Interview — *The Story Behind The Normalized Scoring System*

A conversation with Ed Carnes, developer of The Normalized Scoring System, about why he created it, how it works, and what he's learned from using it in real competitions.

Q1. Where did this come from? What were you seeing in competitions that made you think there had to be a better way to score them?

Ed: I've competed in and judged fiddle competitions for many years, so I've seen this from both sides — being the person standing on the stage and being the judge deciding what score to write down. And over the years, I've seen so many different scoring systems I couldn't even count them.

One thing I kept seeing was that even when you have good judges who know what they're doing, the numbers don't always work the way you think they do. Two judges can hear the same competition pretty much the same way and rank the players very similarly, but the actual numbers they write down can be very different.

One judge may score everybody pretty high. Another may be a much tougher scorer. Another judge may put only a point or two between players, while somebody else may put five or ten points between them. In other words, even though the judges are using the same scoring scale, they may use very different point spreads. When you start adding all those numbers together, those differences can affect the final results. I call this “point spread bias.” It can let a judge with a wider scoring range have disproportionate influence on the outcome

And then there's just bad judging. Anybody who's competed very long has probably seen some of that, too. No scoring system is going to turn a bad judge into a good judge, and bad judging can quickly destroy a competition.

But I kept thinking there had to be a better way to take the scores the judges give us and come up with a fairer result. That's really where The Normalized Scoring System came from. I wanted to make the scoring better without trying to change the way judges judge.

Q2: What exactly was wrong with the traditional method? If every judge is using the same scoring scale, why isn't simply adding or averaging their scores necessarily fair?

Ed: On a 100 point scale, one judge might keep most of their scores between 70 and 90. Another might use 60 to 95. And another judge may use almost the entire 100 point spread. I've seen some pretty extreme examples. I've seen a judge give one contestant a 92 and another contestant a 9 when there really wasn't anywhere near that much difference in their performance or skill. When a judge does that, they've effectively taken control of that category. The point spread they've created can be so large that it overpowers the opinions of the other judges.

Then there's this reluctance people have to give anybody a 100. Giving somebody 100 out of 100 doesn't mean they're perfect. It means they earned all the points you had available to give them based on what you were judging. If you're judging a really good group of players and the competition is close, sometimes you have to go high in the point spread to separate them. But a lot of people just don't want to give a 98, 99 or 100. I think some of that probably comes from school. We're taught that 100 means perfect, and that's hard to get out of your head.

Interview — *The Story Behind The Normalized Scoring System*

And then there's the skeleton in the closet that people don't like to talk about. A judge can manipulate the outcome with the way they use their scores — and I'm not talking about doing it accidentally. I've seen it happen both as a contestant and as a judge.

But here's something else I noticed over the years. Most people don't have much trouble telling you who they thought was first, second, third, fourth and so on. What they struggle with is figuring out how many points should separate those people. And that really stuck with me. If the judge knows the order they think the contestants should finish, why should the final result be distorted because one judge thinks first and second are one point apart and another judge thinks they should be ten points apart?

That was another big driver behind developing The Normalized Scoring System.

Q3. Was the problem really the judges? So are you saying judges are scoring incorrectly?

Ed: No. And that's an important distinction. Can you have bad judges? Absolutely. Can you have biased judges? Absolutely. I've seen both. But that's not really the problem The Normalized Scoring System was designed to fix.

You can put three very good, very experienced judges in the same room, let them listen to exactly the same contestants, and all three can do a perfectly good job judging the competition — and still use their numbers completely differently. One judge may be very conservative with points. Another may score everybody higher. One may put two points between first and second, while another may put ten points between them. That doesn't necessarily mean any of them judged incorrectly. They may even agree on the order the contestants should finish.

The problem comes when we take those very different sets of numbers and simply add them together. Now the way a judge uses numbers can have as much influence on the outcome as who they actually thought performed the best. That's what I wanted to get away from. My feeling is pretty simple: judges should judge. They shouldn't have to change the way they naturally evaluate contestants because of the mathematics we're going to use afterward.

The Normalized Scoring System takes the scores each judge actually gives us, determines how that judge ranked the contestants, and then gives each judge an equal voice in determining the final result. It doesn't try to replace the judge's opinion. It tries to represent it fairly.

Q4: So what does NSS actually do? Explain The Normalized Scoring System to me without getting too deep into the mathematics. What does it actually do with those scores?

