Published in: *Behavioral and Brain Sciences, 24* (3), 2001, 408–409.
© 2001 Cambridge University Press.

Are We Losing Control?

Gerd Gigerenzer
Center for Adaptive Behavior and Cognition,
Max Planck Institute for Human Development, Berlin, Germany

Abstract. Most students are trained in using but not in actively choosing a research methodology. I support Hertwig and Ortmann's call for more rationality in the use of methodology. I comment on additional practices that sacrifice experimental control to the experimenters convenience, and on the strange fact that such laissez-faire attitudes *and* rigid intolerance actually co-exist in psychological research programs.

Methodological practices are rarely the subject of reflection. Most of us have not chosen a practice; someone else did this for us. Yet we tend to defend what we received, following habit, group loyalty, and peer pressure. Hertwig and Ortmann (H&O) do a great service in challenging this proclivity. They ask us to reflect on experimental practices. One might object that differences in practice directly mirror differences in the subject matters of psychology and economics—just as it is natural to use microscopes for viewing things tiny and near but telescopes for ones large and distant. One might thus conclude: Leave the psychologists in peace and let the economists do what they do. My first point is that this "naturalistic" argument is invalid: experimental practices, and their enforcement, in fact vary strikingly within psychology and often resemble those in economics.

What Wundt and Skinner have in common. In Wundt's laboratory, known as the first psychological laboratory, experiments were run more in the spirit of today's economics labs rather than psychology labs. An explicit experimental script went without saying—because the experimental subject was Professor Wundt himself or someone else who held a Ph.D. (Danziger, 1990). The idea of routinely studying undergraduates rather than experts would have been seen as science fiction, and not very good science fiction at that. The experimenter was merely a technician who controlled the instruments, whereas the subject often published the paper. Repeated trials with the same subject were the rule; they allowed the observation of intra-individual error and systematic changes in performance. Performance-contingent payment was not necessary in order to enhance attention and achievement; the choice of experts as subjects guaranteed sufficient intrinsic motivation. Finally, deception was impossible since the experts knew the script and understood the purpose of the experiment.

In B. F. Skinner's laboratory three-quarters of a century later, a script in the literal sense of a written instruction was not applicable, but trials were repeated and conditions well-controlled enough that even a pigeon could eventually figure the script out. Performance-contingent payment was the rule and, moreover, a central concept of Skinner's theory of reinforcement schedules. Deception in the sense of misinforming the pigeon about the purpose of the experiment was hardly possible.

Wundt's and Skinner's research programs can scarcely be more different in nature. Nevertheless, they illustrate that there have always been practices akin to the four guidelines of today's experimental economists. Therefore, H&O's call for rethinking experimental practice should not be simply put aside by "disciplinary" arguments, such as: OK, economists do different things that demand different practices; that's not psychology, so let's return to business as usual.

A bear market for control. In some parts of cognitive and social psychology, we seem to live in a bear market for experimental control. The reasons for this devaluation can be traced beyond the four practices described by H&O. For example, consider the following puzzling observation. In studies of reasoning, German-speaking students have often been reported as performing at higher levels and more consistently than American students do. For instance, the proportion of Bayesian answers elicited with natural frequencies was substantially higher with German-speaking students (Gigerenzer & Hoffrage, 1995) compared to American students (e.g., see Gigerenzer & Hoffrage, 1999, Figure 3). The proportion of students showing perspective change in the Wason selection task was higher with German students (Gigerenzer & Hug, 1992) than in most follow-up studies. The same holds for the proportion of students who reasoned according to the conjunction rule in the Linda problem when the question was phrased in frequencies (Hertwig & Gigerenzer, 1999). Note that Americans and Germans received the same reasoning problems. What is the explanation for this difference?

Prominent American colleagues have suggested that the cross-Atlantic gap in performance could be due to the higher intelligence of German students compared to American students. Others have suggested that the reason might be the higher average age of German students.

I propose an explanation that attributes the puzzling performance difference to experimental practice rather to the students' traits. In our lab, we typically run participants one by one, or in small groups. Engaging in face-to-face (or "monitor-to-face") contact with each participant, the experimenter can make practically sure that each participant understands the task or script and that the participant is not distracted and can focus her attention on the task at hand. In contrast, experimenters who reported substantially lower performance generally did not study participants individually. Their students were tested in large classrooms or even in "take-home experiments." The take-home experiment is a recent development in the art of fast data collection. Here, the researcher distributes a booklet with reasoning tasks in the classroom, asks the students to take it home, try to solve the tasks at home, and return the solutions later. Testing students in large classrooms necessarily means losing experimental control, and take-home tests probably mean losing even more. A researcher has no way of knowing under what conditions the student attempted to solve the tasks at home—some students may have been faithful to the instructions, others may have tried to be, but were distracted by noise or interrupted by friends.

My hypothesis is that this loss of experimental control causes, in part, the differences between the performances of German students and American students. This hypothesis can be experimentally tested by systematically studying the effect of one-by-one testing, large classroom studies, and take-home experiments, while keeping culture constant. I would be curious to learn what H&O think of and know about take-home and large classroom studies as a potential factor number 5 in their list of anti-control devices.

Note that the problem of control has already gained a new dimension: data collection on the internet. The internet offers a rapid way to collect large amounts of data. Little, however, seems to be known about the circumstances under which the participants respond on the net, and how these affect the reliability of the resulting data.

Why laissez-faire here and control there? Let me end with another puzzling fact. Cognitive and social psychologists practice laissez-faire, as described by H&O, but at the same time care a great deal about enforcing strict rules for other parts of experimental methodology. For instance, psychologists tend to insist upon the randomized control group experiment as the only legitimate form of experimentation and null hypothesis testing as a "must" for statistical analysis. However, Fisher's randomized group design is only one of several experimental practices used today in the sciences. For instance, another is the demonstration experiment, in which one makes something

happen—without the statistical principles of randomization and repetition. This type of experimentation is known from Gestalt psychology, such as when an experimenter tinkers with the spatial and temporal relations between two points of light to produce the phi-phenomenon; it is as prominent in Newton's *Opticks* as in today's molecular biology. Similarly, Fisher's null hypothesis testing is only one form of statistical analysis, and a poor one.

Thus, even within cognitive and social psychology, laissez-faire attitudes and a strict enforcement of rules go hand in hand. The big question is, why laissez-faire here and strict control there? Part of the story seems to be historical accident, followed by blind loyalty to institutionalized habits. Or is there a hidden logic?