**Performance Criteria Tips**

**Cornell Cup USA – Arm Enabled**

**Prof. David R. Schneider, Cornell Director of MEng Studies in Systems Engineering, drs44@cornell.edu**

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/or help you to make decisions. This guide works along side:

* The Defining Your System Part II guide, which demonstrates how performance criteria can be found from a functional analysis of your System (you may want to read at least its “Discovering Performance Criteria through Functional Requirements” section first if you’re new to performance criteria).
* The Decision Matrix guide which gives you a well-respected way to use performance criteria and metrics to formally and objectively assess a solution and/or compare various options to prove their value.
* The Defining Deliverables guide which aims to help you define what your stakeholder/client is asking for you to “deliver” (which can include things such as a decision or recommendation) and what you need and need to do to achieve those things.

**Tip 1: Talk the Talk.** Talking about performance criteria professionally with your stakeholders/clients helps to inspire in 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 *criteria* or an *attribute*. The result of the metric used to measure a criteria is referred to as the metric or criteria *score*, a.k.a. 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*, i.e. the combined scores from the group of metrics laid out in a matrix can be used to help you to 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.

Finally, the importance of any single criteria vs other criteria is referred to as that criteria’s *weight* or *priority* which is often represented as a percentage, i.e. this criteria counts towards 25% of the total performance score.

**Tip 2: Don’t expect to be told all of your performance criteria.** Although it may seem counterintuitive, it is very rare that all of the performance criteria for a project are specifically stated by the stakeholder/client etc… They 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 the customer was given multiple deliverables to choose from that were exactly the same in every way except for one difference. What might be some possible ways they could differ and how much better would one would be over the other because of that difference?
* How would you measure the quality of each one of those ways they could differ? Or if you how would you know a deliverable is better or worse in that 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 that could be looked for early on to tell whether the deliverable will be good, bad, or ugly?

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

**Tip 3: Performance criteria & 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 fairly 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, i.e. 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, or 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 criteria is important as well.

**Tips 4 & 5: Watch out for subjective deliverables / performance metrics, metrics that can only be measured at the end of your work, and metrics that are difficult/costly to measure.** Performance metrics should *always* be written in an objective fashion; the way one person would assign a performance score is that same way 1,000,000 other sane and sufficiently trained people would assign that score, because the performance metric is defined by having an objective way of accurately measuring the associated performance criteria.

Butnow 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.

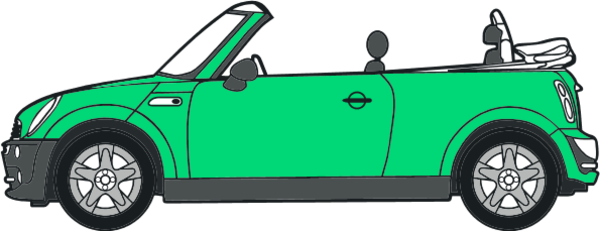
The first problem is that it would be quite time consuming to perform this test, especially depending upon how many data points you needed. And hence it could also be costly to pay the person monitoring the children.

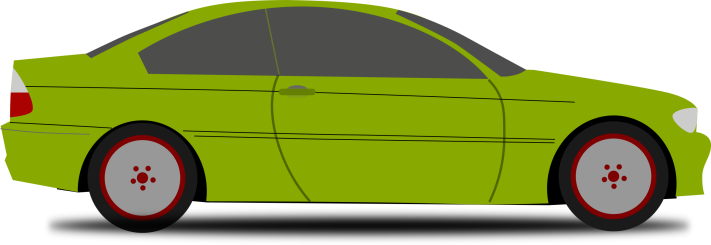
A potentially larger problem is that these measures *require the toy to be fully created* in order to be measured, i.e. only at the end can you tell how well you’ve been doing. If you find out only at the end you’re doing poorly in one area, you’ll have a lot of rework to do -- as compared to if you caught an even an indication of the deficiency earlier on, you could make adjustments earlier on before too much time/effort/resources have been expended. It is important though to get feedback (i.e. at least estimates of your final performance) on your deliverables during your work progress.

