

An Interview with Gene V. Glass

conducted by Daniel H. Robinson

Gene V. Glass is presently Regents' Professor of both Educational Leadership & Policy Studies and Psychology in Education at Arizona State University. He won the Palmer O. Johnson Award for best article in the *American Educational Research Journal (AERJ)* in both 1968 and 1970. He served as President of the *American Educational Research Association (AERA)* in 1975, co-editor of *AERJ* (1984–1986), editor of *Review of Educational Research* (1968–1970) and *Psychological Bulletin* (1978–1980), is executive editor of the *International Journal of Education and the Arts* (since 2000), and serves as editor of *Education Policy Analysis Archives* (since 1993) and *Education Review* (since 2000). Dr. Glass has also served on the editorial boards of 13 journals and has published approximately 200 books, chapters, articles, and reviews.

Dr. Glass is perhaps best known for his role in the development (starting in 1975) of the quantitative research synthesis technique known as meta-analysis, which has had a huge impact on the field of educational research. His work with M. L. Smith on the meta-analysis of psychotherapy outcomes was cited in *Forty Studies that Changed Psychology* (1999). An ERIC search shows well over 1,500 articles on meta-analysis since 1975. Most would agree that meta-analysis has certainly had a huge impact on the field of educational research.

I conducted this semi-structured interview entirely by e-mail using an original list of questions I developed to include topics Dr. Glass has written about, such as

The Research News and Comment section publishes commentary and analyses on trends, policies, utilization, and controversies in educational research. Like the articles and reviews in the Features and Book Review sections of *ER*, this material does not necessarily reflect the views of AERA nor is it endorsed by the organization.

electronic journals and meta-analysis, and others he has been silent about but ones I was sure he would have interesting responses to (e.g., the federal government's recent push for randomized experiments, the statistical significance testing controversy, and the rumored imminent breakup of AERA). I sent the questions to Dr. Glass in December 2002 and received his responses by February 2003, with follow-up questions sent (regarding the publishing of "garbage" and what we should do as educational researchers) and subsequent responses received in March 2003.

A few words may be in order to explain this interview and my role in it. I first made Gene Glass's acquaintance in 1989 while I was a graduate student at Arizona State University. It was no secret that Dr. Glass held many unconventional ideas about educational research and its methods. However, it was nearly impossible to find them in print and frustratingly difficult to provoke him to express them in public. The formal voice of the typical scholarly monologue may not be the nearest we approach to the truth.¹ In the spirit of digging behind the all too carefully circumscribed rhetoric of the academic set piece, as well as hoping that others are provoked to speak more plainly about our business, I present our exchange.

Randomized Experiments, Statistical Significance Testing, and the Government

DR: What are your thoughts about the recent push for randomized experiments in educational research that the federal government is calling for?

GVG: I wouldn't take seriously epistemological questions adjudicated by the federal government. I am sympathetic with the FDA's requirement of randomized clinical drug trials. But that hardly justifies the reliance on a single method for the verifi-

cation or discovery of claims in the soft sciences. And where is the randomized experiment that proves that randomized experiments are the royal road to truth? There is none, of course. Either the primacy of randomized experiments was arrived at through a process of a priori reasoning (as R. A. Fisher improved upon J. S. Mill by operationalizing the joint method of agreement and disagreement via probabilistic approximation) or else someone in the Bush administration thinks that phonics is backed up by randomized experiments and whole language instruction isn't. Ask Reid Lyon, chief of the Child Development and Behavior Branch of the National Institute of Child Health and Human Development at the National Institutes of Health. If the federal government wishes to be consistent (a dubious assumption), then they will have to back off their policies on smoking, coal dust, speeding on interstate highways, and a host of other things that have never been verified by randomized experiments.

DR: You have managed to avoid getting involved in the statistical significance testing controversy that has emerged as several ping-pong battles in the journals. Why the avoidance? You spoke about it at an AERA Division D symposium in 1999. Have you even paid attention to what has been said in these exchanges? Do you predict the field will be influenced at all by this controversy?

