Effects of Mass Consciousness:
Changes in Random Data during Global Events

Article for Explore: The Journal of Science & Healing

Roger Nelson, Princeton, NJ
rdnelson@princeton.edu

Abstract
A long term, continuing experiment is designed to assess the possibility that correlations may occur in synchronized random data streams generated during major events. The project is motivated by numerous experiments which suggest that the behavior of random systems can be altered by directed mental intention, and related experiments showing subtle changes associated with group coherence. Since 1998, the Global Consciousness Project (GCP) has maintained a global network of random number generators (RNGs), recording parallel sequences of random data at 65 sites around the world. A rigorous experiment tests the hypothesis that data from the RNG network will deviate from expectation during times of “global events,” defined as transitory episodes of widespread mental and emotional reaction to major world events. An ongoing replication experiment measures correlations across the network during the designated events, and the result from over 345 formal hypothesis tests departs substantially from expectation. A composite statistic for the replication series rejects the null hypothesis by more than 6 standard deviations. Secondary analyses reveal evidence of a second, independent correlation, as well as temporal and spatial structure in the data associated with the events. Controls exclude conventional physical explanations or experimental error as the source of the measured deviations. The experimental design constrains interpretation of the results: they suggest that some aspect of human consciousness is involved as a source of the effects.
Introduction

In recent decades mind and consciousness have again become a focus of scholarly research after half a century of psychology with a more mechanical approach. Remarkably, it remains difficult to define for scientific usage what these terms mean. What is consciousness? Where is the mind? How shall we explain creative thinking? Is brain activity the answer to such inquiries? What about the possibility of direct effects of mind in the physical world?

These are difficult yet deeply interesting questions. The last, especially, demands not only scientific clarity, but an inclination for adventure in relatively uncharted intellectual waters. Since early in the 20th century, a few researchers working at the edges of physics and psychology have addressed questions like these in research on “extraordinary” capacities of human consciousness, including mind-matter interaction. The Global Consciousness Project (GCP) was created to broaden these efforts. With contributions from scientists and engineers around the world, the Project has generated a unique body of random data collected in multiple parallel sequences, recorded continuously over more than a decade at widely distributed locations. The data can in principle be used to study any potential modulator of randomness, but the original purpose was to assess the possibility of a subtle reach of consciousness in the physical world on a global scale.

A world-spanning network of physical random number generators (RNG) produces calibrated data meeting rigorous standards of randomness. The question we ask is whether these data may show transient deviations from randomness during instances of strongly focused, collective human attention and emotion. The devices produce a 200-bit trial every second at each of 65 locations around the globe, creating a record of random data that can be compared with the history of major events on the world stage. The hypothesis we test proposes that the data will display non-random behavior during times of “global events.” Specifically, we predict systematic deviations in the network data when there is a widespread sharing of mental and emotional responses. An on-going experimental test of the hypothesis, using a replication protocol, finds significant evidence of characteristic anomalies in the data corresponding to a wide range of events. The results indicate that something remarkable may be happening when people are drawn into a community of common attention or emotion. In this review we present the background, methods, and findings of the decade-long experiment, and address certain implications of the results.

Over much of the modern scientific era, mind has typically been studied indirectly as an implicit function of brain physiology. Consciousness science has focused on how consciousness arises more than how it might impinge on or affect its environment. Nevertheless, for nearly a century, a small number of laboratory researchers have persisted in exploring questions at the margins of our understanding, developing over the years the experimental methods needed to study potential interactions between mind and matter. This area of research offers a unique window into the nature of consciousness by proposing direct manifestations of consciousness in the physical world. Evidence of such effects (disregarded by dedicated skeptics as “impossible”) has been gathered under controlled conditions and the evidence raises puzzling questions. How could it be possible to obtain information from distant locations with no physical or sensory connection?
What could explain correlations between physical processes and the purely mental attention of human subjects? Can there be direct effects of intention in the physical world? Is there a sense in which mind is present in the world beyond the brain?

Laboratory experiments which address these questions often focus on the behavior of random systems. Although physical theory takes causality as a guiding principle, it also admits truly random phenomena (that is, phenomena which are in principle indeterminate, and not merely statistically uncertain). Quantum transitions are a familiar example of this weak causality, which is accepted in physics and is potentially of relevance to mind-matter research. Random phenomena are interesting for research on interactive consciousness because, in our current understanding, they are not completely explained by deterministic causes – a characteristic they share with mind-matter interactions which challenge the completeness of conventional scientific models.

Among the early experiments which investigated the interplay of randomness and conscious activity were studies in which subjects were asked to influence macroscopic systems such as the position or face value of mechanically thrown dice. Since the 1960’s, experiments have more typically used the high speed generation of random numbers employing quantum electronic or radioactive sources. With the advent of the computer, automatic recording helped to ensure experimental control while also facilitating the accumulation of large databases. Improved experiments asked whether the random output of quantum sources could be biased by the mental intentions of subjects. In the latter part of the 20th century, replications of random number generator (RNG) experiments were carried out in laboratories around the world.

One prominent research program, the Princeton Engineering Anomalies Research (PEAR) laboratory, was founded by Robert Jahn in 1979 at Princeton University. In carefully controlled RNG experiments, the PEAR lab demonstrated a small, persistent effect equivalent to a few parts in 10,000. Compounded over the full database, the effect is highly significant and cannot be adequately explained by chance fluctuation or methodological error. The research extended the seminal early work of Schmidt and motivated replication experiments in several independent laboratories. While many experimental questions about the RNG experiments remain (most notably the role of psychological variables), the research carefully documents anomalous departures from expectation associated with human consciousness, and specifically with directed intention.

