APPENDIX: TECHNICAL DETAILS

A.1. Example 1 (Breakfast)

To measure my sleep, I used a custom-made device about the size of a book. At night, when I turned off the lights to fall asleep, I pushed a button (*start*). Five minutes later, the device beeped. If I was still awake, I pushed another button (*awake*) in response. The awake button was large, lit from within, and easy to push. If I pushed it within four minutes of the beep, indicating that I was still awake, the cycle continued: Five minutes after the first beep, the device beeped again. If I failed to push the awake button within four minutes of a beep, the beeping stopped. This measured my sleep latency in 5-minute units. If I awoke in the middle of the night and did not get out of bed (e.g., to urinate), I pushed the awake button. In the morning, when I got out of bed, I pushed a third button (*stop*). The device stored the times that the start and stop buttons were pushed, and counted the number of times I pushed the awake button before and after falling asleep (two separate counts). Similar measures of sleep have a long history (Blood et al. 1997).

A.2. Example 2 (Faces)

As an example of what I watched in the beginning, on July 20, 1995, starting at 9:43 a.m., I watched *Politically Incorrect* (13 minutes), *The Real World* (54 minutes), *Road Rules* (19 minutes), and *The Simpsons* (6 minutes); on August 16, 1995, starting at 9:24 a.m., I watched *Newsradio* (23 minutes), *Frazier* (23 minutes), and *Comedy Product* (standup comedy; 22 minutes).

After I determined that the crucial stimulus was a face looking at you, I mostly watched *Washington Journal* and *Booknotes.* Those provided a high density of such stimuli.

A.3. Example 4 (Morning light)

The light source I placed on my treadmill for the main experiment of this example was an ordinary fluorescent fixture with four standard-length (1.22 m) 40-W fluorescent lamps covered by a diffuser. The lamps were General Electric F40SP65, which emit light with color temperature 6500 degrees Kelvin, resembling skylight. I was exposed to the light while I walked on a treadmill watching television. The light fixture lay on the arms of the treadmill, shining upward. My eyes were about 0.6 m horizontal and 0.6 m vertical from the center of the diffuser. The television was directly in front of the treadmill at roughly eye level.

The treadmill was near a window, so the fluorescent light, to be effective, had to be substantially brighter than light from the window. Measured with a light meter (Gossen Luna-Pro sbc) at the position of my eyes pointed at the center of the diffuser, the intensity of the fluorescent light was 1600 lux. During the experiment, light from the window (measured by putting the light meter at the position of my eyes and pointing toward the center of the window) was about 40 lux at 7:00 a.m., and about 100 lux 2.3 hours later. On the days I had taken walks, the outdoor light intensity at 6:30 a. m., measured by pointing the light meter straight up, was about 1000 lux. Light from the television was 10–20 lux.

Open Peer Commentary

How observations on oneself can be scientific

David A. Booth

School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom. **d.a.booth@bham.ac.uk**
http://www.bham.ac.uk

Abstract: The design and interpretation of self-experimentation need to be integrated with existing scientific knowledge. Otherwise observations on oneself cannot make a creative contribution to the advance of empirical understanding.

Seth Roberts is right to argue that experiments on oneself are unduly neglected in contemporary science. Unfortunately, however, Roberts has misapplied scientific method in the studies that he describes running on himself.

There are four types of flaw in his claims to have evidence for some effects of visual exposure, movement practices, or food selection on expressed mood, perceived sleep, symptoms of a cold, and body weight.

1. Roberts' self-observations are contaminated by effects of his knowledge of previous observations that he made of himself. There is little or no point in self-replication when the phenomena depend on perceptible stimulation or controllable action.

2. Roberts' manipulations are confounded by known influences that may provide explanations of his observations that conflict with his hypotheses.

3. Contrary to the argument by Roberts, the unexpectedness of an observation makes no contribution to the strength of the evidence. This is because, if flaws 1 and 2 were avoided by considering only a single observation, the surprise becomes logically indistinguishable from mere coincidence. Worse, a lifetime spent looking for surprises will collect an increasing number of spurious concurrences. For example, this is the basis of the very high proportion of supposed intolerances for foods that, on testing, prove to be misperceived (Booth et al. 1999; Knibb et al. 1999).