Sometimes this can be overcome simply by setting up sub-deliverables (think of a homework assignment that targets only part of your final exam / project). Many times though you need to be more 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.

This is just one example and you may need to understand the problem better before you can develop good objective performance metrics. For example, if it’s a toy catapult, you might figure out that increasing its projectile’s maximum velocity makes it more fun. Of course, you’d also have to probably optimize this criteria against a safety performance criteria as well.

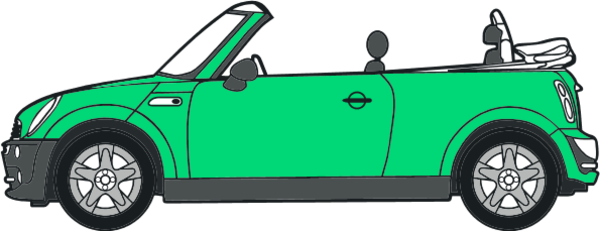
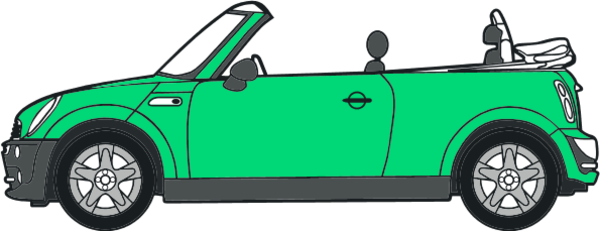
**Tip 6: You are rarely the ideal customer for what you’re designing, so make sure your evaluation criteria match your customers’ needs, not necessarily your own.** This is particularly important when defining objective ways of measuring for seemingly subjective performance metrics like “coolness”. As an example, imagine you are looking at the two pictures below and want to measure coolness.





**Figure 1: Car Coolness Comparison**

Having read the earlier tips you might naturally ask, “What is it about it that makes one thing cooler than another?”, then “How would you quantitatively measure that characteristic?” However, what you think is “cool” may or may not be as cool to your customer. For example, considering the two pictures above you might think that having decals on the side of the car is a good metric for measuring coolness. However, it’s then important to think about how something could meet your objective metric and not meet your intensions, as is perhaps the case with the two pictures below. The new one certainly has decals but may or may not be the “cool” you’re looking for. Asking the questions again “What is it about it that makes one thing cooler than another?”, then “How would you quantitatively measure that characteristic?” can help you refine your objective measures further.



**Figure 2: Car Coolness Comparison 2**

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 of their needs or have weighted the importance of their needs accurately.

**Tip 7:** **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. i.e. once you have 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. Stakeholder: “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 one tough argument to be in. Without objective performance metrics you’ve either have to do more work for the same pay, or you’ve made one very unsatisfied stakeholder who will be unlikely to 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 you stakeholders and explain to them why this change is also in their best interest.

**Tip 8: Make sure that you’re measuring what you think you’re measuring.** Taking the toy example again 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. 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, etc. could all be things you could measure to collectively determine a fun criteria score. (see Decision Matrix guide for more details on calculating scores and setting score thresholds) -- 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, i.e. What defines something as a “feature”?, What defines something as a “mode of play”? Both of these might be measured differently depending upon the kind of the toy, i.e a set of Legos vs. an RC car vs. a doll house vs. an electronic handheld game all 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 & 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 & packaging of a wide variety of RC cars, and then measure your design based upon how many of these features does your system share. 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 have having successfully completed doing 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. (The Brainstorming Risk guide also provides more information on how to quantify failures.)

