Heart Rate Variability Does Not Tap Putative Efficacy of Thought Field Therapy

John P. Kline
Florida State University

Callahan (2001) has offered a series of case reports in an effort to validate the rationale and methods of Thought Field Therapy (TFT). These case reports employ subjective ratings, that is, the Subjective Units of Distress (SUD) rating scale as well as a gross measure of heart rate variability (HRV). My criticisms center around (a) inappropriately strong inferences given exclusive reliance on case reports, a potentially biased sample, and lack of appropriate controls; (b) misinterpretation of statistical artifact as systematic effect; (c) lack of systematic evaluation of HRV changes; and (d) erroneous interpretation of HRV. Callahan's article provides no evidence for the efficacy of TFT nor does it provide evidence for the credibility of TFT's rationale. © 2001 John Wiley & Sons, Inc. J Clin Psychol 57: 1187–1192, 2001.

Keywords: Thought Field Therapy; heart rate variability

> But the fact that some geniuses were laughed at does not imply that all who are laughed at are geniuses. They laughed at Columbus, they laughed at Fulton, they laughed at the Wright brothers. But they also laughed at Bozo the Clown.
>
> —Carl Sagan

I was asked to write this commentary presumably because of my background in psychophysiological methods and basic clinical science. As I read Callahan's (2001) article, it became apparent that its psychophysiological aspects are, relatively speaking, a minor portion of the many things that are wrong with it. Here I will focus on some of those basic issues of research design and inference followed by a brief commentary on

This paper was published as an open review of a published manuscript in this journal. It was not subjected to peer review. The absence of prior peer review of both research papers and the reviews themselves emanated from concerns expressed by Dr. Roger Callahan that the review process was biased against TFT. This paper was published as an open review of the original research paper of TFT. The reader is encouraged to read the original article, along with this accompanying review, and the final critique of the *Journal*'s decision to publish this set of nonreviewed articles in order to gain a perspective on the issues presented. —Larry E. Beutler, editor.

Correspondence concerning this article should be addressed to: John P. Kline, Department of Psychology, Florida State University, Tallahassee, FL 32306-1051.

the psychophysiological aspects of the article. In sum, Callahan has offered a series of vague and selective case reports in an effort to validate the rationale and methods of Thought Field Therapy (TFT). These case reports employ subjective ratings, that is, the Subjective Units of Distress (SUD) rating scale as well as a gross measure of heart rate variability (HRV). My criticisms center around (a) inappropriately strong inferences given the exclusive reliance on case reports, selective sample, and lack of appropriate controls; (b) misinterpretation of statistical artifact as systematic effect; (c) lack of systematic evaluation of HRV changes; and (d) erroneous interpretation of HRV. Callahan's article provides no evidence for the efficacy of TFT nor does it provide evidence for the credibility of TFT's rationale.

Basic Issues of Research Design

When I teach undergraduate students the basic tenets of research methods, or introductory psychology for that matter, I begin the first class of each new semester with this quote by Bertrand Russell (1945): "It is not what the man of science believes that distinguishes him, but how and why he believes it. His beliefs are tentative, not dogmatic; they are based on evidence not on authority or intuition."

This of course implies that the evidentiary standards, and whether they were met, are readily apparent to those who are appropriately trained to evaluate those standards. One scientist can show another the same evidence, and the other scientist should be able to independently draw the same inferences. I can evaluate a colleague's work and then say "Ahhh . . . I now see why my colleague believes as she or he does." Callahan's article sheds no light on why he believes as he does. It represents a disjointed series of unsubstantiated assertions, ill-defined neologisms, and far-fetched case reports that blur boundaries between farce and expository prose.

Callahan suggests that he can use HRV to evaluate the scientific merit of meridian tapping for psychic distress due to potato and corn toxins. In this, he conflates issues of technology and science, overinterprets case-study data, employs only a gross measure of HRV, and then misinterprets the HRV data. Science employs systematic methods of controlled observations; the technology used to make the observations should not be confused with scientific method. A series of case reports with HRV or magnetic resonance imaging scans is still only a series of case reports and is not necessarily more scientific than a series of case reports presenting SUV.

Callahan's report lacks even the most basic methodological rigor. He presents no controlled observations, no statistical analyses, gives no methodological information, and what little methodological information is given illustrates clearly that his entire house of cards is based solely on selective anecdotes. The article communicates nothing of substance, yet Callahan blithely proffers unsubstantiated presuppositions as sacrosanct edicts. Case reports represent the most lax of evidentiary standards, and the ones Callahan presents do not even meet those. He provides little detail on the specific cases, no evidence of controlled observation (e.g., controlled single-subject observation), and inadequate methodological information both in terms of the procedures he uses to quantify HRV and those that he used to validate his algorithms. It is well known that overfocusing on case studies to the exclusion of controlled experimentation can lead to confirmation bias. Callahan would undoubtedly argue that since TFT has such a "high cure rate," there is no danger of this. However, the onus is on him to show evidence of a high cure rate rather than simply reiterating his unsubstantiated assertion. Consider Hume's maxim: "That no testimony is sufficient to establish a miracle, unless the testimony be of such a kind, that its falsehood would be more miraculous than the fact which it endeavors to establish" (Hume, 1758/1952).

