how the Dionysian brain works. Since the Dionysian approach is often undertaken when the outcome of experiments is unpredictable, self-experimentation is rarely popular with granting agencies.

Linking self-experimentation to past and future science: Extended measures, individual subjects, and the power of graphical presentation

Sigrid S. Glenn

Department of Behavior Analysis, University of North Texas, Denton, TX 76203-0919. **sglenn@unt.edu**

Abstract: The case for the value of self-experimentation in advancing science is convincing. Important features of the method include (1) repeated measures of individual behavior, over extended time, to discover cause/effect relations, and (2) vivid graphical presentations. Large-scale research on Pavlovian conditioning and weight control is needed because verification could result in easy and inexpensive mitigation of a serious public health problem.

Several years ago, Irene Grote, a former student of mine and now a research scientist at the University of Kansas, expressed her conviction that increasing the frequency of self-experimentation among behavioral scientists would contribute significantly to the scientific understanding of behavior. I was not convinced and countered that self-experimentation was a fine tool for personal development and self-understanding, but I doubted its relevance to advancing science. Roberts' target article happily proves Dr. Grote right and leads the way in using a time-honored and powerful research method to bring experimental rigor to the task of integrating the principles of evolutionary theory and the principles of learning.

The first feature of Roberts' research that demands my attention is his use of repeated measures, over extended time, of the behavior of an individual organism to experimentally analyze the variables affecting behavior. Roberts' use of long baselines followed by systematic and repeated manipulations of independent variables is characteristic of the work of Pavlov, Skinner, and virtually all behavioral researchers whose work with nonhuman organisms led to our current understanding of behavioral processes operating during the lifetime of individual organisms.

Roberts' use of this experimental method in self-experimentation has added to its usefulness in two ways. First, self-experimentation allows experimenters a kind of extended access to the behavior of adult human subjects that is difficult – if not impossible – to achieve otherwise. Most people simply will not participate for years in someone else's scientific experiment, so the kind of long-term experiment that is usually possible only with nonhumans is now made possible with humans doing self-experimentation. Second, Roberts has extended the use of this experimental method to explore characteristics of human behavior that may have their origin in the history of the species. In doing so, he brings the experimental method to bear on hypotheses heretofore examined only in terms of correlational research, and his amazing graphs powerfully reveal the effects (or noneffects) of his independent variables. As pointed out by Smith et al. (2002b), "graphs constitute a distinct means of inference in their own right" and graphs "are uniquely suited for the coordination of scientific activities (experimental, numerical, and theoretical)" (p. 758).

The second arresting feature of Roberts' article is the potential importance, both theoretical and practical, of his theory that the set point for weight maintenance can be (and often is) changed via Pavlovian conditioning. If the role of Pavlovian conditioning is supported in the case of weight control, the theoretical importance of this finding extends well beyond the realm of weight change. The theory offers an entrée into the vast and almost unexplored borderland between evolutionary biology and principles of learning, and it brings experimental rigor to the task of integrating the principles of evolutionary biology and the principles of learning.

On the practical side, Roberts' data clearly call for a large-scale investigation under tightly controlled conditions to test the Pavlovian conditioning theory of set point. Should the hypothesis be confirmed, an easily implemented, inexpensive treatment will become available for losing unwanted weight and a major public health problem may be abated. The knowledge that conditioning accounts for some of the problems said to derive from "Stone Age" bodies living in modern environments should allow humans to engineer modern environments to make those environments work *for* them rather than *against* them, thus mitigating any mismatch between Stone Age life and modern life.