Later versions of the RNG experiments used portable random sources and by the early 1990s field work was feasible. In the field experiments, rather than instructing a participant to focus his or her intention on a laboratory RNG, the device was brought to locations where groups of people, blind to the experiment, were engaged in communal events and activities such as a rituals, ceremonies, meetings and musical concerts. The experiments asked whether continuously recorded sequences of random data might show structure during periods of group interaction which involved shared emotions or deep interest. These experiments were subsequently replicated by other researchers. The results indicated that deviations in the random data were correlated with periods of group activity or “group consciousness,” especially when people involved reported a sense of coherence or resonance within the group. Tests in which data were collected in mundane or unfocused situations typically conformed to expected random behavior.
The field work raised a number of issues which became the basis of the Global Consciousness Project. Among these are questions about the effects of mere attention or mental engagement as opposed to directed intention: is the latter necessary, or might RNGs be generally responsive to the consciousness environment?\textsuperscript{12} Working in the field with groups also suggested using multiple devices in a distributed network: would multiple, simultaneous data streams reveal different effects?\textsuperscript{15,16} Would the RNGs correlate with each other and would this be a function of their proximity to the event or their mutual separation? Other questions concern the impact of various qualities that characterize events: their size, coherence, emotional tone, importance, human vs. natural origin, etc.

In 1997 an effort was launched to engage these issues using a permanent, world-wide network of RNGs. The result was the Global Consciousness Project, which began data collection in August, 1998 and continues to this day.\textsuperscript{17,18} The GCP network is an instrument designed to capture indications of mind-matter correlations manifesting on a global scale. In practical terms, the project makes a conceptual leap from the single-device laboratory and field experiments which examined individual intention and group attention, respectively, to a multi-device experiment designed to look for effects of synchronized or coherent mass consciousness on a global scale.

**Method**

The proposition of global mind-matter correlations needs to be translated into an experimental hypothesis. Since we are breaking new ground, there is little history to guide hypothesis specification. We can, however, infer from the laboratory and field research described above that the effect would be small compared to the intrinsic noise scale of the data, and would most likely span a broad range of physical, social, and emotive conditions. We therefore work with a general hypothesis describing a range of conditions rather than a narrow set of parameters:

*Periods of collective attention or emotion in widely distributed populations will correlate with deviations from expectation in a global network of physical random number generators.*

The hypothesis avoids premature over-specification, but identifies the main elements we wish to test for: global correlations between collective conscious activity and the material world as represented by the physical RNG network. Experimentally, this general hypothesis is instantiated in a series of specific, rigorously defined hypothesis tests, each of which is compatible with the general statement. Technically, we propose a *composite hypothesis* which formulates our broadest guess of how global mind-matter correlations might be defined for the RNG network. We then proceed experimentally with a series of replications using *simple hypotheses* which are fully specified and can be compared quantitatively against the null hypothesis.

The term “global consciousness” evokes many ideas that differ from our intended usage, so it needs clarification. It is difficult to define consciousness in the abstract even though, as the saying goes, “I know it when I see it.” Since our approach to the GCP hypothesis is empirical, we create an operational definition stating what we do in the experiment, thereby defining pragmatically the object of investigation. That is, for the formal experiment we treat global or mass consciousness as a set of operations, rather than as an intellectual or theoretical construct. We want to study X and we do so by performing operations Y and Z. Specifically, we identify
global consciousness as the outcome of the operations constituting the formal replication series. This produces a precisely-defined experimental database which can be used to evaluate the general hypothesis.

The operational definition of global consciousness has a number of advantages. First, it avoids confusing our experimental proposal with a theoretical conjecture. The GCP hypothesis is not intended to describe a theoretical position, but is an experimental question motivated by prior research findings. Second, it allows us to determine a confidence level for deviations of well-defined statistics as a basis for further analysis. Finally, the replication series at the core of our definition is well-suited to study an effect with low signal-to-noise ratio.

**Procedure**

To set up a formal test, we first identify an engaging event. The criteria for event selection are that the event provides a focus of collective attention or emotion, and that it engages people across the world. Thus, we select events of global character but allow for variation in their type, duration, intensity and emotional tone. In practice, events are chosen because they capture news headlines, involve or engage millions of people, or represent emotionally potent categories (e. g., great tragedies and great celebrations).

Once an event is identified, the simple hypothesis test is constructed by fixing the start and end times for the event and specifying a statistical analysis to be performed on the corresponding data. These details are entered into a formal registry before the data are extracted from the archive. We select and analyze an average of 2 or 3 events per month. The selection procedure allows exploration, while the replication design provides rigorous hypothesis specification for each event.

Because the project is unique, with no precedents to provide information on relevant parameters, we began with guesses and intuitions about what might characterize suitable, informative events. Field research on group consciousness suggested that synchronization or coherence of thought and emotion would be important, so we focused on major tragedies and traditional celebratory events that bring great numbers together in a common focus.

While many observers assume we can and should follow a fixed prescription to identify “global events” this is not straightforward. To give specific examples, we could select a disaster if it results in, say, more than 500 fatalities. But this would likely exclude slow moving but powerfully engaging events such as volcanic eruptions or major hurricanes, and it would fail to identify emotionally powerful, extremely important incidents like the politically disruptive attack that destroyed the Golden Dome Mosque in Iraq in Feb. 2006, but killed relatively few people. What we try to do is to identify, with the help of correspondents around the world, events that can be expected to bring large numbers of people to a shared or coherent emotional state. The following is a partial, illustrative list of criteria that we have learned are useful, with examples:

1. Suddenness or surprise. Terror attacks, especially where they are not usual.
2. Fear and compassion. Large natural disasters, typhoons, tsunamis, earthquakes.
3. Love and sharing. Celebrations and ceremonies like New Years, religious gatherings.
4. Powerful interest. Political and social events like elections, protests, demonstrations.


