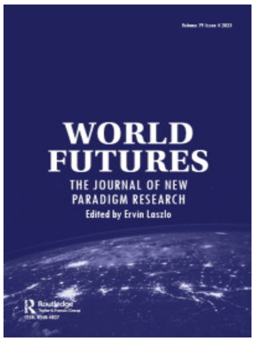

Routledge
Taylor & Francis Group

World Futures

The Journal of New Paradigm Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gwof20

Sentiment and Presentiment in Twitter: Do Trends in Collective Mood "Feel the Future"?

Dean Radin

WORLD FUTURES
https://doi.org/10.1080/02604027.2023.2216629

Sentiment and Presentiment in Twitter: Do Trends in Collective Mood "Feel the Future"?

Dean Radin

Department of Research, Institute of Noetic Sciences, Novato, California, USA

ABSTRACT
Meta-analyses of experiments investigating human behavioral and physiological reactions to unpredictable future events suggest the existence of a poorly understood ability to "feel the future." Is this effect reflected in sentiment metrics based on social media posts? To find out, analysis of 13 years of daily Twitter sentiment data in 10 languages was examined two weeks prior to events assessed as significantly negative and unpredictable, including acts of terrorism, mass shootings, unexpected deaths of celebrities, etc. Results of the analysis was statistically significant ($p = 0.001$), suggesting the existence of a form of collective presentiment.

KEYWORDS
Sentiment; presentiment; social media; prediction

Introduction

This paper describes a study that explores the hypothesis that a daily collective sentiment metric derived from Twitter posts ("tweets") can anticipate future events that are normally unpredictable. The sentiment metric was based on an automated analysis of the emotional quality of words used in tweets in 10 languages, and posted from January 2009 through January 2023 (Dodds et al., 2015).

This unconventional hypothesis was motivated by two types of experiments. The first is known as the Global Consciousness Project (GCP) (Nelson & Bancel, 2011), which began in 1998 and is still active as of this writing. The GCP experiment uses a purpose-built global network of well tested, quantum-indeterminate, electronic random number generators (RNG) located in cities around the world. The GCP's aim was to test if

CONTACT Dean Radin dradin@noetic.org Department of Research, Institute of Noetic Sciences, Novato, CA, USA

events that attract the attention of millions of people, usually via live worldwide media coverage, would be correlated with entropic changes in the truly random data generated by the RNGs. The underlying concept was loosely based on the philosophy of dual-aspect monism (Atmanspacher & Rickles, 2022), which proposes an intimate correlation between mind and matter. From that perspective, during periods of mass mental coherence, similar periods of coherence might be detectable as the emergence of unexpected order in physical systems. In RNGs, such effects could in principle be observed as changes in entropy.

On average, about 40 RNGs have been running continuously in the GCP network, ranging from 3 when the study began to over 70 at its peak. The GCP experiment was in turn motivated by a half-century of prior laboratory experiments using RNGs in mind-matter interactions studies with individuals. Meta-analysis of hundreds of such experiments suggest that the probabilistic outputs of RNGs deviate from chance expectation when subjected to focused human attention and intention (Bosch et al., 2006; Radin et al., 2006; Radin & Nelson, 1989).

The planned portion of the GCP experiment ran from 1998 through 2015, whereupon the formal goal of investigating 500 events had been achieved. In this study, first the researchers noticed, or were alerted by others, especially positive or negative events of worldwide interest. Upon considering the nature of the event, a time-period was selected that was judged to cover the period of interest, generally an hour before the event to 8 hours afterwards. Once that period was formally registered, the random samples generated during that period were retrieved from a central server that continuously collected the data from the RNG network. A standardized analysis was then applied to the data and a statistical assessment made. The result of any one event was summarized with a single z score. Over the 500 events studied the combined results accumulated to an overall z score of 7.3, which is associated with odds against chance of 3 trillion to one (Nelson, 2017). That outcome provided unusually strong evidence of a correlation between inferred periods of collective consciousness and negentropy variations in random physical systems.

