always* be written in an objective fashion. In other words, the way one person would assign a performance score is that same way 1,000,000 other sufficiently-trained people would assign that score. This is because the

performance metric is defined by having an objective way of accurately measuring the associated performance criteria.

But now imagine you were designing a toy and the stakeholder said, "We want it to be *fun*". This is a good start as it informs you of the spirit of what your stakeholder is looking for. But how do you define *fun*? And is it fun enough? And fun in the right ways?

When you encounter a naturally subjective criteria, you still need to develop an objective way to measure the criteria. For example, maybe you measure the length of time a child plays with it, or how often the toy is selected first when the child comes home from school. These are measurable indicators of the criteria you're looking for, which is good, but taken alone they also have some problems.

The first problem is that it would be quite time consuming to perform this test, especially depending upon how many data points you needed. And hence it could also be costly to pay the person monitoring the children.

A potentially larger problem is that these measures require the toy to be fully *created* in order to be measured. In this case, only at the end can you tell how well you've been doing. So, if you find out only at the end you're doing poorly in one area, you'll have a lot of rework to do. Now compare that to if you caught an indication of the deficiency earlier on, you could make adjustments sooner, before too much time, effort, and resources have been expended. So, it's important to get feedback on your deliverables during your work progress.

A simple way to receive feedback during the process is by setting up sub-deliverables. Consider this like a homework assignment that targets only part of your final project. Many times though, in setting these sub-deliverables up, you need to be creative. Perhaps another measure of fun could be how many features the toy has. You'd have to define what a "feature" is but that could also be an informative process to learn about the stakeholder's needs. Regardless, the "number of features" is something you could focus on early in the design to be able to guarantee at least some level of performance and it's also something that comes at practically no additional cost to measure.

Tip: Don't expect to be told all of your performance criteria

Although it may seem counterintuitive, it's very rare that all of the performance criteria for a project are specifically stated by the stakeholder/client etc. Stakeholders don't know how to express what they want, but they'll know what they want (or don't want) when they see it.

Therefore, it's your job as a designer to determine what all the performance criteria are. Some good questions to ask are:

- What would make for a good deliverable vs. a bad deliverable?

- What would make for a good deliverable vs. a great deliverable?
- Imagine there were multiple deliverables to choose from that were the same in every way except for one difference. What might be some possible ways these deliverables could differ? How much better would one deliverable be over the other because of that difference?
- How would you measure the quality of these differences? Or how can you know a deliverable is better or worse in a certain way?
- Which of those ways are most important? How much more important are some over another?
- If you had to rate an employee on their ability to produce this deliverable, what would be some of the key characteristics you would look for?
- Are there any indicators to be looked for early on to tell whether a deliverable will be good, bad, or ugly?

Not only will these questions help to bring out deliverable expectations, but these questions also assess the deliverables by the performance criteria.

Tip: Only by having the performance metrics established can you objectively prove that you have in fact produced one of the best possible, if not the current very best solution.

Establishing performance metrics and thresholds of achievement for those performance metrics are an objective way to establish contract terms of completion. In other words, once you've shown that your system measures at a certain performance criteria value, you have shown that you have met your stakeholder/client's needs and fulfilled your contractual obligations. Making performance metric scores a part of the establishing contract process can also help make it very clear where their priorities lie.

If you don't develop objective measures though, you run the risk of never being able to prove you've sufficiently completed your deliverables. So your stakeholder could say, "Yeah your toy is kinda fun. But that's not the kind of fun we were looking for. Your contract isn't complete, and we can't pay you until we think you've made it fun enough." That's a tough argument to be in.

Without objective performance metrics, you either have to do more work for the same pay, or you will have one very unsatisfied stakeholder who will unlikely want to work with you again.

Therefore, if you ever feel you have to change your performance criteria or metrics during the design process, it may also be a good idea to check with your stakeholders and explain to them why this change is also in their best interest.

Tip: Make sure that you're measuring what you think you're measuring.