Experience has led to considerable standardization, and for some kinds of events, pre-defined parameters can be applied. For example, events that repeat, such as New Years, Kumbh Mela, or Earth Day, are registered with the same specifications in each instance. For unexpected events, such as earthquakes, crashes, bombs, the most used protocol identifies a period beginning at or near the moment of occurrence, followed by time (typically 6 hours) for the spread of news.

About half the events in the formal series are identifiable before the fact. Accidents, disasters, and other surprises must, of course, be identified after they occur. To eliminate a frequent misconception, we do not look for “spikes” in the data and then try to find what caused them. Such a procedure, given the unconstrained degrees of freedom, is not statistically viable. There is no data mining, and there is no post hoc inclusion or exclusion of events. All events are entered into the formal experiment registry before the corresponding data are extracted from the archive. For details, see http://noosphere.princeton.edu/pred_formal.html. The analysis for an event then proceeds according to the registry specifications. All registered events are reported, whatever the outcome.

**Equipment**

The GCP is Internet-based and employs a network of RNG devices installed at host sites (nodes) around the world. A central server receives data from the distant nodes via the Internet and incorporates them into a continually growing database archive. Each local node comprises a research quality RNG (Mindsong MicroREG by Mindsong, Inc., Orion RNG by ICATT Interactive Media) which is connected to a host computer running custom software. The software collects one data trial each second, a trial being the sum of 200 consecutive random bits of RNG output. The bit-sum is equivalent to tossing a fair coin 200 times and counting the heads, yielding random values with a theoretical average of 100 and standard deviation 7.071. The bits are generated from physical random processes (specifically, quantum tunneling) in the RNG circuitry and are not created by a computer algorithm. Each RNG is calibrated with at least one million 200-bit trials, processed using a custom suite of tests developed at the PEAR laboratory which examines statistical distribution parameters (four moments), the arc-sine distribution, extreme value counts, run lengths, correlations and autocorrelations. The devices are shielded, and an exclusive or (XOR) logic operation eliminates bias from physical causes in principle.

The trials are time-stamped, written to the local disk and then uploaded from the local hosts to the network server in Princeton, NJ at 5-minute intervals. Custom data management software on the server stores all raw data in permanent archives. The result is an accumulating database of continuous parallel data sequences. The GCP design requires that the data be synchronized at one-second resolution. Host computers use network time protocol (NTP) or an equivalent for synchronization, and though we are aware of some failures, most hosts successfully maintain 1-second accuracy. (Unsynchronized data cause no problems other than increased statistical noise in the analysis.) Synchronous data generation means that we can treat the network as a single instrument, using statistical measures that address the whole network rather than treating the RNGs individually.
Figure 1 shows the location of host sites in the network, which grew to approximately 60 nodes in the first years of the Project, and since 2004, has been relatively stable with 60 to 70 operational nodes. We rely on volunteers to host and maintain the RNG device and software at each node. The geographical distribution of nodes is constrained by infrastructure limitations. While we aim for a world-spanning network – ideally a deployment representative of world population densities – network coverage is poor in areas where Internet access is limited. For example, we do not have coverage in many parts of Africa and Asia.

Figure 1. Google map showing locations of all RNGs that have been in the network and contributed data. The distribution depends on Internet infrastructure.

The GCP website at http://noosphere.princeton.edu describes all aspects of the project, ranging over its history, context, and technology. One of the important features defining the Project is transparency, and the website is a public access repository of information, including the entire archive of raw trial data, which is freely available for download. We maintain a complete record of the formal hypothesis tests and preliminary results from ongoing analyses, as well as contributions and critiques by independent, third-party investigators.

**Results**

Through January 2011, over 345 rigorously vetted, pre-specified events have been registered in the formal replication series, including tragedies and celebrations, disasters of natural or human origin, and planned or spontaneous gatherings involving great numbers of people. The events generally have durations ranging from a few hours to a full day. The Project registers about 30
formal events per year, and the data taken during these events comprise somewhat less than 2% of the 12-year, 25-billion trial database. The cumulative experimental result attains a level of 6.2 \( \sigma \) (standard deviations) relative to the null hypothesis. The odds of a chance deviation of this magnitude are about a billion to 1.

The formal result is obtained by first converting the test statistic for each event to a standard normal Z-score. The scores are averaged and the confidence level against the null hypothesis is given by the deviation of this average from zero. We find an average event Z-score of \(~0.33 \pm 0.06\) which yields the composite deviation cited above. The calculations assume that the RNGs have stable output distributions, and this has been extensively verified across the 12-year database.\(^{19}\) We do not, on the other hand, assume that the RNGs are perfect theoretical devices; the normalized Z-scores of the formal series are based on empirical estimates of mean and variance for each device, calculated from its entire data history. All analyses are checked for validity by running simulations on pseudo-random data sets, and the results are compared not only with theoretical expectation but with control distributions.

Figure 2 is a scatterplot of 346 Z-scores from the formal trials. The dashed horizontal line shows expectation and the solid line shows the mean deviation of all trials. This is obviously a small shift relative to the distribution, but it is highly significant because of the statistical power of so many replications. Examination of the scatter gives a visual impression of the distribution, which tests as normal about the mean value; it also clearly displays homogeneity over time.
To display the reliability and the compounding significance of the small effect, we can use a “cumulative deviation” plot of the running sum of deviations from expectation as the replication series accumulates. The event data are shown together with results from a random simulation in Fig. 3. The cumulative deviation of the actual event Z-scores is compared with the distribution of cumulative traces for 250 simulation series of Z-scores drawn randomly from the (0, 1) normal distribution. It is clear from Figure 3 that the event data are from a different population: they have a positive bias which is not present in the control distribution.

