The Dead Man Test: a Preliminary Experimental Analysis

Thomas S. Critchfield¹ · Elva Z. H. Shue²

Published online: 23 February 2018
© Association for Behavior Analysis International 2018

Abstract

Behavior analysts often invoke the “Dead Man Test” as a means of distinguishing behavior from other things, but the assumption underpinning this test, that behavior is absent in vitality-challenged individuals, lacks systematic empirical support. In a field experiment, three individuals who reasonably could be considered as deceased each were observed under three conditions in which behavior might have been observed. None was detected. These results are consistent with predictions derived from the Dead Man Test, although, due to limitations of the experiment, that foundational concept of behavioral measurement cannot yet be considered as validated.

Keywords Behavioral assessment · Dead Man Test · Behavioral thanatology · Vitality-challenged individuals

People who are interested in problems of behavior show an astonishing capacity to direct their efforts toward things that are not actually behavior (e.g., Boring, 1953; Chomsky, 1965; Freud, 1900; Stephens, 1998). As a ready heuristic for addressing this problem, Ogden Lindsley (e.g., 1991) developed the Dead Man Test (DMT), which asserts that nothing a deceased person can do qualifies as behavior. This simple maxim has become a key feature of contemporary applied behavioral measurement. The DMT “has been used successfully by a large number of applied practitioners,” (Lindsley, 1991, p. 457), and today popular applied behavior analysis textbooks (e.g., Cooper, Heron, and Heward, 2007; Malott & Trojan Suarez, 2004) invoke the DMT when introducing the concept of behavioral measurement.

In view of the DMT’s widespread acceptance, it is strange that its central claim—that only living organisms may emit behavior—appears to have been widely embraced without critical evaluation (e.g., Critchfield, 2016). Within behavior analysis, “critical evaluation” means “experimental analysis,” but a major hurdle to any experimental test is that the DMT, as Lindsley (1991) described it, is circular—it equates *behavior*

with *being alive* and therefore precludes the labeling of anything a deceased person might do as *behavior*. Some scholars have attempted a philosophical work-around by providing face-valid examples of actions that are performed frequently by at least some living persons and, presumably, can be generally agreed upon as *behaviors*—for instance, “Bobbie’s crossing his legs” (Malott & Trojan Suarez, 2004, p. 9), or “yelling ‘Whoa!’” (Cooper et al., 2007, p. 26). This suggests that the DMT could be evaluated partly by assessing the frequency with which deceased individuals perform such face-valid actions. Although a few studies have detected operant behavior in persons who lack consciousness (e.g., Fuller, 1949), and even in neurons that have been separated from a living body (Stein & Belluzi, 1988), apparently no previous study has undertaken the required analysis, which is the purpose of the present investigation.

Methods

Observations were conducted on a weekday between 9:00 and 10:30 AM in an exhibit (“Inside Ancient Egypt”) at Chicago’s Field Museum of Natural History. Participants were three individuals whose bodies had been exhumed from Egyptian tombs and were on display within glass cases. No medical verification of *deceased* status (state-certified death certificates) could be obtained for these individuals. However, using archaeological methods, Field Museum scientists have determined that the individuals were interred

✉ Thomas S. Critchfield
tsritic@ilstu.edu

¹ Department of Psychology, Illinois State University,
Normal, IL 61761, USA

² Soule Chapel Methodist Church, Smoot, WV, USA

during periods of Egyptian history that, in all cases, are separated from the present day by more than the apparent maximum human life span of about 120 years (Weon & Je, 2009). This provides strong circumstantial evidence that, if subjected to a proper medical examination, the participants would in fact qualify for a diagnosis of *deceased*.¹

Target behaviors were selected that reasonably could be performed by the participants within the limited space available inside the museum's glass cases. The first participant was a boy of unknown age from Dynasty 26 (approximately 664–525 B.C.), referred to here as Tim (not his real name), who was displayed prone on a flat surface without linen wrappings. His target behavior, *sitting up*, was chosen because this behavior commonly terminates prone resting for living individuals and is topographically similar to “tucking and rolling,” which Cooper et al. (2007, p. 26) advanced as one of their face-valid behavior examples. *Sitting up* was defined as raising the head and torso while leaving hips at least partially in contact with the resting surface.