***Ed*:** I have a story I use that shows both the problem and what The Normalized Scoring System does. Let's take three judges — Benny, Major and Clark — who are judging a horse show.
In the finals, we have two horses: a white horse and a brown horse. Each judge can give each horse up to 100 points based on how well it walks around the arena. At the end, we'll add the three judges' scores together, and whichever horse has the most points wins.

Interview — *The Story Behind The Normalized Scoring System*

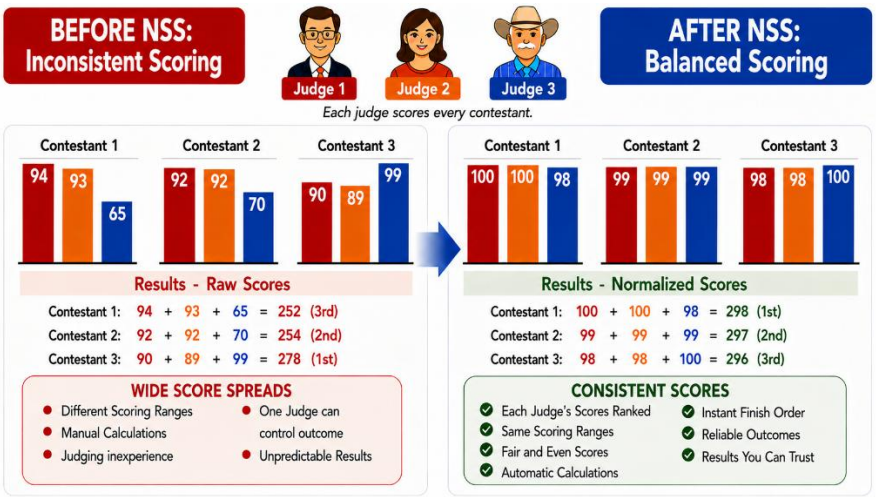
The white horse goes first. Benny really likes it and gives it a 95. Major likes it too and gives it a 93. Clark isn't quite as impressed and gives it an 80. Then the brown horse goes. Benny likes it, but not quite as much as the white horse, so he gives it a 93. Major feels pretty much the same way and gives it a 90. But Clark likes the brown horse better and gives it a 90.

Now we add up the scores. The white horse gets 268 points. The brown horse gets 273. So the brown horse wins. But wait a minute. Benny picked the white horse. Major picked the white horse. Clark picked the brown horse. Two of our three judges thought the white horse was better that day, but the brown horse won. Why? Because Clark used a much bigger point spread between the two horses. Benny had two points between them. Major had three. Clark had ten. Clark's ten-point spread was enough to overpower the opinions of the other two judges. That's the problem.

Now let's take those exact same scores and run them through The Normalized Scoring System. With NSS, we're going to look at how each judge ranked the horses. Whoever that judge had in first gets 100 points. Second gets 99. If we had more contestants, we would just continue down the scale. So Benny ranked the white horse first and the brown horse second. The white horse gets 100 and the brown horse gets 99. Major did the same thing. White gets 100 and brown gets 99. Clark ranked the brown horse first, so brown gets 100 and white gets 99. Now add those together. White horse: 299. Brown horse: 298. The white horse wins.

We didn't change Benny's opinion. We didn't change Major's opinion. And we didn't change Clark's opinion. We didn't ask any of them to judge differently. We simply took their original scores — what I call their “raw scores” — ranked them, and applied the same point spread to every judge.

That's really the heart of The Normalized Scoring System. It doesn't change how the judges ranked the contestants. It gives each judge's ranking equal weight in determining the final result. I will attach an illustration for the written transcript.



Interview — *The Story Behind The Normalized Scoring System*

Q5: Does normalization change a judge's decision? Some people may hear "normalization" and think you're changing the judges' scores or overriding their decisions. Is that what happens?

Ed: No. And that's something I want people to understand about NSS. The judge still makes the decision. NSS doesn't look at a score and decide the judge was right or wrong, and it doesn't try to change who that judge thought was first, second, third or fourth. It takes the scores the judge actually gave and uses those scores to determine how that judge ranked the contestants. Then it applies the same point spread to every judge's rankings.