Figure 3: The bold jagged line shows the cumulative sum of deviations from expectation in the formal data. Grey lines show 250 simulated datasets drawn from the (0, 1) normal distribution. The horizontal line is null expectation and smooth parabolas show confidence levels.

A still more powerful control background is produced by resampling the non-event data (98% of the database) to generate clones of the formal data series using the same parameters, but randomly offset start times for the events. Repeated resampling (also known as bootstrap sampling with replacement) produces the empirical distribution of expected scores, which is statistically indistinguishable from the random simulation. It provides a rigorous confirmation that the GCP database as a whole conforms to expected null behavior, whereas the behavior at the times of events displays a persistent deviation. Resampling also verifies that our analytical procedures do not introduce spurious correlations. This de facto control database necessarily contains any systematic non-ideal behavior also present in the event data. Since the non-event database exceeds the size of of the event dataset by nearly two orders of magnitude, we can check for spurious effects with high precision.
The experimental trace in Figure 3 reveals several other important facts about the event data. First, although the trend is fairly steady, it fluctuates randomly about the average slope, as is expected for a weak effect dominated by random noise. Second, it is evident by inspection that the deviation is distributed smoothly over events; the cumulative rise is not dominated by a few outlier events. Formal testing shows the distribution of event Z-scores to be statistically indistinguishable from the expected normal distribution. Third, the average contribution of events is small.

This is an important point. While selected examples may look convincing, the small effect size means that a single event cannot discriminate against the null hypothesis. Given an average effect size of 0.33, an estimated 80 events are needed to attain a significance of 3σ (p-value 0.001) for a comfortable confirmation of the hypothesis. Even with a less demanding criterion, or a subset of events with a larger effect size, many replications are needed for an effect to be reliably identified. Simply put, the analysis of individual events cannot confirm the GCP hypothesis or identify anomalous effects during individual events with certainty. This is true even for extreme cases, such as the terror attacks on Sept 11 2001. Statistical noise is such a large component of the data that it can partially obscure – or, equally well, masquerade as – anomalous effects. Only in an accumulation of replications will real effects gradually become clear while the noise washes out.

**Secondary Research Program**

The essential experimental problem for the GCP is how best to study effects of the hypothesized global consciousness in data dominated by random noise. The solution is a two-stage research program. First, the replication series, which we refer to as the formal experiment, yields an aggregate score which estimates the overall significance of the composite hypothesis against the null hypothesis. The formal experiment is ongoing, and it can be likened to a continuing meta-analysis which updates the significance of a measured effect size with each new event.

The formal series is the foundation for a broader research program to examine parametric details and potential models. We characterize the data based on the fundamental RNG trial scores: the second-by-second outputs for each RNG in the network. Whereas the event Z-scores concisely summarize the formal result, the trials index a complete description of the experiment: trial values with their time-stamps for each device, the geographical position of the RNGs, and the event labels. A trial-level description permits analysis of any aspect of the experiment.

The secondary analysis program is motivated by the need to test various explanatory proposals against structure shown to be present the empirical data. It is largely the work of Peter Bancel, who has been studying the GCP data since 2002. Some of the basic results are publicly available,18-20,26 but several important findings remain to be published and can be presented only in descriptive form based on personal communications. Bancel has used a reduced dataset excluding trials from the first few months when the network had less than 10 nodes. Events longer than 24 hours, as well as a few events where the original analysis did not use the standard correlation statistic also are excluded. The reduced dataset yields virtually the same composite statistics as the full database.
Inter-RNG Correlation

Trial-level analyses demonstrate that the formal result is driven by the 1-second network variance, while the RNG state probabilities and autocorrelation conform to expectation. The network variance can be decomposed to show its relation to synchronized RNG-RNG correlations. Complete details are presented in a previous publication in which we show that analytical expressions of the formal result can be reduced to synchronized correlations between the RNG trials.\(^{18}\)

The correlation elements are expressed as the products of pairs of trial values, \(C1 = z_i z_j\), where \(z_i\) is the (normalized) trial value of the \(i^{th}\) RNG for one second, and similarly for \(z_j\). The elements of \(C1\) include all possible combinations of RNG pairs, subject to the restriction that the pair-products have identical time-stamps. It can be shown that the average value of \(C1\) is proportional to the average linear (Pearson) correlation between RNGs.\(^{18}\) Under the null hypothesis, the expected average value of \(C1\) is zero and, in this reformulation, a deviation in the mean value of \(C1\) corresponds to the non-zero average of the event Z-scores.

The event and trial-level formulations lend themselves to different interpretations. The event results confirm the formal predictions, and thus successfully identify an effect which we identify as operational global consciousness. The pair-product formulation provides more detailed information. Specifically, the \(C1\) measure shows that the effect is associated with synchronized correlations of RNGs in the network, thus providing physical insight into how the effect arises during events.

It is perhaps useful to provide an intuitive picture of the synchronized correlations represented by \(C1\). Imagine that the network of RNGs is replaced by buoys tethered at scattered locations across the ocean, and that the data acquisition consists of monitoring the height of each buoy, at each second, as it bobs up and down with the waves. The null hypothesis for \(C1\) describes them bobbing randomly, without apparent correlation. A significant positive value of \(C1\) corresponds to a substantial number the buoys bobbing up and down in unison. This represents an anomaly because we do not expect wave motions at distant ocean locations to be correlated.

A Second, Orthogonal Correlation

In addition to the primary statistic, \(C1\), which is specified in the GCP hypothesis registry and posted to the results page, there are other independent effects and correlations. These, as we will see, show structure that was not expected by anyone involved in the experiment, including the main experimenter, prior to about 2007, some 9 years into the project.