GVG: I don't want to mislead anyone who might take my silence as indicative of some displeasure or lack of interest in these issues. Rather, it should be taken as an indication that I have opinions about them and little confidence that my opinions are right. And these are issues on which I can quickly arrive at a state of confusion. Suffice it to say, perhaps, that my opinions on these matters are the result of an experience I had many years ago as a

member of the National Assessment of Educational Progress Analysis Advisory Committee. ANAC, as it was known, was a very active group that met several times a year to go over the NAEP results and try to make sense of them. In actual fact, it was a group of statisticians and measurement types who sat and watched John Tukey pour over stacks of computer printouts and discover things that the rest of us were blind to. John would scribble numbers wildly on yellow pads and arrange them in squares and smooth them and trim them and suddenly announce a finding that no one else had a prayer of seeing before his revelations; and as soon as he pointed it out, everyone could see it, too. One ANAC member of a particularly traditional stripe kept insisting that one could not do such things unless one had first stated a “model.” That person left ANAC of his own choosing. I began to question whether the approach to statistics that I had been taught and taught to hundreds of others in turn was really what data analysis should be about. There was little talk at ANAC meetings of significance levels or *p*-values. Now one might protest and say that Tukey was working with population data, a census in fact, for which traditional inferential techniques would have been superfluous. But the opposite was true. NAEP was the closest we have ever come in education to a legitimate sample survey, and yet the inferential techniques with which the journals and textbooks are filled played little part. Data analysis was exploration and discovery.

Another encounter with Tukey had a great influence on me. We were “seat mates” on a flight back from an ANAC meeting on Cape Cod to Denver in about 1977. John used to spend several weeks each year in Boulder at the National Center for Atmospheric Research analyzing data on the putative destruction of the ozone layer by fluorocarbons, and I lived in Boulder. (The only time his friends can recall seeing him in a tie was on the evening news in the late 1970s when he was announcing the National Academy of Sciences’ conclusion that there was insufficient evidence to ban fluorocarbons.) I was struggling through a crossword puzzle and John asked, “What are you stuck on?” “A five-letter word for a Greek market,” I replied. “Agora,” he said. (To myself: “Damn! I should have

thought of that! Agoraphobia, sure. He must think I’m really stupid.”) Emboldened by this display, perhaps, he turned to more important things. He had actually read some of the stuff I had recently done on meta-analysis, and I am proud to say he was intrigued. “What’s the biggest problem you’re having with this stuff?” he asked. “I can’t make up my mind about what role inferential stats should play. Nothing is sampled from anything. There aren’t any ‘populations’ to speak of. If the population is really hypothetical itself and any inferences are to a ‘population from which these studies might have reasonably been sampled,’ then the population is just the sample writ large and what do we need the inferential methods for?” I think I said. He didn’t hesitate to respond, which now leads me to believe that these qualms were not in the least unfamiliar to him, and may have once been shared. “What seems to be the biggest source of controlling variation in your bunch of effect sizes?” he asked. “I don’t know what you mean,” I confessed. “Do the results seem to vary most across types of patients, or the year in which the study was published, or the type of instrument used to measure the outcome, or what?” he queried. I had to think a bit, and finally I suggested tentatively that they seemed to vary most according to which researcher was doing the study. Jones always got nice big effect sizes around 1.0 and Smith always got effects around .50. “Then jackknife your findings on ‘researcher,’” he replied. “Agora!” I said to myself. “Jackknife on what’s causing the variation. Simple. I’m still thinking about what he said.” I’m thinking that the act of interpreting data is so complex—so very much more complex than our models—that about the only way to guard against “delusion by variation” is to split databases in half or thirds or whatever and cross-validate anything you think you see in one half on the other half. Perhaps it is true that the future of what we thought of as inferential statistics lies at the end of a path marked by methods like the jackknife and bootstrapping techniques so unfamiliar to most that they cannot be mentioned without a reference (Efron, 1979).

The Hegemony of Scholarly Journals and Organizations

DR: Could you talk about your withdrawal/avoidance of some of the journals in terms of

publishing your ideas/work in the last few years. Some of your best and most interesting papers are published on your Website rather than in any journals (which is why I wanted to capture some of your thoughts in ER). I know you are disenchanted with the journal review process and I was hoping you might comment on that issue. Ten years ago, you spoke at a session at the annual meeting of the Midwestern Educational Research Association about electronic journals and predicted that they would soon overtake paper journals. The prediction hasn’t exactly come true, although there are many more cyber-journals now than in 1994. Do you have any idea why there has been so much inertia?