A third feature of the target article that stands out for this commentator is a sometimes subtle but pervasive message that our understanding of scientific method is limited by the grip of traditional formulations. As a graduate student, I was puzzled by textbooks on experimental psychology that proclaimed the researcher begins by "stating a hypothesis." I wanted to know how the experimenter came by the hypothesis and I suspected that it usually derived from casual observation (unsystematic induction), or worse, from a "theory" that was itself based on casual observation. This seemed almost to guarantee a great deal of wasted effort. In the target article, Roberts shows that self-experimentation is on a short list of methods that have been shown to generate scientific ideas systematically. By a concatenation of data derived from such experimental analysis and correlational data consistent with theories already supported by converging lines of evidence (theories such as the theory of natural selection), Roberts shows us a way to begin filling the "missing methods" cell in his table of scientific activities. And the way he goes about doing it is entirely consistent with the way the great behavioral researchers have always generated the ideas that eventually led to new discoveries: by systematically investigating what appeared to be a cause/effect relation in the behavior of an individual organism – often discovered while collecting data on something else.

From methodology to data analysis: Prospects for the *n* = 1 intrasubject design

Joseph Glicksohn

Department of Criminology and The Leslie and Susan Gonda (Goldschmied) Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat Gan, 52100, Israel. **chanita@bgumail.bgu.ac.il**
http://www.biu.ac.il/soc/cr/cv/glicksohn.htm

Abstract: The target article is important not only for *black-box* studies, but also for those interested in tracing cognitive processing and/or subjective experience (via systematic self-observation). I provide two examples taken from my own research. I then proceed to discuss how best to analyze data from the $n = 1$ study, which has a factorial design.

Self-experimentation, or what some term *systematic self-observation,* has a venerable history in psychology, going back to the Würzburg school (Humphrey 1951) and even earlier to the work of Ebbinghaus. Roberts eschews a focus on subjective experience, discussing, as he does, cause-effect relationships in behaviour. Yet, I would argue that the importance of his target article is not only for *black-box* studies, but also for those interested in tracing cognitive processing and/or subjective experience. I provide two examples from my own research.

Schacter (1976, p. 475) presents the "entoptic explanation" of hypnagogic imagery, which states that entoptics serve as the raw data for hypnagogia. To test this, we manipulated awareness of both entoptics and hypnagogia in the same observers – my two coauthors, Friedland and Salach-Nachum, serving in time-locked single-subject designs (Glicksohn et al. 1990–91). Systematic self-

https://doi.org/10.1017/S0140525X04250063 Published online by Cambridge University Press