The second participant was a middle-aged man from Dynasty 25 (approximately 747–656 B.C.), identified on his coffin as Harwa, “Doorkeeper in the Temple of Amun” (see <https://ancientneareast.org/tag/field-museum>). Harwa was displayed vertically, facing forward. The target behavior *head turning* was chosen because linen wrappings, which covered Harwa's body but not head, would not interfere with it and this behavior is topographically similar to an example behavior, “looking around the room,” provided in one online explanation of the Dead Man Test (<https://behaviorintervention101.wordpress.com/2013/02/21/what-is-behavior-the-dead-mans-test>). *Head turning* was defined as the entire chin breaking an imaginary line running through the navel and sternum.

The final participant was an apparently juvenile cat (*felis catus*) of unknown gender from the Late Period (ca. 26th Dynasty to 332 B.C.), referred to here as Penny (not the animal's real name). Penny was displayed horizontally on a flat surface with only the head unwrapped, facing upward. *Head turning*, defined as for Harwa, served as the target behavior.

Using 10-s momentary time sampling, an observer noted the presence or absence of a participant's target behaviors while also video recording using an Apple iPhone 5S®. A

second observer independently recorded from the video. The plan was to compare observer records using the kappa statistic (Cohen, 1960), but it requires separate evaluation of agreement for behavior occurrences and nonoccurrences, of which only the latter was observed. Thus, simple percent agreement (Berk, 1979) was used and was determined to be 100% for all three participants.

For analysis purposes, each data point reflected a block of 6 consecutive observation intervals and was expressed as the number of these observations (0–6) in which the target behavior occurred. An experimental condition remained in place until visual inspection of the most recent 4 data points suggested an absence of trend. In an ABCA design, a no-intervention baseline was followed by an Auditory Prompt condition in which the in-person observer, at the start of each block of 6 intervals, lightly knocked three times on the display glass and firmly stated of participant's name, with the process repeating twice afterward (thus, for example, three total cycles of: tap-tap-tap-“Penny!”). Because of concerns about sound transmission through display case glass, a Visual Stimulation condition also was employed in which, once per interval, the in-person observer blinked a 30-lm pocket flashlight (C. Crane Company model PVF) three times toward the participant's eyes. On several occasions, experimental events (auditory prompt or light flash) failed to take place as scheduled due to the unwanted social influence of museum personnel who, for unknown reasons, appeared not to appreciate the importance of the research project. These instances, reflecting a loss of between 33 and 100% of the scheduled prompts, verified by a second observer's inspection of the video recordings, are noted in Fig. 1. The study concluded with a return to baseline, which was cut short when museum officials, in an apparent display of atonement for their earlier interference with the project, offered to personally escort the observer to the museum exit. It was decided that this opportunity to

Fig. 1 Occurrences of target behaviors during baseline and two intervention conditions. Asterisks indicate lapses of treatment integrity; see text for explanation

¹ In planning the research, uncertainties arose about the applicability of ethical guidelines that commonly are used to protect the interests of living individuals. Ultimately, informed consent was not sought because Section 8.05 of the American Psychological Association Code of Ethical Principles of Psychologists (2010 edition; <http://www.apa.org/ethics/code>) states that, “Psychologists may dispense with informed consent only ... where research would not reasonably be assumed to create distress or harm and involves ... naturalistic observations ... for which disclosure of responses would not place participants at risk of criminal or civil liability or damage their financial standing, employability or reputation, and confidentiality is protected.” This clause was deemed to apply since being observed was part of the participants' daily experience and target behaviors (described below) were not apparently relevant to reputational, financial, or legal risks to those individuals.

support positive collegial interactions superseded the need for more data, and the study was terminated at this point.