4. In the case of his most "weighty" conclusions, Roberts' theory has been refuted by extensive prior research.

It is with some grief that I see Roberts spoil his case for self-study, because I began my research in molecular neuroscience and my education in cognitive psychology with experiments on myself.

My initial brain/mind interest was the neurochemistry of psychosis. In that context, I once ate nothing but a large bar of chocolate for lunch and analysed its metabolic products, to show that the origin of a compound seen more often in hospital patients diagnosed with schizophrenia was less likely to be in their brains than in the boxes of chocolates which visitors gave to them more often than to the nurses who served as the control group in that project (Booth & Smith 1964). The printer's block for the key figure still sits on my office shelf, labelled "In Memoriam: experimenter as his own subject" – although the nausea that I suffered after eating a half-pound of chocolate was clearly not fatal! Note that my metabolism (or the hospital visitors' gifts) could not be affected by my perceptions or my actions, given that I kept the chocolate down.

Eight years previously, I worked by myself through a little book of experiments in psychology (Valentine 1953). My memory is of laboriously training out the Müller-Lyer illusion and replicating the primacy and recency effects in recall of lists of words. Later, however, I found that I had a capacity for direction of attention sufficient for observable dissociative effects as autosuggested movement. So I became aware that directed forgetting could modulate that curve of deficits in serial recall. No one these days should estimate the size of an effect without comparing performance between people who know and don't know the correct hypothesis.

https://doi.org/10.1017/S0140525X04220064 Published online by Cambridge University Press

Some of my professional discoveries about hunger provide a basis for sympathy with Roberts' thesis but also for criticism of his examples. In some instances, I can't tell if personal experience stimulated the hypothesis, my theorising triggered the self-observations, both, or neither.

For example, Booth et al. (1970) provided the first (group) evidence that protein is better than carbohydrate in a meal at keeping hunger at bay some hours afterwards. The finding has been replicated several times (most recently by Long et al. [2000]); indeed, the original pair of experiments was limited, like all single studies, and so needed to be extended by different designs (very differently by French et al. [1992]). The effect, moreover, may be the key to low-carbohydrate diets (like Atkins'): weight loss occurs only when energy intake is lowered, and reduction of hunger by the raised protein content may enable this self-restriction to be better sustained (Bravata et al. 2003). The autobiographical twist on this is that I had gained the impression that meals based on rice, even when I had eaten enough to feel very full, left me hungry again little more than an hour later. For a long while now, I have believed that I can prevent this (other) "Chinese restaurant syndrome" by including enough flesh food of some sort in the meal. Yet I can't tell if this effect is self-experimental evidence for the hunger-delaying effect of protein or autosuggestion from – or valid application of – my theory of late mobilisation and utilisation of assimilated amino acids through the alanine cycle.

On Roberts' ideas about sleep, he pleads that the basis of his anticipatory awakening is "surely not expectation." Yet it is likely that he was aware that he was not going to have his usual breakfast: I regularly wake early when I know I have something unusual to do when I get up. Quite apart from autosuggestion, Roberts fails to allow for some obvious mechanisms. For example, eating fruit in the evening could induce earlier waking because fruit fills the bladder more than drier foods do.

The most disastrous moves made by Roberts deploy the notion of a "set point" for body weight. This concept of a reference value is simply redundant when there are opposing negative feedback functions (Booth et al. 1976; Peck 1976). Furthermore, even for body temperature regulation, the hypothalamus has only countervailing networks for heat production and dispersal – no 37°C-setting neurones. The urgent scientific issues about obesity are the mechanisms by which a person can most easily lose more energy than they gain during and between meals (Blair et al. 1989; Booth 1998): "set" points that move (!) divert attention from the real scientific problems. Similarly mind-numbing is the unoperationalised notion cited by Roberts that flavour-calorie associations increase "palatability" (Booth 1990; Conner et al. 1988).

Basic mistakes undermine these designs and interpretations. One example will have to suffice here: sucrose is a compound of fructose and glucose and so is useless as a control for fructose. Indeed, a lot of fructose without glucose is poorly absorbed and the resulting upset could reduce hunger (Booth 1989, p. 249). Roberts can swap anecdotes with his readers for a very long time, but scientific understanding is not advanced until a literature-informed hypothesis is tested between or within groups in a fully controlled design shown to be double-blind.