There is no reason, \textit{a priori}, to assume that the original network variance measure captures all anomalous deviations present in the event data. While there are in principle many statistics we could investigate, the simple expression for \(C1\) suggests a few forms to test. First, and most obvious, is the value of individual trials, \(z_i\), or more generally, the single trial moments of the form \(z_i^n\) which, taken together, represent the full statistical distribution of individual trials. We find that the single trial statistics in the aggregate conform to null behavior. This is an important result since it says that, within the accuracy of the experiment, direct perturbations of the individual trial scores are too small to measure. The formal experiment provides evidence of
significant correlations among RNGs, but we do not see evidence of anomalous deviations in the trial values themselves.

Second, the C1 statistic suggests a class of correlation products, $z_i^n z_j^m$. A straightforward algebraic analysis shows that, for integer (m,n), only the case $z_i^2 z_j^2$ is independent of C1. We refer to this correlation statistic as C2. It is a particularly interesting statistic because it has exactly the same structural form as C1, but represents a unique, orthogonal correlation channel. The identification of C2 comes solely from analytical considerations, and it is not measured by the formal replication. As with C1, the average value of C2 is zero under the null hypothesis, and a positive value indicates the presence of correlations. A calculation of C2 yields an effect size of that is statistically indistinguishable from the C1 effect size, but less significant because of larger variance. Control analyses using re-sampling on the entire database show that C2 conforms to null expectation in off-event data, and also confirm empirically that C1 and C2 are uncorrelated. *(Bancel, personal communication)*

It may be helpful to think of C1 as analogous to the mean of a distribution, while C2 is a measure of its variance. In the image of ocean buoys, we envision C1 as an anomalous correlation of their bobbing. A non-zero value of C2 might be visualized as a discovery that the magnitude of the movement is also correlated.

Our characterization analysis thus finds that the RNG network exhibits two orthogonal trial-level correlation channels. The C1 statistic underlies the formal result, while C2 is revealed by analysis to be a unique, alternate correlation channel, not measured by the formal experiment. The finding that the two effect sizes are of the same magnitude is useful for interpretations of the experiment. It suggests that mass consciousness, when defined in terms of pair correlations, is a more general effect than is indicated by the formal experiment alone. This is also an important result for modeling the apparent mind-matter correlations.

**Distance and Time**

So far, we have shown that operationally defined global consciousness, corresponds to correlations in the RNG network, and that independent correlations also appear in a parallel channel. We would like to know if the event data contain further structure and if the structure might relate C1 and C2 more directly. Two important questions to consider are whether the correlations depend on the location of RNGs, and whether the correlation strength evolves in time as an event unfolds. The trial-level description provides a basis for spatial and temporal analyses since the correlation statistics contain the RNG locations and trial times as parameters.

An immediate challenge is the choice of an appropriate measure for the tests. In the case of spatial structure, even events with a definite location, such as earthquakes or catastrophic accidents lack a ready parametric description for the distribution of global reactions. Consider the Indian Ocean Tsunami December 26 2004. Although the disaster was localized, the response

---

* Strictly speaking, the mutual correlation of C1 and C2 is identically zero under the null hypothesis. For convenience in calculations we employ a modified form which uses the pair-products of zero-mean quantities: $C2 = (z_i^2 - 1)(z_j^2 - 1)$. Integer powers of C2 are also uncorrelated with C1, but not with C2 itself, so we need only examine the lowest order, C2. Details will be presented in a forthcoming publication.
to the news of the event was widespread and complex. It is not clear without careful study what aspects of the reactions are relevant to the global consciousness we posit, or how to determine the impact on different regions of the network. Similarly, while the GCP hypothesis tacitly implies that effects will correspond to the event timing, it does not provide a metric for durations.

Despite these difficulties, both spatial and temporal structure are in principle detectable. Arguing from minimal assumptions based on the GCP hypothesis we expect that a characteristic of structure in the data correlations will be smooth variation, both in time and across the network. Smooth, large-scale heterogeneities in the data can be regarded as signatures of the posited global consciousness because they are not characteristic of excursions which occur purely by chance.

**Spatial Structure**

For spatial structure, a test can be devised from a linear regression of the correlation strength against the distance between RNGs. A general observation from the physics of spatially distributed complex systems is that correlations among interacting constituents tend to weaken as their separation grows. Thus a prediction based on physical intuition suggests that the correlation strength will decrease as a function of RNG pair separation. A test of this conjecture is constructed as follows.

The geometrical separations of the RNG pairs are calculated for each of the $10^{10}$ elements of C1 in the event data. The elements are sorted by distance into bins and the average values of the correlation strengths are calculated for each bin, allowing a regression of correlation against distance. A non-zero regression slope provides evidence of smoothly varying spatial structure, and the expectation is that the slope will be negative, meaning that the correlations weaken as separations increase. The broad deployment of the GCP network allows us to perform the test over distances which range from a few meters out to the earth’s diameter.

We learn that there is a distance effect by this metric. The inter-RNG correlations decrease as the geographic distance between RNGs increases, with a distance scale on the order of 8,000 to 10,000 Km. It is present in both C1 and to a lesser degree, C2, and the composite across both measures has a significance level approaching 3 $\sigma$. The regressions thus give empirical evidence for spatial structure and indicate that models will need to incorporate distance-dependent correlations in order to adequately describe the event data. The form of the dependence and whether it applies uniformly or only for certain kinds of events, are issues that remain to be resolved. These are challenging questions for analysis because of the small effect size. However, simulations of a numerical model demonstrate that a linear dependence on distance does provide a good representation of the data. Specifically, when we model the dependence by a pure linear decrease which declines to zero at the earth’s diameter, the distribution of expected model slopes is well distinguished from a no-slope model. We find that the regression value for the actual event data agrees well with the linear slope model, but is incompatible with models that do not incorporate distance dependence. (Bancel, personal communication.)
**Temporal Structure**

The GCP hypothesis proposes that data correlations will correspond to the human response to events, which first grows as an event becomes the focus of global attention, then persists for a time as people attend to the focus, and finally dissipates as attention wanes. The GCP test events are likely to incorporate sections of null data before or after the correlations because the formally specified periods make generous estimates of the event durations in order to maximize the likelihood that the full response is included in an event. The expected temporal pattern in event data will thus be substantial periods of correlation during the actual effect, typically bracketed by null sections. If this hypothetical picture is correct, physical intuition suggests ways to characterize the time structure.