In short, incredible claims require incredible evidence. Callahan does not provide credible evidence, let alone incredible evidence. The internal validity of his so-called study is highly suspect. Consider, for example, his section entitled "TFT raises low HRV and lowers too high HRV." Here he gives deep meaning to a simple and obvious statistical artifact, citing a non-peer-reviewed manuscript (Pignotti & Steinberg, this issue) in support of this claim. Again, he relies solely on anecdotal case reports, including those reported by Pignotti and Steinberg. In choosing to focus on only those cases that were either extremely high or extremely low on HRV, he has apparently interpreted regression to the mean (see Ray, 2000) as a meaningful effect!

The following scenario may serve to illustrate this general issue. At Time 1, Callahan defines two groups of individuals by scores on HRV: One group has "low" HRV and the other has "high" HRV. Assume for a moment that between Time 1 and Time 2, we do absolutely nothing to either group. Statistical regression would lead us to predict Callahan's results, that is, that the low HRV group will have higher HRV at Time 2 whereas the high HRV group will have lower HRV. This prediction is made solely on the basis of a statistically commonplace issue that should be covered in any undergraduate course on experimental design. *Non sunt multiplicanda entia praeter necessitatem*. This is Occam's razor, which holds that we should favor explanations that make the fewest possible assumptions. Regression to the mean (i.e., statistical regression) can explain Callahan's reported raising and lowering of HRV; we need invoke neither meridian magic nor potato toxins.

Consider the second sentence of the abstract: "TFT algorithms are effective but the specificity of diagnosed treatment gives results that are superior to algorithms." TFT algorithms have not been shown to be effective nor efficacious. By this, Callahan appears to be referring to what follows the heading of "Algorithm Versus Causal Diagnosis": "Specificity of treatment resulting in superior results has been demonstrated and is documented, with thousands of clients who have showed little progress until more thorough and correct specific procedures were used."

Larry Beutler's invitation was intended to provide a forum for Callahan to air his putative evidence for TFT's efficacy, which he claims is not published because of journal editor bias. In essence, this article *is* supposed to provide part of this documentation. In inviting this work, Dr. Beutler certainly has demonstrated an editor's openness to what Callahan has to say. In response to this more-than-gracious invitation, Callahan has demonstrated yet again the invalidity of his putative "therapy" (see Lilienfeld & Lohr, 2000).

Psychophysiological Considerations

The use of HRV as an "objective" measure necessitates that it is properly measured, properly quantified, and properly interpreted. Callahan provides no evidence that his measurements meet any of these criteria. Simply stating that the measures are objective and scientific because they are physiological or medical is not enough. This is like saying that we can use brain waves to communicate telepathically simply because we can measure electroencephalograms (EEGs) and they seem sort of like radio signals. The ability to measure EEGs is not tantamount to the ability to read minds. Neither is the ability to measure HRV tantamount to a scientific validation of treatment efficacy.

Callahan suggests that HRV can provide better information about therapeutic progress than SUD. However, he then cites another non-peer-reviewed paper (Callahan, this issue) as evidence that there is a "close correspondence between HRV changes and improvement ratings given by the client after TFT treatment." I obviously cannot comment on this work because I have not yet had an opportunity to read that article (i.e., as of the time I write this, the paper has not been published). However, I can say this: Callahan cannot have it both ways. If SUD correspond closely to HRV changes as he claims, then the incremental validity of HRV changes for indexing therapeutic progress is questionable, and he makes no effort to assess this. Callahan apparently is oblivious to this concern.

When relating psychological processes to physiological measures, five types of relations are possible, generally speaking (Cacioppo & Tassinary, 1990). These include one-to-one, one-to-many, many-to-one, many-to-many, and null relations. The null relation is, of course, our beginning assumption. It is incumbent upon investigators to demonstrate the other types. Cacioppo and Tassinary (1990) go on to define types of psychophysiological relations according to their specificity and generality. A one-to-one relation would be considered to be highly specific. Furthermore, one that is context-free would be considered to be very general as opposed to one that is highly context-bound. This is important because Callahan assumes that HRV stands in both one-to-one and context-free relations to better health. According to Callahan, increased HRV is sine qua non of better mental and physical health. This is a questionable assumption to say the least. HRV has been related to myriad factors (see Berntson et al., 1997), some of which Callahan reviews in his article. Callahan neglects to mention the multiple determinants of HRV. Among these are the respiratory sinus arrhythmia (RSA) apparent from spectral analyses and occurring in as 0.15 to 0.4 Hz band. RSA is likely mediated by fluctuations of vagal–cardiac neural connections, and ipso facto is taken as an index of vagal activity. There are also R–R interval oscillations at low or mid frequencies (e.g., 0.05–0.15 Hz), which are thought to reflect sympathetic outflow. Finally, there are very low frequencies (VLF; between 0.003 and 0.05 Hz), and ultra-low frequencies that include circadian rhythms. VLF frequency R–R interval fluctuations may reflect thermoregulatory cycles or plasma renin activity (see Berntson et al., 1997).