Using the toy example, where you're trying to measure fun, imagine one toy comes with a "free" specialized action figure, another toy has two figures, and another one doesn't have any. If you use the number of included action figures as a measure of fun, is it fun that you're measuring or is it a measure of an added perceived economic value from getting a "free" figure? Perhaps it's both. So taking the number of included figures alone would therefore probably be improper way to measure fun. However, it could still be one aspect of how you measure fun.

For example:

- the number of free figures,
- the number of features,
- the popularity of the character branding,
- the number of recommended players,
- the number of modes of play
- could all be things you could measure to collectively determine a fun criteria score.

This is not to say though that the same measurement can't be used as part of the calculation of many criteria. For example, it's perfectly fine to use the mass of your system as a measure included as part of both an "ease of operation" criteria calculation and then an "ease to ship" criteria calculation.

Of course, to implement the above suggestions, you'll also have to make sure these are all objective measures. In other words, what defines something as a "feature"? Or what defines something as a "mode of play"?

Both of these might be measured differently depending upon the kind of the toy. For example, consider a set of Legos vs. a doll house vs. an electronic handheld game vs. an RC car. All of these might count modes of play or features differently. Modes of play for the Lego set might be the number of alternative instructions that an expert could design. For the doll house it might be how many common household activities can be simulated using the toy from a larger possible set of real world common household activities. For the electronic game, it might be how many games and game modifying options are available, or how many game and option combinations are available. And maybe for the RC car there is really one main mode of play, driving the car, so perhaps only measuring the features makes for a better metric.

Again, features might also be measured in different ways but one way to do this for RC cars would be to make a list of all of the features that are currently highlighted in the specs and packaging of a wide variety of RC cars. Then measure your design based upon how many of these features your system shares.

Although your list may not be completely inclusive of every feature ever made, or be could made, it still helps to create a standard of comparison which is far better objective design tool than perhaps even the most expert RC car specialist's opinion. Other common metrics that need to be defined well are what defines a system as having successfully completed something and what defines something as a failure. Again, think about what is most important to your customer's needs. For example, common ways to measure a failure are the impact on the time and cost it takes to address the failure and the frequency at which the failure occurs.

Tip: Make sure your evaluation criteria match your customers' needs, not necessarily your own.

You are rarely the ideal customer for what you're designing, so make sure your evaluation criteria match your customers' needs, and not necessarily your own.

This is particularly important when defining objective ways of measuring for seemingly subjective performance metrics like "coolness". As an example, imagine you are comparing coolness in cars. Since how what you think is "cool" may or may not be as cool to your customer, be sure to ask the questions:

What is it about the component that makes one thing cooler than another?

- How would I quantitatively measure that characteristic?
- Such questions can help you refine your objective measures further.

In a similar way, you may be a potential customer of your design work but that doesn't mean all of your needs and opinions match those of the target "ideal" customer. When your customer is actually very different from you, it's sometimes easier to create objective metrics. When you are similar to your customer but still significantly different in some ways, it can sometimes be harder to distinguish what is actually most important to the customer versus what is your natural intuition as to what's important to you.

In these cases, it's crucial to actively look for the differences between you and your customer so you can still use your intuition but be professional in checking it against your measures of what your customer actually wants.

Otherwise, you may find yourself in the not uncommon position of "I don't understand why they don't like our idea. Why can't they see how cool it is? It's the most awesome thing I've ever seen." But it may not be the most awesome thing your customer has seen, because you aren't addressing all their needs or have weighted the importance of their needs accurately.

Tip: Performance criteria and metrics may need to be adjusted during the design process.

After you've done your initial performance criteria identification, while working with your stakeholder, continue to look for additional criteria as well. It's common to recognize criteria during the design, building, or even testing phases of your work. For example, you know your stakeholder mentioned "maintenance" specifically but they didn't mention "installation". Chances are they're going to evaluate your system on that as well.

It is also common to determining criteria that are internally important to your own work.