For example, given two independent measures, C1 and C2, that show effects during the events, it follows that they both are driven in part by the same source. Making the assumption that there is a temporal correspondence between the effect and whatever drives it, we expect correlations between the two measures during the actual effect, but not otherwise. Testing this conjecture, we find that the correlation between the two measures approaches significance after 45 minutes to an hour, and then diminishes slowly. Correlations between the two orthogonal measures of network structure thus suggest that the effect duration is on the order a few hours, and typically is shorter than the event prediction.

Another perspective on time structure comes from consideration of the variance of the data during the formal events. If we concatenate the data for all events into a single vector and divide it into time blocks of equal length we can calculate an averaged block variance. Repeating this calculation for a range of block sizes gives the variance as a function of block length. In the absence of an effect, the block variance is determined only by the intrinsic fluctuations of the random data. However, for deviations surrounded by null periods as hypothesized above, some data blocks will straddle the cross-over region between deviating and null data sections, and the change in deviation within the block will make an extra contribution to the variance. When the block length is small, the excess contribution to the variance is small. As the block size approaches the length of the deviations, the excess is more substantial and the variance increases. The block length at which the variance attains its maximum value will indicate the time-scale of the effect. Analysis shows the block variance of the event data does change as predicted, and suggests structure with a time-scale of a few hours. (Bancel, personal communication.)

**Categorization Structure**

The events can be categorized to test the relative importance of psychological and sociological factors. Most variables of interest are subjective, but can be estimated with acceptable reliability using standard tools from the social sciences. Given the large sample of events, we can split the database into a small number of category levels (2 or 3, typically) which give reliable estimates. There are important caveats to note. Most important, there is often overlap or confounding of variables so the results of categorization are generally not independent. For example, we may be interested in comparing scores of events that engage millions of people against events which draw only a few thousand. It is easy to identify large and small events, but if we ask what determines the size, we quickly see a confound with the “importance” of the events.
Nevertheless, examining the characteristics and nature of the events in this way provides insight into what factors contribute to the anomalous correlations. Fig. 4 shows a sampling of results for categorization on various dimensions. We see that “importance” as indicated by estimates of the number of people engaged by the event is a substantial contributor to effect size. Events judged to be large or medium in size (importance) have an average effect size approaching 4 $\sigma$, while small events average about 2 $\sigma$. Similarly, events that are judged to evoke or embody great compassion have a much larger effect size than those showing little or none. On the other hand, assessment of the positive or negative valence of events indicates that, while many guess otherwise, there is no difference. Both positive and negative events produce an average effect size of roughly 3.5 $\sigma$, and events that are neutral an average effect of about 2.5 $\sigma$.

![Differential Effects Within Categories](image)

**Figure 4.** Formal events categorized by estimating the importance or the number of people engaged differ (left). Events rated for positive vs. negative emotional valence do not differ (middle). The rated level of compassion shows clear differentiation (right).

When we consider generalizations across the emotional categories, regularities appear. An instructive meta-category identifies the “direction” of the emotional or attentional source as external vs. internal. For example, terrorist attacks and natural disasters are external sources which impinge upon us, while meditations and celebrations have an internally generated quality, often with an outward focus. We find a characteristic distinction of effect size in the two correlation measures, C1 and C2, depending on this meta-categorization. Events in the external group tend to produce stronger effects in C1, while the internal group show stronger effects in C2, on average.
Models and Theory

We have two orthogonal measures of linear correlation in the GCP data: the original statistic specified in most of the formal hypothesis tests, and a second representing a correlation of variances. Both measures are significantly large compared to expectation. These two measures are correlated with each other during a portion of the formal events, and in addition they both show a decline in magnitude as the separation of the RNGs increases. These parametric results represent temporal and spatial structure in the event data, and they can serve as input for theoretical models of the deviations.

Three classes of models to consider are: 1) conventional explanations in terms of physical and electromagnetic fields, or conventional methodological errors or biases, 2) unconventional information transfer via fortuitous selection of events, experimenter intuition, or retroactive information flowing from future results, and 3) field-like models of consciousness or information sourced in individual human minds, or a non-linear field representing a dynamical interaction among minds.

Explanations of the formal experiment based on spurious effects can be rejected for the reasons detailed in descriptions of the GCP research program, and on the basis of empirical studies. Methodological leaks and systematic biases are precluded, respectively, by event specification and registration procedures which effectively blind the analysis, and by resampling controls which find no evidence of biases in the off-event data. Such explanations are also inconsistent with the multiple indications of unexpected data structure.

Proposals based on electromagnetic perturbations are among the most frequently advanced conventional explanations of the GCP results, but they can be challenged on a number of points. Design features of the RNGs and the network protect the data generation from biases, as previously described. Even if these protections should fail, it is unlikely that local EM fields could give rise to distant correlations among the RNGs. Lastly, direct analysis shows no evidence of diurnal variation in the RNG outputs, whereas ambient electromagnetic fields arising from the daily cycle of human activity would presumably induce a corresponding variation in the data. We do not see current proposals based on ordinary EM fields as viable explanations for the measured global correlations and data structure, but it would be premature to exclude entirely the possibility of subtle EM effects.