The aforementioned specific components are quantified through spectral analyses. Callahan employs an unspecified measurement period, making frequency resolution uncertain. Furthermore, there is no description of any artifact scoring procedures, making it possible that any changes observed are due simply to gross movement artifact (i.e., maybe people are more likely to keep still while they are being tapped). To complicate matters, he uses the standard deviation of normal to normal (SDNN), which is calculated as the SD of all normal R–R intervals. This measure makes the source of HRV difficult to interpret. Since SDNN is frequency nonspecific, we do not know whether any changes, if indeed legitimate changes had been documented, would have been mediated by sympathetic and/or parasympathetic nervous systems. Since Callahan does not provide convincing evidence of sound methodology, this may be a moot point. However, it is an important methodological consideration for any study purporting to assess HRV.

HRV has been related to emotional states, anxiety, and depression, so assessment of HRV is most certainly interesting in the context of treatment outcome (cf. Balogh, Fitzpatrick, Hendricks, & Paige, 1993). There is evidence, for example, that RSA is decreased in infants who are more irritable and that higher RSA is associated with faster habituation and increased reaction times to novel stimuli in infants. Decreased RSA also has been reported in patients diagnosed with panic disorder (see Hugdahl, 1995). Furthermore, decreased RSA appears to be associated with trait anxiety (see Watkins, Grossman, Krishnan, & Sherwood, 1998).

Although HRV is interesting within the context of treatment outcome assessment, it is erroneous to hinge evaluation of treatment efficacy solely on HRV measurement. As mentioned earlier, this assumes that higher HRV is isomorphically related to better health. HRV is related to several factors, some health-related and some not. Among its many correlates, HRV has been related to frequency of sexual intercourse, but not masturbation, in cohabiting couples (Brody, Veit, & Rau, 2000). RSA is greatest when respiration is slowest, which provides a more parsimonious explanation of anxiety and panic effects than does a blocked meridian or an individual energy toxin. Unfortunately, there is no evidence that Callahan controlled for respiration rate. Of course, as mentioned earlier, he does not present any experiments, and his loose method of case presentation does not really control for anything. It is certainly plausible to suggest that the attentional demands of being tapped could lead to alterations in respiration rate, which in turn could lead to alterations in HRV.

Individual Energy Toxins and Chemical Sensitivities: Beyond SUD and Spuds

The "individual energy toxin" construct used throughout the article is vague. Callahan states, "what we call the energy system is the first system affected by the toxin." He then goes on to say ". . . and so it therefore seems appropriate to use the term "energy." This unwarranted logical leap presumes that he has demonstrated reliable and valid measurement of what he calls an "energy system." There is no evidence that he has done so. His implicit assumption that HRV measures this "system" simply because it is an electrophysiological measure ignores even the very basics of physiology. As mentioned previously, most changes in HRV can be traced to autonomic and endocrine factors. In contrast, putative "meridians" are not theoretically constrained by neural, lymphatic, or blood vessel boundaries (Shang, 2000). If Callahan wishes to infer changes to flow of Qi energies through meridian points, he will need to devise and validate appropriate measures. These need to be developed before he can meet his objective of being able to assess the individual energy toxic effects of corn and potatoes. For now, these notions have no plausible scientific foundation.

After the heading "More Evidence of the Power of Specificity," Callahan gives an account of treatment of a woman with multiple chemical sensitivity (MCS) who became fatigued at exposure to wheat and Irish breakfast tea. MCS refers to a clinically severe variant of chemical odor intolerance. MCS usually (a) involves numerous lifestyle changes to avoid chemicals, (b) overlaps symptomatically with other controversial conditions such as chronic fatigue syndrome (CFS) and fibromyalgia, and (c) leads to increased rates of disability. Unlike CFS, MCS per se does not yet have a generally accepted case definition (see Bell, Schwartz, Hardin, Baldwin, & Kline, 1998a). MCS is a controversial diagnosis, and some take the position that it represents nothing more than another manifestation of somatization.