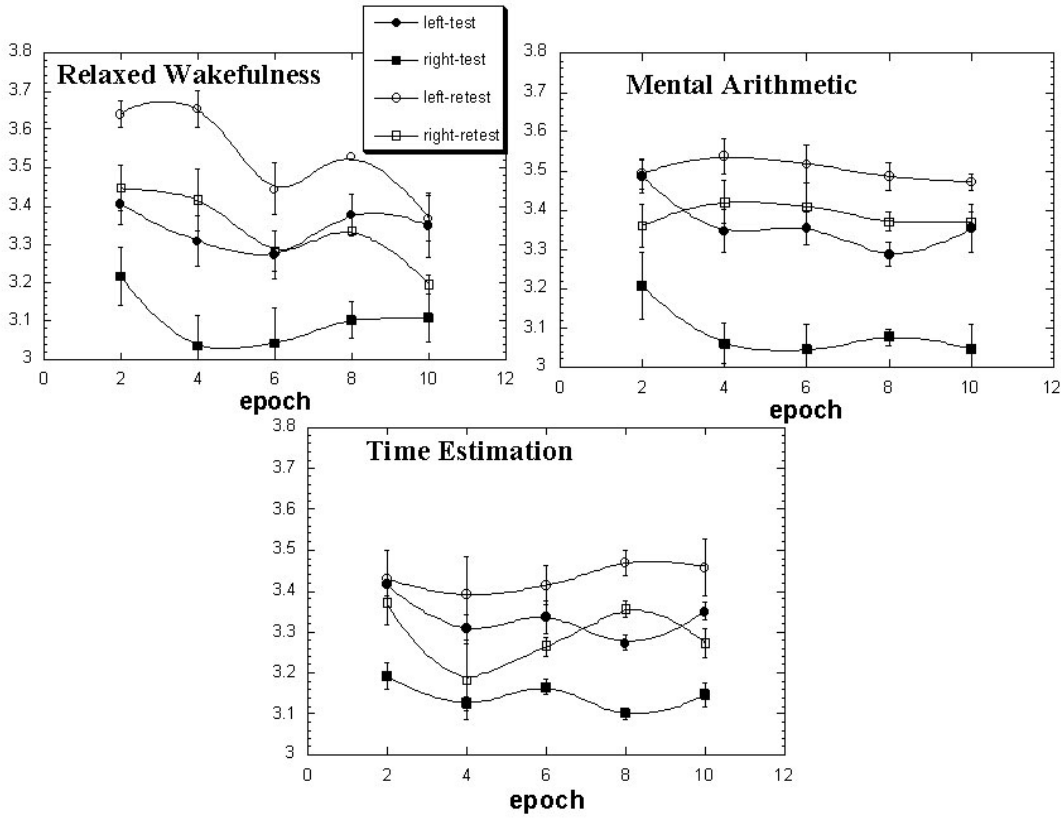


Figure 1 (Glicksohn). Mean alpha power values (after log transformation) together with SEM of one participant, as a function of epoch size, task, hemisphere, and session.

observation was conducted daily for 10 weeks, during which an experimental regime was followed: (1) exposure to a strong source of light for 10 minutes (week 2); (2) vigorous exercise (week 4); (3) cessation of smoking (week 6); exposure to drawings of entoptic phenomena (week 8). Vigorous exercise and cessation of smoking were incorporated because the observers had noted in the past that these factors tended to enhance the appearance of entoptics. The manipulations employed to explore hypnagogia were as follows: (1) use of an alarm clock to awaken approximately one hour after falling asleep (week 3); (2) exposure to monotonous sources of stimulation immediately prior to falling asleep (week 5); (3) exposure to unfamiliar surrealist pictures prior to falling asleep (week 7); exposure to monotonous sources of stimulation, and subsequent alarm-clock awakening (week 9). The entoptic hypothesis could not be supported: there was no correlation between the frequencies of incidence of the two. Furthermore, the content of the hypnagogic imagery was, in the main, not directly related to entoptics. In many respects, this study is similar to some of those described by Roberts. But note the difference in focus – not only on cause-effect relationships, but also on subjective experience.

My second example concerns the bipolarity of mood. One option is of reciprocal inhibition of polar opposites, whereby when one mood is dominant, its polar opposite is suppressed, but can later be released. To investigate the alteration of mood across time, two cooperative participants recorded daily mood (Glicksohn et al. 1995–96). Pleasant and unpleasant affect bore a lawful relationship, indicative of a cyclical, reciprocal inhibition of mood: assessed over the span of three months, they correlated negatively ($r = -0.64$ and -0.71, respectively). Furthermore, using spectral analysis, we found that the two were in opposite phase. Again, in many respects, this study is similar to some of those described by Roberts. But while Roberts reports on simple *t* tests to show reliable changes in sleep duration as a function of phase of diet, his data give clear evidence of cyclicity, better analyzed using spectral analysis.

This brings me to the issue of how best to analyze data from the $n = 1$ study, having a factorial design. There seems to be a confusion in the literature regarding the appropriate analysis of such data (see Crosbie 1995). Yet there is a readily available rationale for implementing an ANOVA with repeated measures, that is, to pool interactions in order to create an error term (Cox 1958, pp. 128–29). Let me demonstrate, using data from an unpublished EEG study.