Results and Discussion

Figure 1 shows that, for all three participants, the target behaviors failed to occur during the initial baseline period and were not occasioned by the stimuli introduced during the two intervention phases. During the second baseline phase, behavior rates mirrored those of earlier parts of the experiment. These results might be viewed as support for the assumption that behaviors commonly seen in living people do not occur for deceased individuals. The support is strictly provisional, however, because negative results can be an artifact of when and how observations were undertaken. Observing the participants at different times might have revealed instances of the target behaviors. Consider, for example, that living individuals, in the transitory condition known as *sleeping*, can remain motionless for periods far exceeding the duration of the present study, only to emit behavior afterward. If something similar holds for the deceased, more extended observations than those of the present study may be needed to detect behavior.

It is equally difficult to interpret the failure of the present interventions to occasion any behavior. These interventions might have been effective had treatment integrity not been compromised (Fig. 1), which illustrates a common risk of conducting research in the field: extra-experimental events can intrude upon the protocol. This suggests that laboratory investigation might be preferable for future research on this topic. As Wacker (e.g., 1996) has explained, it can at times be useful to address questions that arise in applied settings under better-controlled conditions. It should be noted, however, that needed laboratory studies cannot be undertaken until participant recruitment practices (e.g., Pilgrim, 1998) have been suitably adapted for the type of individual currently under study. Based on the experiences of other disciplines, considerable pitfalls during this process can be expected (see <http://hbswk.hbs.edu/item/a-market-for-human-cadavers-in-all-but-name>).

Even if properly implemented, the present interventions may have been ineffective because they were antecedent based (Cooper et al., 2007). Better outcomes might be obtained with consequence-based interventions, but this approach was not attempted in the present experiment due to uncertainties about how to adapt existing reinforcer assessment protocols (e.g., Pace, Ivancic, Edwards, Iwata, & Page, 1985) for the type of participant under study. With the benefit of hindsight, this was a fortuitous decision, because, in order to be effective, consequences usually must be delivered contingent on behavior (Cooper et al., 2007), and none was observed in the present study.

Finally, as always is the case in small-*N* research, the present findings have uncertain generality to individuals other than those who were specifically studied (Robinson & Foster, 1979). For example, nothing is known of participant learning histories (Tatham & Wanchisen, 1998) that might have influenced behavior frequency or intervention outcomes. It may be fruitful in future studies, therefore, to observe individuals whose histories can be determined. A contemporary example might be former Chinese Communist Party Chairman Mao Zedong, whose experiences are thoroughly documented (e.g., Zhi Sui & Zhisui, 1996), and whose preserved body has been publicly displayed in Beijing since his death in 1976. We note, however, that the persistence of history effects may correlate with the duration of the history (e.g., Tatham & Wanchisen, 1998), and the histories of some deceased individuals are far longer than what is normally encountered in living subjects. This makes it difficult to predict the generality of the behavioral history literature to deceased individuals.

To summarize, the present study found no evidence, in three individuals who gave every indication of being *deceased*, of actions that reasonably can be called *behaviors*. This is consistent with predictions based on the DMT which, despite its popularity, has gone unquestioned for more than 50 years (Critchfield, 2016). The present findings thus provide encouragement that our field's applied measurement traditions, which have been so heavily influenced by the DMT, may in fact be sound, though additional studies are needed to corroborate these findings.

First Author's Postscript

This paper was originally created for consumption by students in a graduate course on single-subject research. One goal was to examine single-subject design considerations for an atypical research question, because one's understanding of design principles is never tested so thoroughly as in an unfamiliar context. In the present case, the fact that the research question is illogical does render the logic of design irrelevant to empirically answering it. Quite secondarily, the present paper might be taken as a good-natured prompt for healthy skepticism about time honored conventions in our field, something about which I have written occasionally (e.g., Critchfield, 2014; Critchfield & Reed, 2017). Although the proposition that deceased persons might emit behavior is advanced here strictly with tongue in cheek, only slightly more bizarre notions have received some empirical support. For example, although learning is often said to require a nervous system, learning phenomena have been demonstrated in organisms that lack one. Here, for illustration, I refer the reader to an interesting account of possible learning in slime mold, which is neither an animal nor a discrete organism: <https://www.theatlantic.com/science/archive/2016/12/the-brainless-slime-that-can-learn->