GVG: Actually, I’m not particularly disenchanted with peer review. I just don’t see any necessity of it in many circumstances in this day and age. In the hard sciences, xxx.lanl.gov—the e-print Archive for math, physics, and computer science—has shown the superfluity of traditional peer review for the most scientific of the sciences. The likes of philosopher Joseph Ransdell have raised important points regarding the proper place of peer review in the sciences that apply with greater force, it seems to me, to the peculiar intersections of science and practice like educational research. (In particular, see Ransdell’s and Harnad’s exchange in “Scholar’s Forum: A New Model for Scholarly Comment” at <http://library.caltech.edu/publications/scholarsforum/>.) I don’t mind publishing in traditional journals; I just don’t see much need for it anymore. Some papers that I have made public through my own Website are downloaded a half dozen times a day, often from places that have no access to the traditional journals. That’s reaching a wider, larger audience than paper journals reach. Now, the questions in most people’s mind are: Won’t there be chaos if everybody “publishes” anything they want? How will we be able to separate wheat from chaff, truth from error, pure gold from garbage? There are answers, of course, and more questions.

Did I miss on my prediction in 1994 that scholarly communications would soon be taken over by “cyber-journals”? I guess I did. I underestimated the strength of the vested interests that certain individuals, companies, and organizations have in restricting the communication of scholarly information. Yes, I said “restricting.” Certain

publishers of scholarly journals are realizing profits of between 25% and 50%! That's in a world in which profits of 10% are phenomenal and under 5% are common. These companies have played academics for chumps for decades, but the game is slowly coming to an end—and not because of the efforts of individual scholars caught in the publish-or-perish trap, but because administrators and librarians are forcing the termination of journals that they have purchased for years, which no one reads. My library at Arizona State recently sent around a list of over 50 periodicals in education that they were considering dropping, and not a single person on the faculty objected; in fact, few faculty could recognize more than a handful of the titles.

However, the weeding out of the obscure has done nothing to threaten the hold that a few journals have on the market. Scholarly organizations are to be blamed or credited—depending on your point of view—with their existence and monopoly. A few years back, the American Psychological Association adopted a policy that read, in effect, that anyone posting a preprint of a manuscript on the Internet would have that manuscript disallowed for submission to any APA journal. That policy survived scarcely a few months before it was withdrawn amid cries of “Foul!” Although the American Educational Research Association shows annual income of only about \$10,000 from the sale of old publications, they have recently rejected an offer by a group of members to scan and make all old journals freely available to all on the Internet. Past AERA journals are now available on a CD, but it costs money. That's too bad.

Let's get down to brass tacks. Scholarly communications are not essentially about paper vs. Internet “packets”; they are about commercialization vs. open access to knowledge. My experiences with publishing e-journals over the last decade have taught me that many more people than we ever dreamed want access to educational research: parents, teachers, professionals of many types, students and scholars far from the United States who cannot afford our books and journals. In the equation that relates the cost of scholarly knowledge to the demand for it, there is a beta-weight known as the “elasticity” coefficient. The elasticity of scholarly knowledge is very, very large, and negative. Charge even a

nominal fee for access to educational research, and the demand will quickly fall to zero. And why wouldn't it? Why wager even a small bet that something about which one has no earthly notion might be of some benefit? Open access to knowledge is about democracy; it's about citizen participation. Perhaps that is why the current administration in Washington, DC, has decided to dismantle the ERIC system.

And beyond the obvious financial vested interests are intellectual vested interests that will long outlive the financial ones. Believe it or not, there are people in the academic world who would prefer that ideas opposed to their pet ideas not see the light of day. Now, this is hardly a big problem in the more refined regions of mathematics and cell biology and nuclear physics; but in the social sciences, it's a big problem. Remember Jensen, remember Herrnstein, remember Burt—just to unravel one tiny thread of the problem. I happen not to agree with what these particular individuals said, but I never would have infringed on their opportunity to say it. The gatekeeper function of a few scholarly organizations has been rationalized as an economic necessity; after all, there were only so many pages of print available. Of course, that rationalization is out the window in the cyber-world, so those vested interests will seek other bases for maintaining control. Expect to hear a lot more talk about Truth and “the responsibility not to mislead the public” when the topic turns to scholarly publication now.

DR: You talk about a balance between publishing anything and having some control over what gets published. Obviously, if everyone simply published anything they wanted on their websites, no one would read all of it and there would be a bias to read only the Websites of the more rich and famous (like Gene Glass). Can you offer any suggestions as to how we counter the so-called “evil journal empires” that seek to restrict the amount of potentially useful information without ending up with piles of garbage?

