Journal of Business Research 65 (2012) 1060–1066

Contents lists available at ScienceDirect

Journal of Business Research

The reception of my self-experimentation

Seth Roberts *

Tsinghua University, Beijing, China
University of California, Berkeley, United States

ARTICLE INFO

Article history:
Received 1 June 2010
Received in revised form 1 November 2010
Accepted 1 February 2011
Available online 12 March 2011

Keywords:
Self-measurement
Introspection
Self-tracking
Sleep
Mood
Innovation

ABSTRACT

Self-experimentation makes some experiments much easier. They might be impossible without it. It can generate plausible new ideas by (a) producing surprising results, which suggest new ideas; and (b) allowing implausible ideas to be cheaply tested. For example, one of my self-experiments showed that seeing faces in the morning raised my mood the next day. Another found that standing more than 8 h while awake made me sleep better. A long article about my self-experimentation (Roberts, 2004) got a chilly reception within my department (psychology). It got a much better reception elsewhere. Blog posts about it led to a popular book (Roberts, 2006) based on one of the results. My self-experimentation combined insider knowledge, outsider freedom, and the motivation of someone who personally benefits from the research — a potent combination.

1. Introduction

Researcher introspection and my self-experimentation (Roberts, 2004) share something important: their reason for existence. Researcher introspection makes it easier to study hard-to-study topics. Gould (1991) used it to study how various products changed his "vital energy" (Gould, 1991, p. 194), meaning how focused he felt, how energetic he felt, and so on. "Much of consumer research [using conventional methods] has failed to describe many aspects of my own consumer behavior," he wrote (Gould, 1991, p. 194). He used introspection to fill the gaps and described in detail experiences over many years. To collect such data in other ways would have been impossible. Earl (2001) used his own experience to shed light on why people pay high prices to attend rock concerts when recordings are cheap. The question would have been much harder to answer in other ways. Hirschman (1990, p. 115) described how a "near-death experience altered her perception of consumption [and] consumer behavior research." To learn from near-death experiences using common research methods is probably impossible. Reid and Brown (1996) used introspection to argue that shoppers classified *apathetic* experience a lot of emotion during shopping. It would have been much harder to gather similar data from others.

To say researcher introspection should not be used because of flaws (Wallendorf and Brucks, 1993) ignores this. "One might as well … junk the Cuisinart because it won't chop firewood" (McKibben, 1983, p. 44). All methods have flaws and limitations. Reasonable criticism of an unusual method is to argue in specific cases – specific uses of the method – that the conclusions were false or misleading. The main critique of researcher introspection (Wallendorf and Brucks, 1993) did not do this.

My self-experimentation (Roberts, 2004) filled a similar niche: It made difficult or impossible experiments much easier. The otherwise-impossible experiments (a) produced surprising results that generated new ideas and (b) tested ideas I couldn't have tested otherwise.

2. Background

My self-experimentation began when I was a graduate student in psychology. I wanted to learn how to do experiments. I believed "the best way to learn is to do" (Halmos, 1975, p. 466) so I tried to do as many experiments as possible. Self-experiments were much faster than the rat experiments I'd been doing.

One of my first self-experiments tested two acne medicines (tetracycline and benzoyl peroxide) my dermatologist had prescribed. In the beginning, I thought tetracycline worked and benzoyl peroxide did not. In a few months, my results showed the opposite: benzoyl peroxide worked and tetracycline did not. I was stunned. It had been so easy to learn something surprising and useful.

A few years later, I started waking up too early. Because of my acne experience, I began to self-experiment (record my sleep and try various treatments) to improve the situation. I tested a wide range of non-drug treatments (e.g., exercise, eating cheese). For ten years, nothing worked. Nothing I tried made a difference. Then something curious happened (Example 1 of Roberts, 2004). My sleep records showed that at the same time I had lost weight I had started to sleep

* 1152 Arch Street, Berkeley, CA 94708, United States.
E-mail address: twoutopias@gmail.com.

0148-2963/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbusres.2011.02.014

less. I showed this to my students. A week later, one of them came to my office and suggested a diet with high water content. It would make me lose weight and need less sleep, he said. I tried it. Nothing happened. The next time I saw the student, I told him that the water diet had made no difference. "How much fruit are you eating?" he asked. "Four pieces a day," I said. He said he ate *six* pieces a day. So I changed my breakfast from oatmeal to two pieces of fruit. As soon as I made the change, my early awakening got *worse*. Apparently breakfast mattered. I tried several different breakfasts. Eventually I concluded that any breakfast with calories caused early awakening. This made more sense than you might think. If a mammal is fed at the same time every day, it will become active a few hours earlier, a phenomenon called *anticipatory activity*. I'd been eating breakfast at 7 am and waking up at 4 am.

These results also suggested a broader point. Our bodies were shaped by evolution to work well under Stone-Age conditions. Stone-Age people didn't eat breakfast. So it made some sense that breakfast caused trouble. My breakfast results made me think perhaps many health problems were due to differences between modern life and Stone-Age life.

After that, I focused on Stone-Age/modern differences. *Can this or that element of Stone-Age life improve health?* I asked. This seemed to be a good approach because my rate of discovery increased. Soon after the breakfast results, I used an idea about Stone-Age life to try to improve my sleep (Example 2 of Roberts, 2004). Even after I stopped eating breakfast, I continued to wake up too early. Apparently breakfast was not my only problem. What else about my life might be causing trouble? I knew that conventional experiments had shown that social contact controls when we are awake: We tend to be awake at the times we have social contact. I believed that Stone-Age people had plenty of social contact in the morning, whereas I lived alone. Putting these two things – the research result and my idea about Stone Age life – together suggested this: Perhaps lack of human contact in the morning made my sleep worse. Data also suggested that TV could have the same effect as social contact. Perhaps if I watched TV early in the morning, it would improve my sleep.