To conclude, personal experience can be a good way to get new ideas. Deliberate manipulation of the environment and keeping an eye open for unusual consequences may accelerate the generation of hypotheses. Yet the only way that science progresses with new ideas is to test novel hypotheses against existing theories in a competent design. Individualised analysis of complex performance (Booth & Freeman 1993; Booth et al. 1983; 2003) is also grossly neglected but requires adequate design and aggregated data. Don't compare conclusions; find out about mechanisms.

Dionysians and Apollonians

Michel Cabanac

Department of Physiology, Faculty of Medicine, Laval University, Quebec, G1K 7P4, Canada. **michel.cabanac@phs.ulaval.ca**

Abstract: There are two sorts of scientists: Dionysians, who rely on intuition, and Apollonians, who are more systematic. Self-experimentation is a Dionysian approach that is likely to open new lines of research. Unfortunately, the Dionysian approach does not allow one to predict the results of experiments. That is one reason why self-experimentation is not popular among granting agencies.

The title of this commentary is borrowed from a letter to *Science* by the Nobel laureate Albert Szent-Györgyi (Szent-Györgyi 1972). In it, Szent-Györgyi recalled that the terms Apollonian and Dionysian, proposed earlier by J. R. Platt (in a personal communication), reflect extremes of two different attitudes of mind that can be found probably in all walks of life, but apply especially well to scientific research. Apollonians tend to develop established lines to perfection, whereas Dionysians rely on intuition and are more likely to open new, unexpected lines of research

Self-experimentation, as demonstrated by Roberts, is evidently Dionysian. Roberts shows how the results of each experiment led to a new hypothesis, then to another self-experiment, and so forth. I can modestly confirm that I have also experienced a similar concatenation of ideas. For my M.D. thesis, I trained in the area of temperature regulation in relation to central nervous system physiology under J. Chatonnet (researching exclusively with dogs, which was still possible back then); then I did a post-doctorate under P. Dejours, a respiratory physiologist. Dejours proposed that I should check whether what he had discovered and termed "accrochage" – that is, the peripheral nervous signal for hyperventilation at the onset of exercise – would also take place with shivering. It did, but that is not my point here. Both of us served as subjects in an experiment with acute cold exposure aimed at arousing shivering. During these sessions, I experienced what J. Barcroft and F. Verzár had discovered and described as "basking in the cold": a sudden relief and relaxation between the bursts of shivering that accompany intense cold discomfort (Barcroft & F. Verzár 1931). This observation largely determined my future life in research, because after returning from Dejours' lab I started to study what caused "basking in the cold." This led to another fruitful experience. One day I had subjected myself to a hot bath in order to study the influence of core temperature on thermal comfort and was in a hurry because another subject was supposed to arrive in the lab any minute. While I was cleaning the bathtub with a hose, I became aware that the very cold water from the hose felt pleasant on my hand. The ice-cold water, usually unpleasant, felt good! It felt good because I was still hyperthermic and sweating. This was the beginning of a still-active line of research: first, on sensory pleasure, then, on other pleasures like taste: then, on phenomena like earning money or enjoying poetry. I have described the process of self-experimentation leading to new lines of research, and the sequence of events that determine lines of research in my book *La Quête du Plaisir* (Cabanac 1995).

Thus, my experience confirms Seth Roberts' experience. I congratulate *BBS* for giving a voice to such an unusual, and unpopular, view. The advantage of self-experimentation has been demonstrated in the target article. However, a profound disadvantage lies hidden in this type of approach, as recalled by Szent-Györgyi himself. The Dionysian approach of self-experimentation indeed favors "accidental" discoveries (by well-prepared minds), but it is impossible to predict what will be found, and what the next question will be. Yet, nowadays research is a profession, managed by professionals, with strict criteria. Large amounts of money are usually needed for research and this often comes from the public; hence, taxpayers must be confident that their contributions are well spent. In turn, granting committees tend to demand detailed descriptions of the research projects, with a cascade of experiments and their expected results predicted in advance. This is not

https://doi.org/10.1017/S0140525X04220064 Published online by Cambridge University Press