For example:

- how easy it is to assemble (especially when the client expects it delivered already assembled, like a car),
- the efficiency or safety of performing a process
- whatever else your team may care about but does not directly impact the stakeholders

You can always double check with your stakeholder as you determine additional performance criteria but remember your stakeholder may not know everything (or even that much) about how to get the deliverables they've asked for (that's why they're asking you to do it). So, in some cases you may need to develop your own arguments as to why another criterion is important as well.

Tip: All performance criteria needs their weight, or relative importance, determined.

It is unlikely that all identified criteria are equally important, so it's important to weight the performance criteria.

In deciding criteria's relative weights, some common influences are:

- the frequency at which criteria will affect a situation/use case and the importance of those situations
- the stakeholders' preferences
- marketing data
- how you estimate competitors' systems weight their performance criteria

You don't have to always take your stakeholder's preferences (or marketing data, etc) as law, but instead, as a designer, you should consider all the perspectives and make your own conclusions. For example, you as the designer may feel more ethically responsible to increase the weight assigned to a safety criterion than what your stakeholder may think is necessary. This also doesn't necessarily mean that your stakeholder is "bad", just perhaps misinformed and hopefully through your performance metric definitions you can demonstrate why it's advantageous to do so.

You may also find that the stakeholder tells you, "A is more important than B" then "B is more important than C" but "C is more important than A."

Wait a second, how is that possible?

It could be due to a stakeholder who doesn't understand their priorities. But a common logical explanation is that all these statements are true, but only in different situations. To help unravel this then, it's up to you to work with your stakeholder to determine what those situations are, and perhaps which situations are more important or more common.

Tip: Take a good first guess on weights.

To get a possible good first guess at your performance weights, arrange your criteria in an inequality statement like the following example done for the performance criteria generically named A thru F.

$$C \gg A = D > B = E = F$$

Then assign a number, like perhaps 50 or 100, to the top criteria (C in the example given). Then go from left to right and assign a number relative to the next criteria according to your statement.

So if we started with C =100, we'd then see C is far more important than A. So we might assign A a value of 60.

Criteria D is considered roughly as important as A so we'll assign D a value of 60 too. D is more important than B, E and F so we'll then assign those last three each a value of 40.

Then as a final step, we sum up all of the numbers (100 + 60 + 60 + 40 + 40 + 40 = 340) and then divide each criterion's number by that sum to give an initial percentage weight:

$$A = 100 / 340 = 0.30$$

$$C = 60 / 340 = 0.17$$

$$D = 60 / 340 = 0.17$$

$$B = 40 / 340 = 0.12$$

$$E = 40 / 340 = 0.12$$

$$F = 40 / 340 = 0.12$$

This is unlikely to be perfect, and probably you would want to make some adjustments to this yourself. However, both the initial inequality statement and the resulting weight guesses are ways to give you, your team, and your customer, something to begin to react to.

Tip: Be sure to ask questions.

Be sure to ask important questions. For example, ask your stakeholders if an option gets a very low score in any one criterion, or several criteria, but still achieves a desired minimum overall performance score would the stakeholder still be willing to accept that option as a solution. Often asking the question this way, also helps them recognize that low scores in one criterion may be possible and help everyone to truly see how important that criterion is to them.

Tip: Make adjustments to your weights.

Establishing your performance criteria weights to accurately represent your customer's good preferences is an important, but often challenging, process. Often, once you're able to get a good first initial guess, discussions between you and your teammates, or you and your customer can be far more productive. They may not be able to fully articulate their preferences, but their reactions to different weightings can be very informative.

It is perfectly acceptable to create an initial guess. Then iterate and make changes until you believe you can make an argument from your understanding of the customer's needs as to why these weights were selected. Then starting with your weights, repeat the process with your team, adjusting, and recording what the team's refined and defensible argument is.

Finally, repeat this process using the team's agreed upon weights with your customer.

Of course, you do not need to follow these steps precisely, and you can talk to anyone, at any point, as you see best. However, the important thing is that you communicate to the customer that this is how you are going to assess their needs being met. If they aren't comfortable with that, continue the discussion until they can agree that a solution that performs to a certain level in your decision matrix is one they will be willing to accept as their solution.