**Tip 9:** **Make sure you determine the relative importance of all of your performance criteria, i.e their weights**. It is unlikely that all identified criteria are equally important, so it’s important to weight the performance criteria and some common influences to decide their relative weights are listed below

* the frequency at which it will affect a situation/use case & the importance of those situations
* the stakeholders’ preferences
* marketing data
* how you estimate competitors’ Systems weigh their performance criteria
  + you may want to compete head on or focus in other areas

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 of 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 criteria 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 of these statements are true but only in different situations. In order to help unravel this then, it’s your job to work with your stakeholder to determine what those situations are and perhaps which situations are more important or more common.

**Tip 10:** **Beware of Using Surveys:** Surveys may sound like a great way to measure a seemingly subjective performance criteria like “coolness” or “aesthetics” or “fun” but there are a lot of pit falls to surveys. First off surveys are expensive. They take a lot of time to create properly, to find the proper audience and run them, and then to analyze them properly.

Furthermore each step has significant potential for error. Creating survey questions that actually target the criteria without other having other factors influence the surveyed audience’s responses is quite difficult to do. They may say your design is cool but do they think it’s cool for the reasons you think they’re saying it’s cool? Or is there something else that is making them say it’s cool that you’re not recording properly in your survey? Or is there something else that would make it even cooler to them that you’re not asking about? Maybe the audience doesn’t even know how to express themselves or their needs well enough? If someone is asking for a car, but they never heard of an airplane, they might not know enough to ask for an airplane, even if it meets their needs much better.

Plus getting access to an audience that represents your target customers and is of reasonable size to conduct and later analyze the survey is a task in itself. Then when you finally have the survey results, can you be confident that they will be conclusive enough – or will you look at the results and say “Shoot, we should have asked about this as well. Looks like we’ll have to run another survey to tell what’s going on.” Trying to think about all the possible kinds of responses to a survey and how you would react to the possible responses can help you to create a far better survey in advance.

If you do decide to do a survey, remember a survey is a test. Like any test, a survey should try to do more than just give you a yes/no conformation but rather provide you with enough information that you are confident in which direction your design should move in next; whether it’s the same, or a different direction. Surveys if done properly can be great for validation (or disproval) of your ideas but remember that you need to be smarter than your target customers – after all if they were able to understand their needs and what could solve their needs well enough, they might not need your help in the first place.

Perhaps one of the largest pitfalls to the way most people initially think to use surveys to measure their performance is that the survey requires a significant part of the product to already have been designed and implemented in order to conduct the survey. One of the last things you want to do is to create your system and then only find out at the end if it isn’t meeting your expected levels of performance. So even if you use a survey for end validation, make sure you have another means for at least estimating your performance early in the design process.

**Tip 11:** **Taking time to really understand the needs and performance criteria of the challenge you’re trying to undertake is never time wasted as they influence almost every part of the design process.** Performance criteria are touched upon in a number of guides in addition to the ones mentioned at the beginning of this guide, because they can influence just about every part of the design process. The Delving Into Your System’s Needs guide provides some in-depth techniques for determining needs and performance criteria while working both with and outside of your stakeholders as well. Even if you don’t have the time or resources to follow all of these guides to the letter, the concepts are important and valuable to include in your design and will help you become an even better professional designer.

**~~Tip 12:~~** ***Not a Tip*, Do This. No, Really Make Sure You Do This!:Create a decision matrix for all of your performance criteria.** In the end, it is almost always valuable to arrange the identified metrics, their means of assessing/measuring their criteria, and the relative importance of the metrics into a decision matrix. The Decision Matrix guide provides valuable information on how to do this professionally as well as set and normalize metric scoring levels for the criteria you’ve identified.

It’s fairly common that the decision matrix is a deliverable that your stakeholder may likely not mention specifically (or even know enough to ask for, of even know what it is…) but your stakeholder will appreciate seeing it both during the development of your System and at the end. Hence, it is worthwhile to read the Decision Matrix guide *before* talking with the stakeholder as reading that guide will give you the information you need to do that decision matrix deliverable well.