by-fusing/511295/. A colony of slime mold consisting of millions of individual organisms apparently can show effects that meet criteria for detecting classical conditioning in individuals. Even more strangely, when a colony that has learned fuses with a naïve colony, the hybrid colony also shows learning effects. Cool stuff ... showing that one never knows what may be found until one goes looking for it!

Compliance with Ethical Standards

The research was not overseen by an Institutional Review Board or Animal Care and Use Committee because no living organisms were studied. Informed consent was not sought or required because the potential behaviors of interest were publicly observable.

Conflict of Interest The authors declare that they have no conflict of interest.

References

- Berk, R. A. (1979). Generalizability of behavioral observations: a clarification of interobserver agreement and interobserver reliability. *American Journal of Mental Deficiency*, 83, 460–472.
- Boring, E. G. (1953). A history of introspection. *Psychological Bulletin*, 50, 169–189.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Cooper, J.O., Heron, T.E., & Heward, W.L. (2007). *Applied behavior analysis* (2nd ed.). Upper.
- Critchfield, T. S. (2014). Skeptic's corner—punishment: destructive force or valuable social “adhesive”? *Behavior Analysis in Practice*, 7, 36–44.
- Critchfield, T. S. (2016). Requiem for the Dead Man Test? *The Behavior Analyst*. <https://doi.org/10.1007/s40614-016-0082-5>.
- Critchfield, T. S., & Reed, D. D. (2017). The fuzzy concept of applied behavior analysis research. *The Behavior Analyst*, 40, 123–159.
- Freud, S. (1900). *The interpretation of dreams*. London: Hogarth.
- Fuller, P. R. (1949). Operant conditioning of a vegetative human organism. *The American Journal of Psychology*, 62, 587–590.
- Lindsley, O. R. (1991). From technical jargon to plain English for application. *Journal of Applied Behavior Analysis*, 24, 449–458.
- Malott, R. W., & Trojan Suarez, E. A. (2004). *Principles of behavior* (5th ed.). Upper Saddle River: Pearson Prentice Hall.
- Pace, G. M., Ivancic, M. T., Edwards, G. L., Iwata, B. A., & Page, T. J. (1985). Assessment of stimulus preference and reinforcer value with profoundly retarded individuals. *Journal of Applied Behavior Analysis*, 18, 249–255.
- Pilgrim, C. (1998). The human subject. In K.A. Lattal & M. Perone (Eds.), *Handbook of research methods in human operant behavior* (pp. 15–44). New York: Plenum.
- Robinson, P. W., & Foster, D. F. (1979). *Experimental psychology: a small-N approach*. New York: Joanna Cotler Books.
- Stein, L., & Belluzzi, J. D. (1988). Operant conditioning of individual neurons. In M. L. Commons, R. M. Church, J. R. Stellar, & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Vol. 7. Biological determinants of reinforcement and memory* (pp. 249–264). Hillsdale: Erlbaum.
- Stephens, W. (1998). Witches who steal penises: impotence and illusion in Malleus Maleficarum. *Journal of Medieval and Early Modern Studies*, 28, 495–529.
- Tatham, T. A., & Wanchisen, B. A. (1998). Behavioral history: a definition and some common findings from two areas of research. *The Behavior Analyst*, 21, 241–251.
- Wacker, D. P. (1996). Behavior analysis research in JABA: a need for studies that bridge basic and applied research. *Experimental Analysis of Human Behavior Bulletin*, 14, 11–14.
- Weon, B. M., & Je, J. H. (2009). Theoretical estimation of maximum human lifespan. *Biogerontology*, 10, 65–71.
- Zhi Sui, L., & Zhisui, L. (1996). *The private life of Mao Zhedong*. New York: Random House.