The second class of experiments motivating the present study were studies suggesting that human physiology unconsciously responds differentially before exposure to randomly presented calm vs. emotional stimuli. This phenomenon has been dubbed a “presentiment” or “predictive physiological anticipation” effect (Levin & Kennedy, 1975; Mossbridge & Radin, 2018; Radin, 1997, 2004). The effect manifests a half-second to 10 seconds pre-stimulus, depending on the typical time-course of the physiological measure. To date it has been observed in electrodermal activity, blood pressure, heart rate, pupil dilation, EEG, and blood oxygenation levels in the brain (Bierman & Scholte, 2002; May et al., 2005;

McCraty et al., 2004; Radin & Lobach, 2007; Tressoldi et al., 2009, 2011). Meta-analyses of some two dozen presentiment experiments reported from 1997 to 2018 indicates that the effect is (a) independently repeatable (mean effect size = 0.21, $z = 5.3$, $p < 6 \times 10^{-8}$) (Duggan & Tressoldi, 2018; Mossbridge et al., 2012), (b) the replicated effect size appears to be stable and shows no decline over time, and (c) the effect is not due to statistical, analytical, selective reporting, or other known biases (Mossbridge et al., 2014). A conceptually similar type of experiment, colloquially known as "feeling the future" (Bem, 2011), has investigated implicit behavioral responses to randomly selected future events. Meta-analysis of some 90 replications of these experiments again indicates that humans have the capacity to unconsciously anticipate future events that are specifically designed to not be inferable or predictable by ordinary means (Bem et al., 2015).

If these two types of phenomena are genuine, it would imply that human minds may be collectively and continually interacting with the physical world in subtle ways, and that we may be unconsciously and continually "feeling the future." This in turn suggests that if many people are about to experience an unpredictable emotional event, especially an event with strong negative affect (e.g., an act of terrorism), then before that event unfolds we may collectively shift toward a darker mood, and that shift might be detectable in social media sentiment data.

To be clear, this hypothesis does not propose that individuals adjust the language they use when posting tweets on Twitter because they are attempting to predict a future event. Nor does it involve examination of the content of social media posts as a way to anticipate future events (Brito & Adeodato, 2023). Rather, it proposes that emotional reactions to a future event that cannot be anticipated or otherwise inferred might "ripple backwards" in time to affect one's present mood, which would be reflected by the emotional valence of words used in tweets.

Method

Daily Twitter sentiment data were retrieved from www.hedonometer.org, a website hosted by the University of Vermont's Complex Systems Center as part of a project in its Computational Story Lab (Alshaabi et al., 2021; Dodds et al., 2011, 2015). That project calculates sentiment on a daily basis by automatically retrieving Twitter posts, calculating a "happiness score" for each word, and then taking the average. To evaluate English text, a corpus of 10,187 words, selected based on word usage frequency, were each rated by users from Amazon.com's Mechanical Turk service (mturk.com). E.g., with happiness ratings ranging from 1 to 9, with 1 most negative and 9 most positive, the word "love" was assigned a mean score of 8.42 (based on 50 user evaluations), and the word "terrorist" was assigned an average score of 1.3. Figure 1 shows a portion of the sentiment metric for a few months of English tweets, as it is displayed on the hedonometer.org site.

The hedonometer.org site also provides similarly evaluated sentiment scores for tweets in nine other languages: Arabic, German, Spanish, French, Indonesian, Korean, Portuguese, Russian, and Ukrainian. Sentiment data can be obtained for original tweets, retweets only (i.e., shares of an original post), and original posts plus retweets combined. The present analysis focused only on original tweets because retweets generally signify users' agreements with someone else's post, rather than adding new words to the sentiment metric.