So we're not changing the judge's opinion. We're making sure one judge doesn't have more influence simply because they used a bigger spread between their scores. And NSS doesn't force a competition to use normalized scoring, either. If a competition wants to continue using traditional scoring and simply add the judges' raw scores together, NSS supports that too. Some competitions have traditions or rules they don't want to change, and that's fine.

My goal wasn't to build something that tells competitions, "This is the way you have to do it." I wanted to build something flexible enough that a competition can use the scoring method that works for them — including normalized scoring when they want the advantages it provides. So whether you're using normalization or traditional scoring, the judges still judge. NSS handles what happens after they put the scores down.

Q6: What difference can it actually make? Can The Normalized Scoring System actually change the final placement of contestants compared with simply totaling the raw scores?

Ed: Absolutely. In fact, that's one of the reasons I developed it. Go back to our horse example. Under traditional scoring, the brown horse won even though two of the three judges had the white horse ranked first. One judge used a much larger point spread, and that was enough to change the outcome. NSS doesn't change how any of those judges ranked the horses. It takes the rankings that came from their actual scores and applies the same point spread to each judge. Now each judge has the same influence on the final result.

So yes, NSS can change the final placement compared with simply adding the raw scores together. But it's not changing the result just for the sake of changing it. It's removing the advantage or disadvantage created by the different ways judges use points.

To me, that's the whole point. If three judges are judging a competition, I want all three of their opinions to count equally. I don't want one judge to have more influence simply because they happened to use bigger numbers or a wider point spread.

Interview — *The Story Behind The Normalized Scoring System*

Q7: Is NSS just a scoring formula? When someone hears 'Normalized Scoring System,' they might assume it's just a different formula. Is there more to the system than that?

Ed: Scoring is the heart of NSS, and that's where it started, but it has grown into a lot more than that.

The scoring system gives the coordinator both the traditional raw-score results and the normalized results, so you can actually see the difference the scoring method can make. And it's flexible. It can accommodate anywhere from one to five judges, different scoring methods, and even competitions that want to throw out the highest and lowest scores. There are actually 14 different scoring options built into the system. But once I started developing NSS, I began looking at some of the other problems we deal with when we're trying to run a competition. One of the biggest is time.

If you've ever run a contest, you know how difficult it can be to figure out when you're going to finish. You may not know how many contestants you're going to have until registration closes. You have different categories, different numbers of tunes or performances, breaks, meals, awards — and you may have a hard deadline for when you need to be finished.

So I developed what I call the "Duration Calculator". You can use it before the contest to estimate how long the event is going to take, and then use it during the contest to see whether you're running early or late. It can help you decide whether you've got time for a break or a meal, or whether you need to make an adjustment somewhere to finish when you need to finish. It became useful enough that I actually made the Duration Calculator available separately for competitions that don't use NSS scoring at all.

I also developed a "Registration Module" to manage contestants, categories and entries. Once registration is finished, it can automatically load the contestants into the scoring files. And just like the Duration Calculator, it can also be used on its own.

One of the newer pieces is the" Judge Score Tracking Worksheet". If judges have a PC or Mac available, they can enter their own scores and notes as they judge. The worksheet automatically ranks their scores as they go, so they can see where they've got everybody and how many contestants are going to advance to the next round. That can be especially helpful when you've judged a lot of people and you're trying to remember exactly where everybody fits.

So yes, scoring is still the heart of The Normalized Scoring System. But what I've really tried to build around it is a set of tools that helps a contest coordinator plan the competition, run it, score it and get the results out — without making the job harder than it already is.

Interview — *The Story Behind The Normalized Scoring System*

Q8: Who is this really for? What kinds of competitions could benefit from The Normalized Scoring System?

Ed: Well, I developed The Normalized Scoring System around fiddle and bluegrass competitions because that's the world I've been involved in for most of my life. But the problem NSS addresses certainly isn't limited to fiddle contests.

Any time you've got several judges scoring the same contestants and then combining those scores to determine the winner, you can potentially have the same problem we've been talking about — different judges using their points and point spreads differently.

So there are a lot of possibilities. It could be other music competitions, talent shows, dance competitions, horse or livestock shows, barbecue or cooking competitions, photography and art contests, pageants, science fairs, public-speaking or presentation competitions, and other judged events where people are assigning numerical scores.

There are also sports such as gymnastics, diving, figure skating and cheerleading that deal with some of these same judging issues, although many of those already have very specific scoring systems established by their governing organizations. I'm certainly not suggesting NSS should replace those.