GVG: The rich and famous, as you put it, are going to be read more often regardless of the medium, and not always undeservedly. “Garbage” doesn't worry me. My beloved adviser in grad school in the early 1960s, Chester Harris (Julian Stanley and Henry Kaiser were also my advisers) would often say that you didn't have to eat

a whole egg to know it's rotten. But seriously, let's suppose that there are 5,000 educational researchers in the world and half of them work on problems that I have some level of interest in; and that half of these researchers write one report a year that they want to share with the world. That's 1,250 reports a year or about five reports a day, if I don't read on holidays and weekends. Now how long will it take me to scan the abstracts and a few pages of 5 reports a day to determine whether it's a “keeper” or another piece destined for oblivion? Not long. In fact this is how scientists in physics, chemistry and many areas of mathematics and chemistry have been getting their information for years now; go to xxx.lanl.gov, it's there for anyone to see. I'd much rather do this than trust journal reviewers whom I don't know to filter out everything they don't think is worth my time (given the fact that the journal reviewers themselves seldom agree and take two years to do their work).

There's an irony hiding in this business of “refereed publication.” If I am vaguely interested in a topic that isn't at the center of my work, then I'm prone to pick up a peer-reviewed journal that rejects 90% of everything sent to it. I simply don't have the background to plow through tons of stuff and judge it for myself; trust the experts. But if you are talking about things that come to the core of my own research (let's say, the re-segregating effects of school choice policies), then please don't filter out anything for me; I want to see whatever anyone is writing on the subject and I want it right away. I'll judge it for myself. When we interviewed a sample of “hard scientists” about their reading habits, they said exactly that.

DR: Should groups of academics get together and publish their own cyber-journals that would be “open access” (as you have done)?

GVG: Absolutely. Groups or individual scholars should create online research journals. No organizational imprimatur is needed. The valuable ones will survive by dint of the quality of their editorial boards and the articles they publish. And should the journals be free? Absolutely. University professors like myself are paid well to pursue a research agenda and serve the public that supports them. Returning to them, the public, the fruits of that research are the least we can do. Nothing more than

our time is needed to publish research in the age of cheap, worldwide telecommunications; no one should have to pay to access research.

DR: There have been some rumblings within AERA that the organization will soon meet with the same fate as NRC (National Reading Conference) and APA (American Psychological Association) in the sense that some members will break away and form a new association (Society for the Scientific Study of Reading and the American Psychological Society, respectively). The talk is that many are dissatisfied with what is being passed off as research at AERA conferences and in their journals. As a former president of AERA, do you have any thoughts on this?

GVG: It's inevitable and to be expected. The noteworthy thing is not that this will happen, but that it has been so slow to happen. The hegemony of AERA (to use a word I never use and of whose meaning I'm quite uncertain) stems primarily from the highly politicized nature of educational research. In fact, educational research more resembles a political movement than a science. Consequently, AERA has more to do with legitimizing certain messages (and assisting its members in finding jobs and achieving tenure) than it does with advancing our understanding of education. Because AERA is so political, people fear separating from it and losing its approval.

Meta-Analysis after Twenty-Five Years

DR: Your "Meta-analysis at 25" is a very interesting paper. How has this tool been used well and also abused over that time period? What are some things potential meta-analysts out there should take care to do and not to do?

GVG: Well, there's nothing that people are doing that I would tell them not to do. But there are some things that can make meta-analysis a bit more useful. I've written about these in the "cyber-paper" you refer to (see <http://glass.ed.asu.edu/gene/papers/meta25.html>).

The idea that what we are about in the social sciences is "doing studies" is fundamentally retarding progress—second only in counterproductivity to the idea that our business is about testing grand theories. This idea that we design a "study," and that a study culminates in the test of a hypothesis and that a hypothesis comes from a theory—this idea has done more to re-

tard progress in educational research than any other single notion. Ask educational researchers what they are doing and they will reply that they are "doing a study," or "designing a study," or "writing up a study." Ask a physicist what's up and you'll never hear the word "study." In fact, if one goes to <http://xxx.lanl.gov>, where physicists archive their work, one will seldom see the word "study." Rather, physicists report data—all of it—that they have collected under conditions that they carefully describe. They contrive interesting conditions that can be precisely described and then they report the resulting observations.