As illustrated in Figure 2 (top) for Spanish language tweets, average sentiment varies slowly over time, so an algorithmic method was required to automatically extract the most positive and most negative events. To do this, a moving average window 30 days in length was employed as a high-pass filter, where the window ranged from 30 days prior to the current day, to the current day. This window ensured that the average sentiment metric on a given day was based only on data from the previous

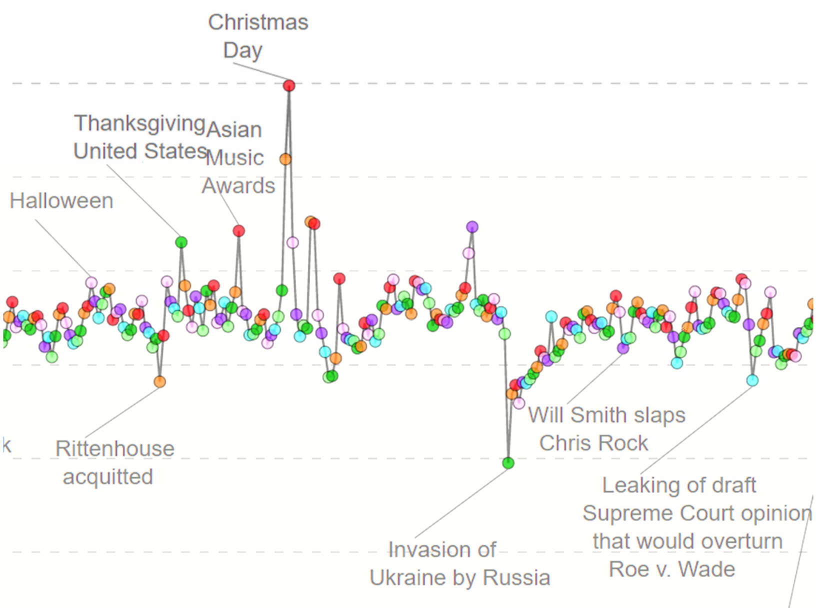

Figure 1. Daily English language sentiment data from October 2021 to February 2022, with annotations for especially notable events. Colors are associated with the day of the week, with Sunday in light purple and Wednesday in green. Note that negative sentiments are typically associated with unanticipated events, while positive sentiments are associated with planned events, such as holidays.

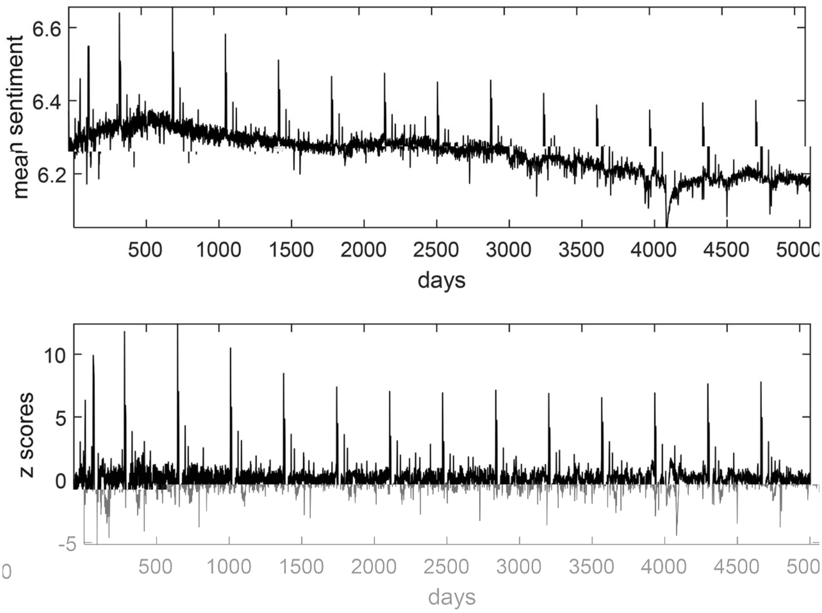

Figure 2. (Top) Sentiment data for Spanish language tweets. (Bottom) Same data after subtracting a 30 day moving average and normalizing.

month. Then that curve was subtracted from the raw data and the resulting residuals were *z*-score were normalized. This procedure generated the curve shown in Figure 2 (bottom).

