

IN SEARCH OF SIGNIFICANCE  
(STATISTICAL AND OTHERWISE)

DAVID GLEN MICK  
University of Virginia

Hubbard and Ryan's article (2000 [this issue]) is a useful—albeit Sisyphean—exercise in reminding social science researchers to stop making nonsense about statistical significance testing (SST). Nearly two decades ago, in my own field of marketing research, Sawyer and Peter (1983) similarly reviewed the assumptions and half-truths surrounding the theory, language, and applications of SST. They also provided several remedial recommendations, including some of the same as Hubbard and Ryan's (replication-extensions, meta-analyses) and others that go beyond them (Bayesian hypothesis tests, strong inference hypothesis tests). But memories fade, and rocks roll back down hills.

Today's marketing researchers remain as enamored of SST as any other group of social science and business researchers. And why not? Finding nonrandom differences or relationships among groups and variables is more interesting than finding random differences and nothing at all. And, as it turns out, the degree of interestingness rises dramatically if the nonrandom differences or relationships are contrary to prior beliefs (I will return to this point momentarily). For these reasons, I think Hubbard and Ryan (2000) overstate their case in charging that SST is of "marginal scientific value." Indeed, it is vital to distinguish in any research enterprise which findings are likely due to chance and which are not. Without this foundation, it is pointless to inquire about the magnitude (effect sizes) of the observed differences or relationships (Levin, 1993).

Nonetheless, my experience as a prior associate editor and now editor of the *Journal of Consumer Research* suggests that Hubbard and Ryan's con-

---

The author of this comment is the editor of the *Journal of Consumer Research*.

Educational and Psychological Measurement, Vol. 60 No. 5, October 2000 682-684  
© 2000 Sage Publications, Inc.

cerns about too much importance and too sole reliance being placed on SST are certainly justifiable. I also concur with their viewpoint on replication-extensions as probably the singly most necessary and effective antidote to the SST addiction in social science and business research. I would add that this is especially true of external replications (i.e., from new independent samples) that also employ different but convergent operationalizations of the same focal constructs or behaviors.

My quibbles with Hubbard and Ryan (2000) are centered in their next-to-last paragraph, in which they briefly outline their recommendations for breaking the compulsion for SST. These include (a) pedagogical initiatives that require graduate students to conduct replication-extensions in their seminar papers and theses and (b) bolder efforts by journal editors to attract replication-extensions of previous important articles. Although in principal these suggestions have merit, they also have little chance of success without major modifications in the philosophy and criteria by which academic research achievement are determined and implemented at many institutions of higher learning. That is, to the question “Why aren’t more replication-extensions completed and published?” Hubbard and Ryan’s implicit answer is (a) because graduate students are not encouraged and do not learn to do so and (b) because journal editors do not champion replication-extensions. I have a different explanation.

It is common and understandable that we design most of our doctoral seminars—especially the ones focused on theoretical developments and empirical studies—as readings and discussions about research that has made substantial scientific advances. We want our Ph.D. students to appreciate and emulate the best of the human quest for knowledge. We hope that they too may one day publish research in esteemed journals that captures large and lasting attention in their fields. Moreover, if they do this, their prospects for academic fulfillment and security (a.k.a. promotion and tenure) will be high. Regarding the latter attainment, we should keep in mind that most academic employers continue to rely resolutely on citation indices and research awards to determine whose scholarship is most impactful.

These points bring us to the question of what constitutes interesting and impactful research. As Murray Davis (1971) argued, the intrigue and impress of any given research project is a function of the extent to which its claims and findings challenge the a priori beliefs of its audience. The prime beliefs to contest are those that are widely and moderately held (but not calcified). For instance, convince the audience that what was thought to be complex is, in fact, simple (or vice versa); persuade them that what was assured to be negatively related is, in fact, positively related (or vice versa); prevail on them that what was resolved to be important is, in fact, trivial (or vice versa). These are the themes and legacies of the Chomskys, Durkheims, Freuds, Marxes, Meads, Skinners, and hundreds of other renowned scholars. Replication-

extensions are not. Most researchers know this tacitly, whether they have read Davis or not. In the end, junior and seasoned scholars have much less to gain in prestige or paycheck from conducting replication-extensions, and, consequently, journal editors continue to receive few of these submissions to review, let alone to advocate.

To curtail the historical tide of SST, Hubbard and Ryan's call for more pointed graduate education and more affirmative journal policies on replication-extensions is logical. Unfortunately, it is also insufficient and probably futile. Without serious broad changes in the written and unwritten values that underlie not only the means by which research is judged and revered but also the very goals of social and business science, the likelihood that replication-extensions will one day equal or supplant SST is less than .05. Of course, whether any value changes should take place in scientific pursuit among social and business researchers is another matter for another article. In the interim, I suspect that most of us will keep searching for significance (statistical and otherwise).

### References

- Davis, M. S. (1971). That's interesting! Towards a phenomenology of sociology and a sociology of phenomenology. *Philosophy of the Social Sciences*, 4, 309-344.
- Hubbard, R., & Ryan, P. A. (2000). The historical growth of statistical significance testing in psychology—and its future prospects. *Educational and Psychological Measurement*, 60, 000-000.
- Levin, J. R. (1993). Statistical significance testing from three perspectives. *Journal of Experimental Education*, 61(4), 378-382.
- Sawyer, A. G., & Peter, J. P. (1983). The significance of statistical significance tests in marketing research. *Journal of Marketing Research*, 20(May), 122-133.