Models involving intuitive selection and retroactive information are variants of a theoretical position from parapsychology advanced to explain psi functioning. The general idea is that expectations and attitudes about the experiment play a role in determining the outcome. In the data selection case, the key notion is that deviations result from a fortuitous choice of timing rather than an actual change in the data. The measured anomalies are attributed to the selection of unlikely data excursions in a naturally varying sequence. The fortuitous selection is assumed to derive from the experimenter's intuition, or, more forcefully put, from precognition of the eventual results, which informs the choice of events, their timing and the test procedures. The C1 data deviations have been analytically tested against an explicit version of this model. The tests nominally reject the proposal, but are not sufficiently powerful to draw definitive conclusions. However, the preliminary conclusions are further supported by the model’s failure to accommodate the C2 correlation and the spatial and temporal structure found in the data.
The retroactive information proposal is based on time symmetry arguments.\textsuperscript{24} It suggests that experimental outcomes are linked to the future in a manner that is analogous to the apparently causal past. It implicates consciousness directly by claiming that unexpected data correlations can be explained as a desired future actualizing in the present. Retrocausal models are not developed to the point where they can be tested quantitatively against the GCP data but we note that no simple version will easily explain the varieties of structure seen in the event data.

Finally we consider field-type models associated with human consciousness. A simple version is similar to ordinary physical models in that it posits a field generated by a distribution of sources. The connection to consciousness is made by associating the field sources with conscious humans, while the field dynamics, which explain the RNG correlations, derive from the coherence of human activity during events. This proposal can accommodate all the inter-node correlations and structure seen in the data. However, it remains phenomenological since it does not explain how the field arises in terms of underlying principles.

A more complex proposal is that individual minds may be mutually interactive. In this view, interactions among the minds of individuals are responsible for an emergent field or property which depends on individual consciousness but is not wholly reducible to it. The proposal suggests that the dynamic and interactive qualities of consciousness also involve subtle interactions with the physical world and that these interactions are responsible for certain anomalous phenomena, such as are found in the GCP event experiment. It can be construed as embodying in a formal way the ideas of such thinkers as Teilhard de Chardin, describing a “noosphere” of intelligence for the earth,\textsuperscript{27} or Arthur Eddington, conceiving a “great mind.”\textsuperscript{28}

**Discussion**

The development of a new experiment presents challenges that can only be dealt with by trial and error illuminated by analytical results. Many aspects of the GCP experiment have no direct precedents. For example, selection and parameter decisions for the early tests were necessarily guesses. Although the hypothesis testing was fully scientific, no objective criteria were available for specifying the target events (other than untested, arbitrary schemes). This has been a concern of critics accustomed to formulaic parameter specification, and it deserves discussion.

As described early in this paper, our research program has multiple levels:

1. A general hypothesis states that we expect to find correlations linking world events and data anomalies. We test it by sampling a variety of events, expecting a range of effects depending on factors such as event importance, emotional impact, valence, and surprise. Events are chosen that are expected to engage large numbers of people and generate shared emotions. But we had at the outset only intuition and opinion to guide their selection.

2. Despite this, specific hypothesis tests rigorously evaluate instances of the general hypothesis. The test parameters are defined \textit{a priori} – i. e., prior to accessing the archived data. The test statistics are standardized, fully characterized, and independent. The results slowly (due to the small average effect size) educate us as to the types of events that do yield correlations, and they teach us, slowly, about appropriate specifications.

3. The composite across the accumulating specific tests is a continuously growing meta-analysis
of formal replications, which yields a confidence level for the existence of anomalous deviations corresponding to events. This constitutes our operationally defined consciousness correlations.

4. Secondary analyses characterize the correlations and establish parameters and constraints for the data anomalies. These become the necessary and appropriate input for modeling effects and identifying promising theoretical directions.

The first two items above are the core of a research program designed to permit exploration of unknowns while accumulating sound experimental data. A very small effect size means we need dozens of replications to achieve reliable statistics, so learning enough to set firm rules for event selection requires many years, given that we identify about 2 or 3 events per month. A decade of experience suffices to establish general (albeit incomplete) guidelines for the types of events we can expect to show effects, and provides guidelines for time periods that are adequate to capture the anomalous effects. The accumulating results of secondary analyses feed back to such standards, showing, for example, that the time period typically needs to be several hours long, but that the exact event length may not be critical.

The discovery of multiple measures of structure beyond the original formal test statistic means that there is convergent evidence for anomalous correlations in the data during defined events. If we consider the two correlation statistics and two structural parameters (time and distance) as independent indicators, their combined weight of evidence contributes to a high confidence level for the experiment, while also providing parametric information needed for theoretical modeling.

No current model is sufficiently developed to explain the experiment. Typically, theory and experiment work together to guide and advance research. However, the interplay between theory and experiment is weak when experimental hypotheses are merely empirical, without a well-developed theoretical basis. This is the case for the GCP event experiment, despite its robust result. It establishes a phenomenon but does not test any proposed mechanism or theory. From this perspective, the result is an extreme example of a scientific anomaly in that it calls for both physical and psychological explanations, without providing a clear theoretical link to either one.\(^{29}\) Of course, anomalies are not off-limits to scientific study, but they require a period of empirical effort before theoretical tools can be brought to bear on the problem.

While we certainly are still in this “empirical effort” mode, the search for data structure has produced results helpful in identifying which classes of model are more likely to be viable. For example our preliminary assessments indicate that a phenomenological field model can in principle accommodate all the structure we measure: C1, C2, and the time and distance parameters, while models based on selection or on EM interactions face serious challenges.

More fundamentally, the empirical results lay the groundwork for a progressive investigation of the hypothesis of operationally defined global consciousness, which we can summarize in three questions. We have partial answers to these questions, and future research will test and elaborate our provisional conclusions.

