Sound the Gong: In Defense of an *Explicandum*

Ruobin Gong, Rutgers University, responds to Chris Burdzy's Presidential Address, delivered at the 2022 IMS meeting in London, and printed in the previous issue.

> **"When I use a word," Humpty Dumpty said to Alice, in rather a scornful tone, "it means just what I choose it to mean—neither more nor less."**
> Lewis Carroll, *Through the Looking Glass*

In his Presidential Address, Professor Chris Burdzy urged the IMS community to "stop using the term 'subjective' in reference to any part of statistics." He further professed emotionally that "I love Bayesian statistics because it is objective. It hurts my feelings when people suggest that Bayesian statistics is subjective or partly subjective." This essay argues that Burdzy's call falls prey to precisely the danger it sets off to apprise. It relies on a hazy interpretation of the word "subjectivity" that does not adequately reflect the complex evolution of its meanings through history, and is incongruent with the modern standards of responsible scientific practice.

Subjectivity is a loaded word. In the context of science, subjectivity carries negative connotations, two of which being "unfettered freedom" and "contaminant of rational thinking" (Burdzy, 2022, pp. 1). By the sound of it, neither quality is quite what we should expect from a respected scientist. Further, much unlike subjectivity which the audience finds detestable, its counterpart objectivity is often met with a lauded reception.

That may all be true. Setting emotions aside, however, are negative connotations a sufficient reason to call for the abolishment of a word from our vocabulary, and positive connotations reason for endorsement? Is subjectivity, a term as old as the human languages that index it in their dictionaries, no more than some fuzzy and terrible feeling that every scientist (data scientist included) is morally obliged to condemn?

Subjectivity and objectivity are contextually encumbered, emotionally charged, yet rarely understood nor explicated. Were we to protest the use of one and parade that of the other, one question would need clarification first. When we speak of subjectivity and objectivity in science, what do we exactly mean?

In a rhetorical sense, objectivity has long been considered one of the virtues in science. It is often discussed in tandem with other desiderata such as truth-seeking, error avoidance, and accuracy (e.g. Levi, 1967, Joyce, 1998). The meaning of objectivity has never been static nor plain. A quick foray into the history of science would tell us that both subjectivity and objectivity carry temporally dynamic and linguistically fluid meanings that are exemplified through centuries of scientific practice.

In *Objectivity* (2007), Lorraine Daston and Peter Galison discuss three competing notions of objectivity that are still at play today. The first kind of objectivity—let us label it as Objectivity$_1$—entails the *idealized* depiction of the object under study. A scientist's pursuit under idealized objectivity is to obtain a universal characterization of a class of objects. She distills that idealization from specimens of a same kind and renders them into a single summary that is nevertheless more perfect than any of them. Idealized objectivity instructs scientists to create exemplars that are devoid of imperfections (and even individualities) of its subjects.

The second kind of objectivity confers a nearly opposite meaning. Objectivity$_2$ commands the *mechanical* representation of the subject under study, in a way that is wholly detached from the idiosyncrasies of the observer rather than the subject itself. The observer is asked to be blindly faithful to her observational apparatus. Everything must be recorded exactly according to what is seen, heard, or otherwise measured: every dent, every blur, every bit of fallen dust or missing corner.

One begins to sense the intricacy of explicating objectivity and the challenge of separating subjectivity from it. On the surface, mechanical objectivity (Objectivity$_2$) is in better agreement with how the ordinary word "objectivity" resonates. Aligned with the Cartesian philosophical tradition, the mechanical definition emphasizes explicit external standards that strips away any human tampering that may cause unreliability. However, mechanical objectivity may be executed to a fault. In the absolute lack of intervention by the observer, every feature must be mindlessly preserved even if they're known or widely acknowledged to be artifacts or consequences of device malfunction. It also demands the scientist to break away from her own identity, which encompasses not only her biases (which it sets off to avoid) but also her informed opinions and educated perspectives. By contrast, idealized objectivity (or Objectivity$_1$) often provides great pedagogical utility as it allows a teacher to reveal most directly and efficiently to a student what she believes to be the "essence" of their subject. However, achieving idealized objectivity relies entirely on the observing scientist to decide what aspect of that she sees is, and is not, part of this essence. In this sense, idealized objectivity agrees surprisingly with the word "subjectivity" as is familiar to most. Thus, if we follow the simple predicate that "what objectivity is not, subjectivity is," we arrive at the conclusion that Objectivity$_1$ is subjectivity in relation to Objectivity$_2$, whereas Objectivity$_2$ is subjectivity in relation to Objectivity$_1$. That might seem absurd, but is not incorrect.