Note that by automatically extracting the most negative days, it was possible that in some cases it was known that, say, a celebrity was dying, in which case the words in some tweets may have reflected that possibility with especially sad or negative words. Analysis of annotated English tweets (at the hedonometer.org site) indicates that most days with especially negative sentiments were associated with events perceived as shocking, like acts of terrorism, mass shootings, and so on. However, to ensure that the events selected for this analysis were unpredictable, more stringent criteria were required.

To extract days with unpredictable negative sentiment, a four step process was devised using English tweets, and then the same method was applied to the other nine languages. The steps were as follows:

1. A threshold of $z < -2.5$ was applied to the normalized residual data (as in Figure 2, bottom) to locate especially negative days.
2. Because sentiment may take a few days to return to baseline after a strong negative event, any selected negative days were required to fall beyond 7 days of a previously identified negative event.

3. Each of the remaining days was examined by retrieving the webpage returned by Google.com using the search phrase, "what happened on *date language*", where the variable *date* was the date of each selected negative day and *language* was one of the 10 available languages. The information on that page was inspected manually to see if the events reported on that day were deemed predictable or not. In general, such events included acts of political or ideological terrorism, lone actor tragedies such as mass shootings or using bombs or vehicles as weapons, unexpected deaths of celebrities, and major environmental disasters.
4. After selecting dates with unpredictable events, those dates were compared across the 10 languages and if there were any duplicates, only one of those dates was retained.

Through this four-step identification and filtering process, non-overlapping unpredictable negative events were determined for each language. The number of selected days after each of these four steps is shown in Table 1. With the final days identified, an ensemble average was formed ±14 days around the normalized residual sentiment data for each selected day, and then a linear regression slope was formed from 14 days before the selected day to two days before (see Figure 3). This slope was associated with a z score as S_v/S_e, where S_v indicates the value of the slope and S_e its standard error.

To assess the statistical likelihood of the observed slope the aforesaid z score would not be appropriate because S_e was calculated based on successive sentiment values that were not independent, and thus the z value for the observed slope would likely be inflated. Thus, a nonparametric

Table 1. The Combined Result Over The 83 Events Surviving the Fourth Pass was Associated with Stouffer $z = -3.087$, $p = 0.001$.

Language	Days	First pass	Second pass	Third pass	Fourth pass	Slope z	Permuted slope z
English	5124	30	15	10	10	−0.188	−0.076
Indonesian	4954	55	26	13	13	−0.249	−0.089
German	5119	43	27	11	9	−0.579	−0.278
Korean	4916	38	24	7	6	−0.232	−0.049
Spanish	5121	42	16	7	6	−4.327	−2.349
Russian	4763	32	11	2	2	−0.582	−0.192
Arabic	4816	51	36	16	13	−2.677	−1.444
Portuguese	5124	54	31	9	7	−2.619	−1.467
French	4763	47	22	12	10	−2.985	−1.466
Ukraine	4173	58	26	7	7	−4.217	−2.353

Column headings: Tweet language, number of days available, number of events identified with negative sentiment at $z < -2.5$ (first pass), number of events more than 7 days apart (second pass), number of events confirmed as unpredictable (third pass), number of events on non-overlapping days (fourth pass), z score associated with slope of mean sentiment from days −14 to −2 prior to negative event, and circular shift test z score to conservatively assess statistical likelihood of the observed slope.