In those with a high degree of self-reported chemical odor intolerance, low-concentration exposures to some chemicals can lead to relatively large changes in psychological and physical symptoms as well as increases in slow-EEG postchemical exposure (Bell, Kline, Schwartz, & Peterson, 1998a). Chemically intolerant individuals generally have chronic symptoms that are not readily ameliorated (Bell et al., 1998a; Bell et al., 1998b). Again, incredible claims require incredible evidence: Without a clear description of how MCS was diagnosed, it is difficult to evaluate the veracity of Callahan's claim. It is possible, for example, that his patient was a suggestible individual with a somatization disorder. He certainly does not provide enough information to rule out this parsimonious alternative hypothesis. Given the demand characteristics of the treatment and the setting in which it was imbedded, the patient may have relaxed, slowed her breathing, and reported lower SUD.

Conclusion

In sum, the evidence that Callahan has offered in support of his far-fetched theory and practices is not convincing. The report is persuasive in one respect: It persuades me that

Gaudiano and Herbert (2000) were accurate in their assessment of TFT: "Thought Field Therapy is marketed as an extraordinarily fast and effective body-tapping treatment for a number of psychological problems. However, it lacks even basic empirical support and exhibits many of the trappings of a pseudoscience" (p. 29)

The present article gives no reason to question these conclusions. In fact, it bolsters them. TFT and related concepts (e.g., individual energy toxins) are based on a rationale without merit. The use of HRV to evaluate its efficacy is based on erroneous premises. Though case studies represent an important, if not indispensable, part of the development of clinical theories (see Davison & Lazarus, 1994), case studies are not sufficient to validate efficacy claims. Evidence for the merit and efficacy of TFT is still lacking. I doubt that such evidence will ever exist, but will remain open to it if and when I see it.

References

Balogh, S., Fitzpatrick, D., Hendricks, S.E., & Paige, S.R. (1993). Increases in heart rate variability with successful treatment in patients with major depressive disorder. Psychopharmacology Bulletin, 29, 201–206.

Bell, I.R., Kline, J.P., Schwartz, G.E., & Peterson, J.M. (1998a). Quantitative EEG changes during nose versus mouth inhalation of filtered room air in young adults with and without self-reported chemical odor intolerance. International Journal of Psychophysiology, 28, 23–35.

Bell, I.R., Schwartz, G.E., Hardin, E.E., Baldwin, C.M., & Kline, J.P. (1998b). Differential resting QEEG alpha patterns in women with environmental chemical intolerance, depressives, and normals. Biological Psychiatry, 43, 376–388.

Berntson, G.G., Bigger, J.T., Eckberg, D.L., Grossman, P., Kaufmann, P.G., Malik, M., Nagaraja, H.N., Porges, S.W., Saul, J.P., Stone, P.H., & Van der Molen, M.W. (1997). Psychophysiology, 34, 623–648.

Brody, S., Veit, R., & Rau, H. (2000). A preliminary report relating frequency of vaginal intercourse to heart rate variability, Valsalva ratio, blood pressure, and cohabitation status. Biological Psychology, 52, 251–257.

Cacioppo, J.T., & Tassinary, L.G. (1990). Psychophysiology and psychological inference. In J.T. Cacioppo & L.G. Tassinary (Eds.), Principles of psychophysiology: Physical, social, and inferential elements (pp. 3–33). New York: Cambridge University Press.

Callahan, R.J. (2001). Raising and lowering of heart rate variability: Some ethical findings of thought field therapy. Journal of Clinical Psychology, 57(10), 1175–1186.

Davison, G.C., & Lazarus, A.A. (1994). Clinical innovation and evaluation: Integrating practice with inquiry. Clinical Psychology: Science and Practice, 1, 157–168.

Gaudiano, B.A., & Herbert, J.D. (2000). Can we really tap our problems away? A critical analysis of Thought Field Therapy. Skeptical Inquirer, 24, 29–33, 36.

Hugdahl, K. (1995). Psychophysiology: The mind body perspective. Cambridge, MA: Harvard University Press.

Hume, D. (1952). An enquiry concerning human understanding. Chicago: University of Chicago Press. (Original work published 1758)

Lilienfeld, S.O., & Lohr, J.M. (2000). Thought Field Therapy practitioners and educators sanctioned. Skeptical Inquirer 24, 5.

Ray, W.J. (2000). Methods toward a science of behavior and experience (6th ed.). Belmont, CA: Wadsworth Press.

Sagan, C. (2001). [On-line]. Available: http://www.pseudoscience.org/definiti-pseudo.htm.

Shang, C. (2000). The mechanism of acupuncture [On-line]. Available: http://www.acupuncture.com/Acup/Mech.htm

Watkins L.L., Grossman P., Krishnan R., & Sherwood A. (1998). Anxiety and vagal control of heart rate. Psychosomatic Medicine, 60, 498–502.