How two graduate students uncovered a critical error in autism screening guidelines

BY **LAURA DATTARO**

22 JUNE 2021

When clinicians suspect someone has autism, they often turn to screening tools — standardized questionnaires or checklists that measure autism traits. If a high score corroborates their hunch, they refer the person to a specialist for a diagnostic evaluation.

These screening tools typically come with **guidelines** to aid clinicians in their use. But the guidelines also inform researchers who are screening potential study participants — a use that ultimately shapes how researchers define autism, says **Punit Shah**, associate professor of psychology at the University of Bath in England.

In the United Kingdom, the **National Institute for Health and Care Excellence** (NICE) authors a variety of screening-tool guidelines, among others. And earlier this year, two of Shah's graduate students discovered that the NICE guidelines for one particular test, called the Autism Quotient-10 (AQ-10), **were wrong**. They **published their findings** in *The Lancet Psychiatry* in April.

Last week, on 14 June, NICE **issued a statement** confirming that they would correct the error. "We were concerned that the guidelines weren't going to be changed," Shah said after the announcement. "We were consequently really pleased to see that NICE have taken the issue seriously and have dealt with it so promptly."

Spectrum spoke to Shah and one of the graduate students, Lucy Waldren, in May about how they spotted the error, why they're spreading the word about it and what needs to happen next.

Spectrum: What are the NICE guidelines, and what was the error you found?

Punit Shah: NICE is a U.K.-based nongovernmental organization. It's an independent, really well-

respected body. The NICE guidelines are often used and adapted by many other countries. So what we do here in the U.K. and what NICE recommends has quite far-reaching consequences.

The guidelines are recommending that people who score above 6 on the AQ-10 should be referred for further assessment and potential clinical diagnosis of autism. That's quite different to what was suggested in the original research, which includes a score of 6.

S: How did you find this error?

Lucy Waldren: Another Ph.D. student, **Rachel Clutterbuck**, was reanalyzing some data that used the AQ-10. The original paper in 2012 **defined the AQ-10** from a larger **Autism Spectrum Quotient**, which is a 50-item measure. Rachel noticed that the original paper and the **one from NICE** had a different cutoff.

We thought that was really odd. So we went and dug through the justification they gave for how they came up with their cutoff score, and we couldn't really find a reason why they chose the 7 over the 6. So we think they made a mistake somewhere.

S: What was your initial reaction to that discovery?

LW: I think just surprise. We kept double-guessing ourselves and thinking, "This surely can't be a thing. They've been around for nearly 10 years, and there's been a mistake in this key guideline." It was almost disbelief.

PS: We use the AQ-10 and other measures of autism traits in our research. And we think, broadly speaking, they've been really helpful for autism research. It's not as if we were going out to attack it or find problems with it. So when Rachel came across this issue, I thought she was making the mistake. I asked her for a screenshot and for Lucy to dig into it.

S: What did you decide to do about it once you were certain of the error?

PS: At first, we just kept checking and double-checking and triple-checking it. We must have checked this dozens of times, thinking, "Well, are we sure here?" Once we were satisfied from an academic perspective that these mistakes can happen, and Lucy did a really good job of interrogating all the appendices and all the nitty gritty of what the NICE guideline group had looked at, we then made the decision at that point to write something on this and send it to an academic journal as soon as possible to just try to make people aware of this.

S: Did you reach out to NICE before you published the paper?

PS: Yes. Our institution did contact NICE to let them know that we had found the error and that we were going to be writing about it. But at that point, at least, NICE hadn't replied to us at all.

S: What was the reaction to the paper, and what's been happening since it came out?

LW: I went to the NICE public annual general meeting, and I raised it as a question that NICE declined to answer. Since then, they have said that they're reviewing the guidelines, which is good.

PS: They never contacted us directly. We got quite a lot of media attention here in the U.K., and a little bit in the United States. There was a big response from the autism community on Twitter. In terms of the NICE response, though, it's been quite underwhelming. They've said things like, "We don't really know what's happened," or, "As part of our broader review of these guidelines, we're going to be looking into this."

S: What is the impact of this one-point error over the past 10 years?

PS: Even though it sounds trivial on the face of it, a difference of just one point on a 10-item scale is actually quite substantial. Because NICE had recommended a higher cutoff, it makes the measure less sensitive. There will be people scoring a 6, many of whom will have autism, who may not have been referred for further diagnosis.

LW: We can't know how many people may have been affected by this. If people have missed out on further assessment and a diagnosis because of it, they then won't have had access to the same level of care and support, including financial support. That will have a knock-on effect on their mental health from the lack of support. They also don't have the validation of receiving the diagnosis that they feel like they should. So it could have had a really serious impact on people's lives over the years.

From a research point of view, the measure is used as a cutoff for participants. So data might have been lost as a consequence of people using the different cutoff as well.

PS: Lucy now is more systematically looking at this. Other autism spectrum trait measures have been quite noisily or messily applied in the literature. We found a couple of papers in which they used the incorrect cutoff to include or exclude people in their studies. I think at some point, those researchers will need to revisit their data or have a think about trying to replicate their basic effect. Some people have been receptive to that idea and have checked whether they've implemented the wrong cutoff. But generally, people have been not as forthcoming about talking about any mistakes they've made.

S: What are you planning to do next?

LW: We're looking more generally at the use of autism trait measures in research and trying to see what future steps may be needed to improve practice.

PS: There's been a long-standing debate about the pros and cons of these measures. It's only in the past few years that people have started **scrutinizing them** a bit more carefully. Even though we're critical, to an extent, of some of these measures, we do use them in our research. They have been incredibly helpful to advance autism research, generally speaking. But the next steps are to involve more autistic people and these more rigorous statistical techniques in scrutinizing them and improving them.