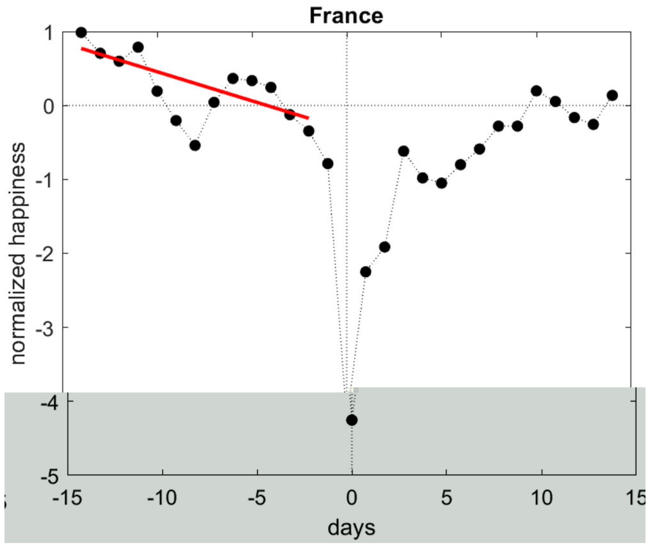

Figure 3. Ensemble average of normalized residual sentiment scores for 10 unpredictable and unusually negative events passing all four selection criteria, for French language tweets. The average score across these 10 days is 4 sigma below the grand average and shown as day 0 on the x-axis. As noted in the text, each daily average sentiment score was determined based on the mean of the previous 30 days, so there is no mundane reason why the slope calculated from −14 to −2 days before the negative events would be negative. By contrast, the curve slowly return to baseline after the negative event would be influenced by the strongly negative sentiment at day 0, because the post-event averages would have included the negative events at day 0.

randomized circular shift technique was employed. This involved circular shifting the original sentiment data by a random amount, then creating an ensemble average as described above, and then determining the associated slope.

The circular shift process was repeated 1000 times to form a distribution of possible slopes (z_d), and then the original z score for the slope (z_p) was compared to that distribution as $z = (z_p - \text{mean}(z_d))/\text{std}(z_d)$. This technique was devised based first on analysis of English tweets, and then the same procedure was replicated for each of the remaining 9 languages, and the resulting z score for each language was then combined via Stouffer Z (Stouffer et al., 1949).

Results

On average only about 0.2% of the days passed all four selection criteria (i.e., ∼10 events/5000 days), and after the permutation analysis only two languages resulted in independently significant slopes (Spanish and Ukrainian). Still, as shown in Table 1, it is noteworthy that all 10 languages resulted in a negative slope, and when the permuted z scores were combined across all 10 languages, the result was a Stouffer $z = -3.087$, $p = 0.001$.

The average normalized residual sentiment score over all 10 languages, with 95% confidence intervals, is shown in Figure 4, which also includes the linear slope from −14 to −2 days before the selected negative events. The slope is calculated −2 days before the event, rather than −1 day before, because events of high negative valence are typically carried live over the news media, and because of the International Date Line it is possible that some negative sentiment reactions would appear to begin the day before an event occurred.

Figure 4 also shows why it is important to limit days selected as negative beyond a 7 day window. By inspection, based on the 95% confidence intervals the mean sentiment value 7 days before the negative events is approximately the same as 7 days afterwards. Of course, in making this comparison, keep in mind that the post-event sentiment means included the negative events by virtue of the way the moving average was designed. Thus, it may take fewer than 7 days for Twitter sentiment to recover to the baseline. But again we used this week-long recovery criterion for the sake of caution.

Discussion

After developing a method to retrospectively analyze English language tweets, and then applying the same method to nine other languages and

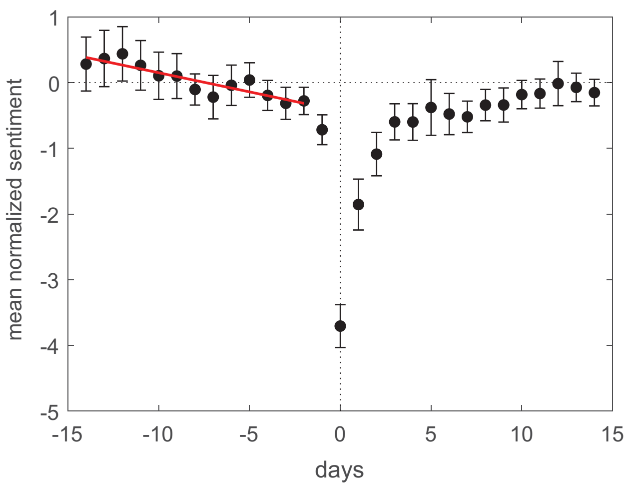

Figure 4. Mean normalized residual sentiment (with 95% confidence intervals) for 84 selected unpredictable events combined across the 10 languages.

combining the results, the results suggest that trends in collective sentiment significantly anticipate unpredictable negative events ($p = 0.001$). Given that these events include acts of terrorism, mass shootings, etc., it may be possible to perform a *prospective* analysis that would paradoxically predict unpredictable events of interest to counterterrorism agencies, law enforcement, mental health advocates, and other important pragmatic applications.