1. Is the effect physical?

We have argued from the data that models based on selection bias, whether from intuition or methodological flaws, are unlikely. In contrast, analytical structure in the data is congruent with field-type models that are consistent with true (physical) data anomalies. All the tests of temporal
and spatial structure as well as the derivation of the orthogonal correlation statistic derive from physical and analytical considerations.

2. Is the effect anomalous?

Conventional physical models are not viable. Beyond the empirical testing which indicates EM fields have no effects on the network, we know from first principles that conventional fields would generate unsynchronized data correlations, in contrast to what we measure. The GCP correlations are synchronized over thousands of kilometers. This synchronization of correlations is both a strong argument against conventional proposals and a useful constraint for any detailed model of an anomalous effect.

3. What characterizes a global event?

The experiment depends on defining “collective attention or emotion” to identify suitable events for study. This is the starting point for determining what underlies the effect. Our approach is empirical and begins with general considerations. Events can be classified into various psychological and sociological categories, and the categories’ relative importance for operational global consciousness can be tested. Preliminary work shows distinctions among category effect sizes, for example, when the events are sorted by emotional type. An important question is whether different types of events have discernibly different structural signatures in the data. The preliminary work showing differential responses for C1 and C2 suggests this is the case.

Conclusions

The GCP is a long-term experiment that asks fundamental questions about human consciousness. Our review describes evidence for synchronized effects of collective attention – operationally defined global consciousness – on a world-spanning network of physical devices. Careful analysis reveals multiple indicators of anomalous data structure which are correlated specifically with moments of importance to humans. The findings suggest that some aspect of consciousness may be a source of anomalous effects in the material world. This is a provocative notion, but it is arguably the best of several alternative explanatory directions. The convergence of several independent analytical findings provides strong evidence for the anomalies, and to the extent these can be integrated into scientific models they will enrich our understanding of consciousness.

Although a full exploitation of the structurally rich database is in early stages, substantial progress has been made in understanding the GCP replication experiment. The analysis already allows us to begin discriminating between theoretical approaches, and it provides tools for refining our general hypothesis. Future efforts will emphasize the human and participatory aspects of the events we study.

We have argued that the GCP experiment is not easily explained by conventional or spurious sources. Instead, we provisionally conclude that the anomalous structure is correlated with qualities or states of collective consciousness activity. While social and psychological variables are challenging to characterize, an obvious suggestion is to look for changes in the level of “coherence” among the people engaged by the events. Defining this construct and developing it empirically will be important for further progress.
In sum, the evidence suggests an interdependence of consciousness and the environment, but the mechanisms for this remain obscure. Substantial work remains before we can usefully describe how consciousness relates to the experimental RNG results beyond the empirical correlations. Our findings do not fit into our current scientific descriptions of the world, but facts at the edges of our understanding can be expected to direct us toward fundamental questions. As Richard Feynman remarked, “The thing that doesn’t fit is the thing that is most interesting.”

It is important to consider different theoretical scenarios. Quantum entanglement, retrocausation, active information fields, and other ideas have been discussed in this context, but these notions drawn from physics have only tenuous connections to the GCP experiment, and it is currently hard to see an entry point to any physical model. Here our research provides much needed input by establishing parameters that can help discriminate models. For example, quantitative modeling can determine whether a linear composition of sources incorporating the known parameters can produce the field-like data structure that appears to be consistent with our findings, or whether a more complex model is needed.

More broadly, the GCP results are of relevance for the study of mind and brain because they bear directly on fundamental questions of consciousness. The starting point for much research in conventional brain science is: What are the neural correlates which give rise to consciousness? This question assumes that consciousness reduces to brain activity. The GCP results urge us to ask a harder question: Are there direct correlates of consciousness to be found outside the brain? The question is challenging because it posits or points to phenomena that are anomalous and hence mysterious from a conventional standpoint. The search for understanding of mind and brain obviously must change dramatically if consciousness correlates are found in the broader world.

Finally, the GCP results inspire deeper questions about our relation to the world and each other. Might we find that the best model, after all, resembles a coherent, extended consciousness akin to Teilhard de Chardin's aesthetic vision of a noosphere? While this is a possibility that is currently beyond the supply lines of our scientific position, the experimental results are consistent with the idea that subtle linkages exist between widely separated people, and that consciousness is implicated.

What should we take away from this scientific evidence of interconnection? If we are persuaded that the subtle structuring of random data does indicate an effect of human attention and emotion in the physical world, it broadens our view of what consciousness means. One implication is that our attention matters in a way we may not have imagined possible, and that cooperative intent can have subtle but real consequences. This is cause for reflection about our responsibilities in an increasingly connected world. Our future holds challenges of planetary scope that will demand both scientific clarity and cooperative intent. On this we should be of one mind.

Acknowledgments

The Global Consciousness Project would not exist except for the contributions of Greg Nelson and John Walker, who created the architecture and the sophisticated software. Paul Bethke ported the software to Windows, thus broadening the network. Peter Bancel is the driving force behind the sophisticated secondary analysis program. Dean Radin, Dick Bierman, and others in the
planning group contributed ideas and experience. Rick Berger helped to create a comprehensive website to make the project available to the public. The Project also would not exist but for the commitment of time, resources, and good will from all the hosts of network nodes. Our financial support comes from individuals including Charles Overby, Tony Cohen, Reinhilde Nelson, Marjorie Bancel, Michael Heany, Alexander Imich, Richard and Connie Adams, Richard Wallace, Anna Capasso, Michael Breland, Joseph Giove, J. Z. Knight, Hans Wendt, Jim Warren, John Walker, Alex Tsakiris, and the Lifebridge Foundation. We also gratefully acknowledge online donations from many individuals. Finally, there are very many friends of the project whose good will, interest, and empathy open a necessary niche in consciousness space. The GCP is affiliated with the Institute of Noetic Sciences, which is our non-profit home.

References