Continues on **page 5**

There is a third kind of objectivity. Objectivity$_3$ advocates for an *interpreted* characterization of the scientific object. It emphasizes the judicious injection of trained judgment when curating and utilizing observations obtained from the material world. Interpreted objectivity seeks a middle ground between idealized and mechanical objectivity, inheriting aspects of both that are conducive to productivity. It recognizes that trained judgment is a small yet necessary component in advancing science, a concept that invariably escapes standardized measurement. This did not deter the scientific community from endorsing interpreted objectivity and implementing dynamic mechanisms to ground and to balance the multiplicity of trained judgments. As Ted Porter (1996) argues, science is fundamentally a social activity. Scientific hypotheses and findings are shaped through personal exchanges, collegial discussions, and community evaluations. The peer-review system, for example, grants the ultimate judgment of right from wrong to individual scientists. While ridden with problems of its own, we have yet to find a competitive alternative solution that might supplant peer review.

The three kinds of objectivity are widely embodied by the prevailing scientific norms on an ever-changing landscape. Daston and Galison's expansive investigations are founded on a wide range of examples from cell biology to astrophysics. I add that these scientific ideals are as vividly illustrated by the histories of statistics and data science. Tracing back to the mid-16th century, statistics made its debut as the *science of the state* (Hacking, 1990). A utilitarian service to governments, our young discipline was tasked with *data curation*—births, deaths, illnesses, for the calculation of taxation and military recruitment. A pursuit of mechanical objectivity was evident through the ever-growing levels of detail and exactness of the tabulations. With the advent of social statistics, idealized objectivity took its turn and found its expression through *data reduction*, epitomized by Adolphe Quetelet and his "average man" (*l'homme moyen*; Quetelet, 1831, see Stigler, 2002). A single numerical summary is artificially construed to describe a group of people: crude and subjective to some, effective and objective to others. After statistics gained formal mathematical ground and developed multiple theories of inference—Bayesian, Frequentist, even Fiducial—it began to exemplify interpreted objectivity through advising the science of *data modeling*. At its finest, modeling is both the artful reduction of rigorously curated data and the rigorous curation of knowledge learned through such reduction. As Sabina Leonelli (2019) puts it, "data are forged and processed through instruments, formats, algorithms, and settings that embody specific theoretical perspectives on the world." It is a perfect showcase of how interpreted objectivity is both a continuation and a combination of mechanical and idealized objectivity. Data science today encompasses all of data curation, data reduction, and data modeling, and every aspect of it is simultaneously objective *and* subjective by nature.

Having taken on a broader and time-transcendent perspective, we see that subjectivity and objectivity are never diametrically opposed concepts, nor are they mutually exclusive. Not only do subjectivity and objectivity rely on one other to derive meanings, but as new scientific contexts form, they too morph into new concepts and pick up qualities that used to be associated with each other. Scientists who aspire to objectivity, however defined, cannot accomplish much without every so often calibrating their compass of inquiries against the respective subjectivity standards. Banishing subjectivity from objective science is as nonsensical as banishing zero from the laws of arithmetic. If subjectivity were gone, what is left of the objective ideal is like a tree with rotten roots, a ruler with faded graduations.

When a community decides that a particular word shall not be uttered, the good reasons are usually that it is obscene, offensive, or otherwise threatening to our collective interest. Subjectivity is none of those things. As debates surrounding objectivity and subjectivity populate the peripherals of modern scientific discourse, an informed discussion around their meanings, as well as how these meanings adapt to questions of our time, becomes a literacy requirement. A true scientific spirit confronts and conquers things that are foreign, ambiguous, or difficult to explain. To explicate what needs explication using a combination of factual evidence and sound reasoning is the responsible scientific practice.

Bibliography

Burdzy, K. (2022). IMS Presidential Address. *IMS Bulletin*, **51**(6), 1-4.

Daston, L., & Galison, P. (2007). *Objectivity*. Zone Books.

Hacking, I. (1990). *The Taming of Chance*. Cambridge University Press.

Joyce, J. M. (1998). A nonpragmatic vindication of probabilism. *Philosophy of science*, **65**(4), 575-603.

Leonelli, S. (2019). Data Governance is Key to Interpretation: Reconceptualizing Data in Data Science. *Harvard Data Science Review*, **1**(1). https://doi.org/10.1162/99608f92.17405bb6

Levi, I. (1967). *Gambling with truth: An essay on induction and the aims of science*. Knopf, New York.

Porter, T. M. (1996). Trust in numbers. In *Trust in Numbers*. Princeton University Press.

Quetelet, A. (1831). Research on the propensity to crime of different ages. Brussels: Hayez. Translated by Sawyer F. Test Sylvester. Cincinnati, OH: Anderson Publishing Co.

Stigler, S. M. (2002). *Statistics on the table: The history of statistical concepts and methods*. Harvard University Press.