Limitations

This study had two main limitations. First, several parameters were selected arbitrarily, including (a) the decision to examine trends two weeks prior to the event of interest, (b) use of a threshold of $z < -2.5$ to identify markedly negative days, and (c) specification of least 7 days between selected negative days. Second, to determine which events were likely to be predictable required a manual assessment of the results of a Google search for each of those days.

To offset the possible biasing effects of these limitations, the analytical method was developed first only using English tweets, and then the same methods were applied to the remaining 9 languages. Table 1 shows that if one only combined the results of those 9 languages, the results would actually end up statistically stronger, with Stouffer $z = -3.229$, $p = 0.0006$. Thus, while the aforementioned limitations are appropriate considerations in evaluating the present results, the first limitation may not play a significant role. However, to offset possible biases introduced by the second limitation, a future improvement might be the development of an automated, objective means of judging if a given negative event were genuinely unpredictable. That might be performed by identifying words in each language associated with unpredictable events (e.g., terrorist, shoot, kill, bomb, etc.), and then simply counting how often those words appeared in web pages retrieved through search engines focused on the selected days of interest.

Acknowledgements

The University of Vermont's Computational Story Lab project is also gratefully acknowledged for use of their Twitter sentiment database, which is publicly available through their hedonometer.org website.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by the generosity of members and donors of the Institute of Noetic Sciences.

ORCID

Dean Radin http://orcid.org/0000-0003-0041-322X

References

Alshaabi, T., Adams, J. L., Arnold, M. V., Minot, J. R., Dewhurst, D. R., Reagan, A. J., Danforth, C. M., & Dodds, P. S. (2021). Storywrangler: A massive exploratorium for sociolinguistic, cultural, socioeconomic, and political timelines using Twitter. *Science Advances*, *7*(29), eabe6534. https://doi.org/10.1126/sciadv.abe6534

Atmanspacher, H., & Rickles, D. (2022). *Dual-aspect Monism and the deep structure of meaning*. Routledge.

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*(3), 407–425. https://doi.org/10.1037/a0021524

Bem, D., Tressoldi, P., Rabeyron, T., & Duggan, M. (2015). Feeling the future: A meta-analysis of 90 experiments on the anomalous anticipation of random future events. *F1000Research*, *4*, 1188. https://doi.org/10.12688/f1000research.7177.2

Bierman, D. J., & Scholte, H. S. (2002). A fMRI brain imaging study of presentiment. *Journal of International Society of Life Information Science*, *20*(2), 380–388.

Bosch, H., Steinkamp, F., & Boller, E. (2006). Examining psychokinesis: The interaction of human intention with random number generators – A meta-analysis. *Psychological Bulletin*, *132*(4), 497–523. https://doi.org/10.1037/0033-2909.132.4.497

Brito, K., & Adeodato, P. J. L. (2023). Machine learning for predicting elections in Latin America based on social media engagement and polls. *Government Information Quarterly*, *40*(1), 101782. https://doi.org/10.1016/j.giq.2022.101782

Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., Mitchell, L., Harris, K. D., Kloumann, I. M., Bagrow, J. P., Megerdoomian, K., McMahon, M. T., Tivnan, B. F., & Danforth, C. M. (2015). Human language reveals a universal positivity bias. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(8), 2389–2394. https://doi.org/10.1073/pnas.1411678112

Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS One*, *6*(12), e26752. https://doi.org/10.1371/journal.pone.0026752

Duggan, M., & Tressoldi, P. (2018). Predictive physiological anticipatory activity preceding seemingly unpredictable stimuli: An update of Mossbridge et al's meta-analysis. *F1000Research*, *7*, 407. https://doi.org/10.12688/f1000research.14330.2

Levin, J., & Kennedy, J. (1975). The relationship of slow cortical potentials to psi information in man. *Journal of Parapsychology*, *39*, 25–26.

