The Next Questions: Who, What, When, Where, and Why?

Steven N. GOODMAN

There is little doubt that the ASA statement will stand the test of time better than many of the textbooks and journal articles purporting to teach the practice of statistics to scientists. And that's what we should be afraid of. As the introduction acknowledges, almost none of the statement is new. That is an understatement; some of these same principles were stated or argued about at least a century ago, with many reminders between then and now. So the question we must ask of ourselves is how could this have happened, and what can we do to change it? How could it be, almost 100 years after the formulation of the Fisherian and Neyman-Pearson approaches to statistical inference, that a body is eminent as the ASA needs to step in and remind scientists how to define an index as ubiquitous in scientific investigations as the P-value, and how to use it or other indices properly? And what will prevent us from dusting this same statement off 100 years hence, to remind the community yet again of how to do things right? If we dont start focusing on that, the most likely outcome of this effort is that the statement will literally stand the test of time, in being as needed in the next century as it is today.

I suspect that for readers outside the statistical community, the most surprising thing about the statement will be the degree of passionate commentary that it will no doubt engender. The fact that statisticians do not all accept at face value what most scientists are routinely taught as uncontroversial truisms will be a shock to many. But if we are to move science forward, we must speak to scientists. How many of these comments can be understood by the broad scientific community? If that proportion is small, we have a clue why we are in the situation we are in, and what needs to change.

In that spirit, I would like to pose those fantastically useful questions that the media is trained to ask in all their reporting: *Who, what, when, where, and why*? These questions can help force us to explore our own responsibility in not just articulating principles, but helping scientists to get things right.

Who?—Exactly who will push this agenda forward? How much can statisticians do alone, and how much must they do in partnership with other scientists? And when we say some statisticians alone, do we mean the thought leaders who write mainly about theory, or the applied statisticians in the field? Or the textbook writers? Or those who labor to teach statistics within specific disciplines? Or the department chairs? How much should we include other quantitative scientists and non-statistical scientists as allies and indeed champions?

What?—The question is two-fold. First, what are we actually recommending scientists do? We say in the statement that they should combine evidence from p-values, confidence intervals, Bayesian measures, false-discovery rates, or other measures with features of the design, conduct, and plausibility of what we are studying. Exactly how our scientists supposed to do that? Where are all the textbook examples? Where are the examples in the published literature? The statement is very good at identifying bad inferential behavior, but can we use it to figure out or dictate what is good? The fact is, bad behavior is often condoned and even encouraged within many scientific disciplines, and good behavior is neither taught nor role-modeled. What do we think would happen to a statistical colleague who tried to convince their collaborator that the $P = 0.01$ just obtained for a main finding was not sufficient to make a claim, because of the design of the experiment, the analysis, or the nature of the hypothesis? Where are the examples they can point to? If we are to make such recommendations, we need to figure out what to tell or teach people.

The second "what" is what is the profession and professionals of statistics supposed to do to maximize the adoption of these principles? Write new books? New software? Change their courses? Write in disciplinary journals? Write more blogs? Post YouTube videos? Tweet more tweets? Reward good behavior? All of the above, or A,B and D? How do we promote the complete and transparent reporting that is recommended, not just in spirit, but in reality?

When? —When do we start? For how long? How will we evaluate when enough has been accomplished?

Where?—This overlaps with the "what?" In what venues and platforms does the agenda have to be pressed forward? In funding review sections? In journal clubs? On journal pages? Which journals? At the ASA? In the halls of congress? In industry? In the OSTP (Office of Science Technology Policy), the NSF, the NIH, the FDA or the DOE? Where are the key policy leverage points, and what should those policies be? Is this amenable to a policy-level fix?

Why?—Why do this? We are all busy; who has time to reform science? Or even their co-author? Or a post-doc who asks for statistical advice the evening before the conference abstract is due? One big "why" is related to research reproducibility, which is the ring by which we can turn the nose of science. When the directors of the National Academy of Sciences, the NSF, and the NIH, as well as industry leaders all tell us that research reproducibility is high on their agenda to improve, and we can see the clear connections between the ASA principles and those concerns, we have at least our collective "why," but perhaps not the individual "why?" The ASA and other groups need to figure out how to push this issue forward in a way that statisticians see as part of their professional responsibility to further this agenda,

Online discussion of the ASA Statement on Statistical Significance and P-Values, *The American Statistician*, 70. Steven N. Goodman, Stanford University, Division of Epidemiology, and Meta-research Innovation Center at Stanford (METRICS) (Email: Steve.goodman@stanford.edu).

and to reward them for it. Perhaps some awards from the ASA for improving the practice of statistics along the lines of this policy statement could provide the role models and inspirations for others to emulate.

There are many more things that could be said about what needs to be done and who needs to do it, but this much should be clear; what follows this statement is as or more important than the statement itself. It must be someone's or some entity's designated or adopted responsibility to carry the essentials of the ASA principles forward, otherwise nothing will change. We need to formulate a vision of what success looks like, and how we will get there. If not, we can start drafting the language of the 2116 ASA statement tomorrow.