Meta-analysis came from the need to extract useful information from the cryptic records of inferential data analyses (t-tests, ANOVAs and the like) in the abbreviated reports in journals. But now meta-analysis needs to be replaced by archives of raw data that we can use to construct complex data landscapes that depict the relationships among independent, dependent, and mediating variables. We need to re-orient our ideas about what we are doing when we do research. We are not testing grand theories, rather, we are charting dosage-response curves for technological interventions under a variety of circumstances. We are not informing colleagues that our straw-person null hypothesis has been rejected at the .01 level, rather, we are sharing data collected and reported according to some commonly accepted protocols. We aren't publishing "studies," rather, we are contributing to data archives. Five years ago, this vision of how research should be reported and shared seemed hopelessly quixotic. Now it seems easily attainable. The difference is the "I" word: the Internet.

The big success of meta-analysis has been in medicine. Over 5,000 meta-analyses of medical treatments have been published in the past 30 years. My physician quotes the results of meta-analyses to me. An internist with whom I play tennis remarked, when he learned that I knew something about meta-analysis, "Well, you know what they say about that: Analysis is to meta-analysis as physics is to meta-physics." Clever, but indicative that meta-analysis has entered the common parlance of practicing physicians. But why not the conversation of educators? Medicine enjoys the advantages of a simpler domain: Treatments are packaged in standard form, out-

come measures are generally agreed upon; success and failure tend to be counted on the same metric. Education research enjoys none of these. That's why our biggest challenge is to tame the wild variation in our findings, not by decreeing this or that set of standard protocols but by describing and accounting for the variability in our findings. The result of a meta-analysis should never be an average; it should be a graph.

DR: If you could offer educational research one suggestion to improve its standing, what would that be? And/or what specifically could we educational researchers begin doing to improve our field?

GVG: It's not a popular suggestion, I predict, and it surely goes against the grain to some extent, but educational research would do well to regard itself not as a science seeking theory to explain such phenomena as classroom learning, teaching, aptitude and the like, but as a technology designing and evaluating lessons, programs, and systems. Some will regard this as a comedown from the search for grand theory. I regard it as a productive advance to a level of relevance and contribution not yet experienced by educational researchers. I don't have the time or words to defend these rather sweeping remarks in detail, and they need defending. But they do have a defender. Everything we have talked about here, Dan, relates in one way or another to the influence that Paul Meehl has had on my intellectual life; and I am not alone in my admiration for his thinking. Lee Cronbach spoke of him in words that he reserved for few others. Michael Scriven named him one of the few psychologists read by philosophers of science. Meehl showed how social scientists' use of inferential statistical methods was not only different from how other scientists used it, how its use was actually retarding the advance of knowledge. He and I exchanged several letters after my first pieces on meta-analysis. He wrote cautiously at first, and then with relief when I confessed that to my mind meta-analysis had nothing to do with testing theories in the soft sciences. Most of what I believe about statistics and theory and scholarly communications is either stated or to be read between the lines of Meehl's (1978) famous "Two Knights" paper. He was working on a paper on the soft sciences' use of the word "theory" when he died last spring. Though in his 80s, his

death was a great loss for those of us who struggle with these questions.

Conclusions

This interview will hopefully inspire some to take up the sword to revive the state of educational research. Overall, Gene Glass's evaluation is that there is much that needs fixing, including the way we conduct, analyze, and present the results of educational research. Our field is plagued by special interests and politics. AERA as an organization is in trouble. For Gene Glass, myself, and several others, the "R" now stands mostly for reunion, reconnecting, and receptions, rather than research. It will

take an organized effort to turn our ship around in terms of once again making contributions to science. As for me, I am on my way to pick up a copy of the Meehl paper (1978).

NOTES

¹ A wide-ranging interview (something seldom seen in the pages of *ER*) places a higher premium on coherence and consistency for explaining one's work than the publication of one article. I recommend it to others.

REFERENCES

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7, 1–26.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting & Clinical Psychology*, 46, 806–834.

INTERVIEWER

DANIEL H. ROBINSON is an Associate Professor of Educational Psychology at the University of Texas, Austin, TX 78712; dan.robinson@mail.utexas.edu. His research interests include uses of technology to enhance learning. He publishes similar interviews in the *Journal of Educational and Behavioral Statistics*.

Manuscript received October 13, 2003

Revisions received October 13, 2003

Accepted October 13, 2003