May, E. C., Paulinyi, T., & Vassy, Z. (2005). Anomalous anticipatory skin conductance response to acoustic stimuli: Experimental results and speculation about a mechanism. *Journal of Alternative and Complementary Medicine*, *11*(4), 695–702. https://doi.org/10.1089/acm.2005.11.695

McCraty, R., Atkinson, M., & Bradley, R. T. (2004). Electrophysiological evidence of intuition: Part 1. The surprising role of the heart. *Journal of Alternative and Complementary Medicine*, *10*(1), 133–143. https://doi.org/10.1089/107555304322849057

Mossbridge, J. A., & Radin, D. (2018). Precognition as a form of prospection: A review of the evidence. *Psychology of Consciousness*, *5*(1), 78–93. https://doi.org/10.1037/cns0000121

Mossbridge, J., Tressoldi, P., & Utts, J. (2012). Predictive physiological anticipation preceding seemingly unpredictable stimuli: A meta-analysis. *Frontiers in Psychology*, *3*, 390. https://doi.org/10.3389/fpsyg.2012.00390

Mossbridge, J., Tressoldi, P., Utts, J., Ives, J., Radin, D., & Jonas, W. (2014). Predicting the unpredictable: Critical analysis and practical implications of predictive anticipatory activity. *Frontiers in Human Neuroscience*, *8*, 146. https://doi.org/10.3389/fnhum.2014.00146

Nelson, R. (2017). Weighting the parameters, a response to Bancel's "Searching for Global Consciousness: A seventeen year exploration". *Explore*, *13*(2), 102–105. https://doi.org/10.1016/j.explore.2016.12.002

Nelson, R., & Bancel, P. (2011). Effects of mass consciousness: Changes in random data during global events. *Explore*, *7*(6), 373–383. https://doi.org/10.1016/j.explore.2011.08.003

Radin, D. I. (1997). Unconscious perception of future emotions: An experiment in presentiment. *Journal of Scientific Exploration*, *11*, 163–180.

Radin, D. I. (2004). Electrodermal presentiments of future emotions. *Journal of Scientific Exploration*, *18*, 253–274.

Radin, D., & Nelson, R. D. (1989). Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics*, *19*(12), 1499–1514. https://doi.org/10.1007/BF00732509

Radin, D., Nelson, R. D., Dobyns, Y., & Houtkooper, J. (2006). Reexamining psychokinesis: Commentary on the Bösch, Steinkamp and Boller meta-analysis. *Psychological Bulletin*, *132*(4), 529–532. https://doi.org/10.1037/0033-2909.132.4.529

Radin, D., & Lobach, E. (2007). Toward understanding the placebo effect: Investigating a possible retrocausal factor. *Journal of Alternative and Complementary Medicine*, *13*(7), 733–739. https://doi.org/10.1089/acm.2006.6243

Stouffer, S. A., Suchman, E. A., DeVinney, L. C., Star, S. A., & Williams, R. M. (1949). *The American soldier: Adjustment during army life (Vol. 1)*. Princeton University Press.

Tressoldi, P. E., Martinelli, M., Semenzato, L., & Cappato, S. (2011). Let your eyes predict prediction accuracy of pupillary responses to random alerting and neutral sounds. *SAGE Open*, *1*(2), 215824401142045. https://doi.org/10.1177/2158244011420451

Tressoldi, P. E., Martinelli, M., Zaccaria, E., & Massaccesi, S. (2009). Implicit intuition: How heart rate can contribute to prediction of future events. *Journal of the Society for Psychical Research*, *73*(894), 1.