EEG alpha is customarily quantified using spectral analysis. It is not clear, however, what should be the optimal length of sampling (epoch size). There is too wide a degree of latitude here, with different studies employing different epoch sizes, ranging from 1–4 seconds (e.g., Hoptman & Davidson 1998) to 15–30 seconds (e.g., Wackermann et al. 1993). I conducted a parametric study entailing three tasks (with eyes closed), with online EEG recording: (1) restful wakefulness; (2) mental arithmetic; (3) time estimation. Two participants provided complete test-retest EEG, exhibiting conspicuous, well-regulated alpha within each of two sessions. The EEG was subjected to spectral analysis, employing epochs varying in size from 2 to 10 seconds. Alpha power density was extracted, defined as total power within that band (8–13 Hz); this was subsequently log-transformed to normalize the distribution. I focus on the data of one of these participants.

An ANOVA with repeated measures on epoch size, task, hemisphere, and session, using the 12 profiles (Fig. 1), was run to assess the stability of alpha power as a function of epoch size. In this analysis, a common error term was created by pooling all interactions. The effect for epoch size was significant [F (4, 51) = 4.33, $MSE = 0.004$, $p < .005$]. In short, for this participant alpha power was highest for the 2-sec epoch ($M = 3.39$) and lowest for the 10-sec epoch ($M = 3.29$). In addition, the main effect for both hemisphere and session was also significant [F (1, 51) = 127.56 and

https://doi.org/10.1017/S0140525X04250063 Published online by Cambridge University Press

113.21, respectively; $p < .005$], but that of task was not [F (2, 51) = 1.37]. (In a similar vein, Roberts has employed a one-way ANOVA with repeated measures, to determine whether standing had an effect on sleep latency.)

A second data-analytic approach is to employ a within-subject multiple-regression model (Lorch & Myers 1990), as has been previously implemented for the individual assessment of structural brain asymmetries (Glicksohn & Myslobodsky 1993) and for studying profiles of line-bisection performance (Soroker et al. 1994). Both hemisphere and session were dummy-coded, and task was defined by means of two dummy variables, using the eyes-closed baseline condition as a suitable reference. The regression was significant [F (5, 54) = 47.76, MSE = 0.005, $p < .0001$, R2 = .82], with significant contributions for epoch size ($b = -0.10$, $p < .005$), hemisphere (b = −0.19, $p < .0001$) and session (b = −0.18, $p < .0001$), and no significant contributions of either dummy variable indexing task. These results essentially replicate the previous findings.

Roberts might well consider more extensively analyzing his n = 1 data base, using the methods described above.

Self-experimentation and self-management: Allies in combination therapies

Irene Grote

University of Kansas, LifeSpan Institute, Lawrence, KS 66044.
grote@ku.edu **http://www.people.ukans.edu/~grote/**

Abstract: Self-experimentation is a valuable companion to self-management in the benefit of pharmaco-cognitive-behavior combination therapies. However, data on individuals participating as active therapeutic agents are sparse. Smoking cessation therapy is an example. Roberts' self-experimentation suggests trying more diversity in research to generate new ideas. This may inform current approaches to the cessation of smoking.

Self-management skills – usually referred to as self-regulation or self-control – are valuable tools relevant to nearly any setting where independence from medical, therapeutic, or teaching agents is a desirable long-term goal. Self-management can involve patients as decision-makers in the long-term outcomes of their treatment (see Grote [1997] for combination therapies of pharmaco-cognitive-behavioral dimensions).

Self-management skills emphasize prevention and, importantly, relapse prevention. Treatments which involve individuals as self-managing decision-makers, and therefore, as therapeutic agents, may empirically turn out to work better than those which do not do so. Given these premises, modern training of medical professionals and therapists seems to mandate instructor skills for teaching self-management, and to suggest that instructors practice self-study. Anecdotes, data, and conceptualizations in the history of science indicate that methodological self-study can contribute important outcomes to medical and other sciences (e.g., Altman 1987/1998; Neuringer 1984).

Where there is self-management, self-experimentation typically is not far away. Select academic programs started teaching students long ago the techniques of studying their own conduct or self-perception using scientific method (cf. Burkett 1987; Hoch 1987; Jacobs 1990). (For example: Jacobs exemplifies studying the effects of smoking, with self-experimentation, Burkett exemplifies studying ways to quit the habit, with self-management, and Hoch exemplifies combining both, studying effects and implementing cessation management.)