Where I think NSS can be particularly useful is with independent competitions — festivals, associations, fairs, community events, regional championships and other organizations that have the freedom to decide how they're going to score their competition.

And NSS doesn't require every competition to be exactly like a fiddle contest. The scoring system can accommodate different categories and different ways of setting up a competition.

The basic question I would ask is pretty simple: “Do you have multiple judges assigning points to contestants and then combining those points to determine who wins?”
If the answer is “yes”, then it's worth looking at whether differences in the way those judges use their points could be affecting your results.

And if they are, The Normalized Scoring System may be able to help.

Q9: You've now had The Normalized Scoring System used in actual championships. What have you learned from seeing it used in the real world?

Ed: Probably the biggest thing I've learned is that it works.

Overwhelmingly, I've seen what I feel are better and more consistent results — results that more closely reflect the collective opinion of the judges instead of being affected by how one particular judge happened to use their points.

Interview — *The Story Behind The Normalized Scoring System*

I've also seen a huge improvement in the actual operation of a contest. There's no longer somebody sitting at a table doing manual calculations on scoring sheets. As the scores are entered, the calculations are being done automatically, you can see what's happening in the competition, and once the last scores are entered, the results are available almost immediately.

And real-world use has caused me to keep adding things. Tie-breaking is a good example. Instead of having somebody trying to work through a tie after the contest is over, NSS can automatically apply the tie-breaking rules the competition has already decided it wants to use.

At the Grand Master Fiddler Championship, I even developed a way to take the contestants in a category and automatically build the slides we display while they're playing, showing their name and where they're from. That has been hugely popular with the audience, and it's turned out to be a great photo opportunity for the contestants and their families.

Something else I've learned that I didn't fully appreciate when I started is how much information the system can give you about judges. Over time, you can begin to see how different judges use their scores. You can see tendencies toward particular styles of playing, judges who use very wide or very narrow point spreads, and whether a particular group of judges gives you a good mixture of viewpoints. I'm not talking about whether somebody is honest or dishonest. I'm talking about understanding how people judge. That's extremely useful when you're trying to put together a good judging panel.

The Duration Calculator has also become much more important to me than I originally expected. For years, a lot of running a fiddle contest was basically gut feeling. You'd look at how many contestants you had, look at the clock and try to figure out whether you were going to finish on time. I remember contests when I was growing up that didn't finish until the early morning hours. That's hard on contestants, families, judges and volunteers — and then people still have to drive home. Now we have a way to actually manage that instead of guessing.

When I originally designed NSS, I looked at some of the more complicated contests I knew about. The national contest in Idaho, for example, publishes estimated performance times within categories. I wanted NSS to be able to do that kind of planning, but also to manage the timing of an entire competition — including competitions with multiple stages running at the same time. And it does. In fact, we've been using it to manage multiple stages at Tennessee Valley Old Time Fiddlers Convention in Athens, Alabama and it works very well.

And maybe the most important thing I've learned is that NSS is never really finished. Every time I use it, or another competition uses it and I help them, I watch what happens. What worked? What was difficult? What did somebody have to do manually that the system ought to be doing for them? Then I take those lessons and put them into NSS.

We now have it being used at national, regional, state and individual competitions, and we're continuing to see interest from other contests. So the real-world use hasn't just proven the concept to me. It's what keeps making the system better.

Interview — *The Story Behind The Normalized Scoring System*

Q10: Where can someone learn more If I'm running a competition and I'm thinking, "We may have exactly this problem," what should I do next?

Ed: The easiest thing is to take a look at The Normalized Scoring System on the Carnes Group website. We've put quite a bit of information there explaining how NSS works and the different tools that are available. You can see examples of the actual screens and reports, pricing and other details, and you can read what a couple of highly respected people in the fiddling community have to say about the system.

The website is: https://www.carnesgroup.com/normalized-scoring-system/

And if you have questions, or you're wondering whether NSS would make sense for your competition, just email us at: nssinquiries @ carnesgroup.com

I'm always happy to talk with competition organizers about what they're doing now, what problems they're trying to solve, and whether NSS might help. And if there's something you're doing in your competition that makes sense for us to add to NSS, we're always open to that too. A lot of the system has grown from seeing what competitions actually need.

You don't have to know that NSS is right for your competition before you contact me. That's what the conversation is for.