One morning I tried it. Nothing happened — or so it seemed. After watching the TV, I felt no different than usual. However, when I awoke the next morning, I felt remarkably good — cheerful, calm, yet energetic. I couldn't remember ever feeling so good that early in the morning. I studied the situation further and, to my astonishment, figured out that if I saw faces on TV early in the morning it raised my mood the *next* day — not the same day. If I saw faces Monday morning, I felt better on *Tuesday*.

I did a simple experiment that showed the effect and helped explain it. On some days I saw about 60 min of faces on TV starting at 6 am; on other days everything was the same except the faces were covered. In addition, I rated my mood every few hours. I used three mood scales. On Scale 1, I rated myself on the dimension sad/happy; on Scale 2, on the dimension reluctant/eager; and on Scale 3, on the dimension irritable/serene. Each scale went from 0 to 100, with 50 = neutral, 60 = slightly positive, 70 = somewhat positive, and 80 = quite positive. For example, if I felt slightly happy I would rate myself 60 on the sad/happy dimension. If I felt slightly sad, 40 on that dimension. If I felt somewhat eager, I would rate myself 70 on the reluctant/eager dimension. To get an overall score, I averaged the three scales.

Fig. 1 shows the results. Each point is a different measurement. The upper panel, which shows my mood at 4 pm, shows the next-day effect I'd noticed. The 6-am faces greatly improved on my mood, but with a one-day lag. Several things made this unlikely to be a placebo effect. First, the initial observation was a huge surprise. Second, the size of the effect depended on many details. For example, East/West travel across many time zones eliminated the effect for weeks. Third, experiments about the effect often produced results often different from what I had expected. Fourth, other results, where a placebo effect was impossible, pointed in the same direction.

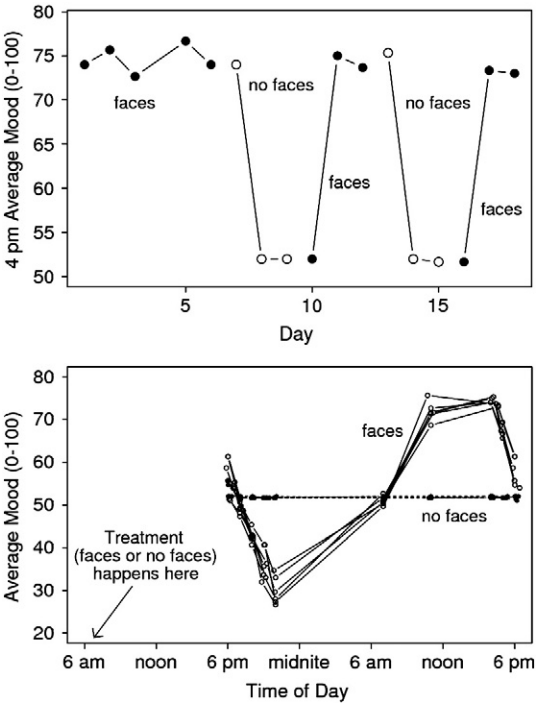

Fig. 1. Mood ratings over 17 days. Upper panel: mood at 4 pm. Lower panel: time course of the effect. In both panels, each point is an average over three ratings, one from an unhappy/happy scale, one from an irritable/serene scale, and one from a reluctant/eager scale. In the lower panel, each line is a separate day. The data start about 12 h after the treatment because that's when the treatment began to make a difference.

The lower panel of Fig. 1 shows how my mood changed throughout the day. It shows that the 6-am faces caused an oscillation in mood (low at night, high during the day) that started about 6 pm of the day I saw the faces – in other words, about 12 h after I saw them – and lasted about 24 h. These results suggest (a) we have an oscillator that controls our mood and (b) that oscillator is given a "push" (caused to oscillate) by seeing faces in the morning. See Roberts (2004) for more about this.

Another Stone-Age-influenced discovery was that standing improved my sleep (Example 3 of Roberts, 2004). One day a colleague said it would be nice if typing counted as exercise. Her remark made me wonder if standing counted as exercise — that is, resembled conventional exercise. If I stood a lot, would I lose weight? My belief that Stone-Age life contained elements crucial for health made this question interesting to me. Surely Stone-Age people stood much more than I did.

I worked, ate, and talked on the phone standing up. My weight didn't change. However, to my surprise, my sleep records showed that I was waking up early less often. (Aerobic exercise hadn't helped.) Fig. 2 shows what happened. I defined an instance of early awakening to be a morning when I fell back asleep between 15 min and 6 h after waking up for the first time. (For example, wake up at 4 am, fall back asleep at 7 am.) The top panel of Fig. 2 shows how the probability of early awakening changed over time. During a baseline period (before lots of standing), I woke up early about 60% of mornings. During the first period I tried to stand more (Phase 1), I woke up early about 30% of mornings.

When I analyzed the data from Phase 1, I saw that standing seemed to have no effect unless I stood 8 h or more (lower panel of Fig. 2). If I stood less than 8 h, the probability of early awakening was close to its baseline value. If I stood more than 9 h, the probability of early awakening was near zero.

After I learned that, I tried to stand at least 9 h every day (Phase 2). The probability of early awakening during that phase was close to zero (upper panel of Fig. 2). The dose–response function during Phase 2