Self-experimentation motivates individuals' curiosity to *understand* cause-effect relationships resulting from variables that they implement systematically, without necessarily wanting to change their lifestyle. Self-management teaches individuals to change some aspect of their lifestyle by systematically changing relevant variables, usually packages of variables, without necessarily wanting to understand or analyze which variable contributes to the outcome. A combination of both approaches, self-experimentation and self-management, motivates individuals to understand the systematic changes in conditions under which they may be able to manage realistic changes in their lifestyle. Thus, Jacobs (1990) is interested in the effects of smoking on the human heart. He does not tell us whether he is interested in self-managing smoking cessation as a result of the findings of his self-experimentation. Burkett (1987) demonstrates that she manages to stop smoking. She does not convey whether she wanted, or needed to understand which component of her package contributed to her success. Hoch (1987) finds that he can quit smoking – to change his lifestyle – and can gradually reduce his criterion, nicotine content, while he counts monies not spent on cigarettes. Carrying Hoch a step further with self-experimentation would have systematically reversed change in nicotine level and money counting. But in clinical practice, reversals for the sake of scientific curiosity can outweigh short-term cost versus long-term benefit.[1]

For the point of Roberts' self-investigations, in each of the three examples on smoking cessation offered in this commentary, some degree of control was in the hands of the individuals who studied themselves, as well as some degree of ingenuity. Therapies on smoking cessation may benefit from the three types of self-exploration to confer some degree of control to participating individuals, and to come up with new ideas for therapy.

Smoking appears to be one of the hardest habits to kick, particularly for individuals with a history of use of other addictive substances: Effective therapies are hard to find: Recidivism is typically considerably higher than maintenance of abstinence. Choosing the best combination remains controversial (Petry & Simcic 2002; Shoptaw et al. 2002). The present comment proposes that the controversy may persist: Therapies may want to try to include smokers themselves as therapeutic agents and as decision-makers in their combination of pharmacological, physiological, and behavioral components. During self-experimentation, interesting findings may surface; at a self-management, level, participants may experience control over their own quality of life changes. Traditionally, individuals receiving therapy are not engaged as therapeutic agents, perhaps on account of the reluctance to accept any data from self-studies as a valid method. Roberts (1998) mentions this reluctance, and in fact likens it to the reluctance many have in accepting smoking as a hazard, despite evidence to the contrary (e.g., Centers for Disease Control [CDC] 1997).

Roberts' effective use of self-experimentation, his conceptual conclusions combined with traditional experimentation, may teach a valuable lesson for smoking cessation research: The persisting controversy about effective components of combination therapies may incite seasoned grant researchers, or any practitioner with an inquisitive experimental inclination, interested in research-based therapy, to invite individuals or groups of individuals to collect self-observation data and to brainstorm their observations and data in individual or in group settings, suggesting self-management, and spurring self-experimentation. Surprising new ideas may spring from such scientific adventure! Visual analysis of basic data designs for self-experimentation and for self-management can easily be taught, as results from self-experimentation replicated across individuals suggest (Grote 2003); accessible visual analysis may enhance motivation or self-analysis.

Roberts' self-studies demonstrate that diversity in scientific method and conceptualizations may promote "selection" for unusual results, and for a richer fabric of results, as confirming or disconfirming traditional techniques by unusual scientific techniques. The history of science supports Roberts' contention that observing diverse behavior *with diverse methods* is better than observing behavior of a limited group and with limited methods. Metaphorically speaking, based on evolutionary conceptions the strongest (data-based and outcome-effective) conclusions inform each other by way of different approaches, and may spur improved approaches. The criteria, based on diversity of approaches, become more stringent – similar to those operating in technology (Grote 1999).

https://doi.org/10.1017/S0140525X04250063 Published online by Cambridge University Press