

The Haunted Life of Data

Lisa Blackman

Introduction: A Case for Small Data

In recent years, the politics of data, its social and cultural life, and the new methods that cultural theorists may need in order to analyze these, have become subjects of intense interdisciplinary debate across the humanities and social sciences (Ruppert, Law, and Savage 2013; Manovich 2013; Gitelman 2013). The politics of data open up the question of exactly what counts as data, especially in the context of the multiple media transactions that register our presences, both in work and play. These transactions leave traces, which potentially accrue “afterlives” (Fuller 2009). As David Beer and Roger Burrows (2013) suggest, data acquires its own social life, or lives, becoming lively in ways that are difficult to see, comprehend, and analyze solely using qualitative methods of inquiry. Data can be extracted, mapped, aggregated, condensed, measured, and translated, acquiring autonomies and agencies that extend and travel beyond the original event or transaction. Dystopian arguments present the increasing application of metrics to life as the final stage in technology acquiring its own agencies and taking command. Reminiscent of nineteenth- and early-twentieth-century dystopian anxieties, machines and, in this context, machine learning are seen now as governing humans in ways that are largely imperceptible, incomprehensible, and thus unpredictable. The so-called “back-end” of social media, for example, provides data that is conjoined with automated practices in complex ways. These recursive relations thus defy calls for transparency and raise ethical questions about data ownership. As many have argued, data repositories potentially create surplus value—revenue and profit for businesses, governments, science, and related actors. Particular data banks and archives are mined, often using proprietary forms of software, which can aggregate vast amounts of data in order to shape and anticipate the future; or this is at least the dream of those invested in a data-driven economy.

These data dreams and anxieties have converged around a new object of interdisciplinary inquiry, big data. This object attracts many scholars and supports a variety of claims, both positivist and post-positivist. The concept of big data also guides new digital methods and practices, including data visualization often based on predictive analytics, hyperlinks, or the frequency or variation of semantic content, such as co-word variation. One of the problems with these methods is that they remediate limited forms of semantic and network analysis and are often designed to produce aggregate correlation. They iron out dynamism, movement, and historical connections that are often significant aspects of **data's** social life (Beer 2013). When big data becomes the primary object, these aspects are often overlooked, especially when scholars mingle vastly different approaches to epistemology and the ontology of knowledge claims. One of the axioms governing these debates is that so-called big data require new forms of processing, new approaches and methods that challenge the well-rehearsed qualitative/quantitative divide (Savage 2013). This axiom is supported by funding councils, journal special issues, capital investment, and software platforms, methods, and practices. Some argue that this will potentially result in the “end of theory” (Anderson 2008). Although quantitative techniques have always been the main stay of more positivist methods, they are increasingly proffered as the solution to a range of problems that have been more central to the humanities: questions of power, agency, subjectivity, technology and embodiment, for example. Many researchers remain suspicious of such pronouncements, **while** at the same time **recognizing** that the *doing* of culture has changed (Beer 2013, 12: author’s emphasis) and that methodological innovation in the humanities is indeed necessary.

This chapter will focus on a case study of what might be called “small data.” I follow danah boyd and Kate **Crawford's** (2012) provocations about the mythology of big data and the assumption that “other forms of analysis are too easily sidelined” (666). The case study presented here is meant to exemplify why “in some cases, small is best” (**ibid.**, 670). It will allow the reader to grasp the micro-dynamics of data generation, and see how data traces can disclose patterns of activity missed if one is guided solely by metrics and computational forms of analysis based on predictive analytics. As boyd and Crawford argue, “all researchers are interpreters of data” (**ibid.**, 667) and big data is not necessarily better data. boyd and **Crawford's** interventions are made in relation to data generated in social media contexts. This will also be the focus of this chapter. But rather than allow the software platforms or API (application programming interfaces) to generate the data set, the analysis will be guided by a “hauntological” mode of inquiry, which follows the trail or traces of a certain event—what I

am calling the “John Bargh priming controversy”—and the different entangled temporalities that it enacts. This event relates to an area of cognitive science known as priming, and the controversy it sparked gained significant traction across social media. It has a historical lineage to psychic modes of experimentation which are part of psychology’s often disavowed past, and offers challenges to liberal conceptions of selfhood based on rationality, will, and cognition.

The different temporalities of what we might call “internet time” carried by the Bargh controversy are disrupted by temporalities which interfere, refract, disrupt, disperse, and cut together—which knot, splice, fracture and create discontinuities. These discontinuous times point toward multiple pasts in the present and futures-yet-to-come. This temporal complexity will be explored in the context of “uncivil discourse” (Ashley et al. 2013) fostered by the controversy as it unfolded online; that is primarily through the communication of sentiment, feeling and emotion, including anger, hostility, defense, pride and the subsequent management of reputation. However, rather than remain at the level of “uncivil discourse” and sentiment analysis, I will show how this controversy also revitalizes earlier historical controversies, which remain far from settled. These different temporalities appear as traces within the controversy, which set in motion and are carried by data that would otherwise be obscured and omitted by aggregate forms of analysis. As Karen Barad (2010) has argued, hauntological methods—associated primarily with the deconstructive writings of Jacques Derrida (1994)—allow one to think with dis/continuity, dis/orientation, and dis/jointedness. She argues that hauntologies work with erasures and the ghostly entanglements that make such erasures visible. Tracking these ghostly traces however takes work, a work that I argue is displaced if one is guided by particular data-logics and digital methods, such as co-word variation and predictive analytics. This work of interpretation implicates the researcher in what I have called, in other contexts and following the work of Grace Cho (2008), an ethics of entanglement (Blackman 2012).

What counts as data when following the Bargh controversy, which I will explore shortly, becomes a pertinent question. To call it a social media controversy would perhaps be inadequate; it would miss the recursive relationships between broadcast and social media that characterized how and why the controversy gained traction. Furthermore, as many scholars have argued, such a distinction is rather arbitrary given the remediation that characterizes media environments (Bolter and Grusin 2000). However, the controversy was enacted across various social media platforms, including blogs, microblogs such as Twitter, Google+,

Facebook, and caught the attention of broadcast media, including newspapers and television news. It also involved online science trade magazines and open-access science journals. The controversy was fueled perhaps inadvertently by the actions of particular individuals: the investigative prowess of a particular science journalist, the irate protestations of a cognitive scientist, and the defensive position of a variety of skeptics. However, their part in the controversy is not reducible to their actions or agency; in this sense each could be considered one of the actors within an assemblage of relations. Furthermore, as complex data scenarios make visible, recursive relations within assemblages of data are difficult to see. Still, the human actors did become intermediaries within a controversy that brings the past into the present and also gestures toward alternate realities, lost futures, or futures-yet-to-come. These lost futures and futures-yet-to-come are carried by the comparison of John Bargh to a Mr. Wilhelm von Osten, the trainer and owner of a horse called Clever Hans in the early twentieth century, who astounded audiences with his remarkable feats and talents (Despret 2004). Why this comparison might upset Bargh and what it discloses about the contemporary field of priming and automaticity research within psychology and cognitive science becomes the subject of the controversy. The significance of this comparison is to be found in digital traces, which throughout the controversy become the object of redaction, removal, erasure, and attempts by many users and interested parties to keep this association *alive*.

The controversy and the data generated is analyzed as an example of “digital hauntology,” particularly if we accept that haunting is a form of mediation, a history of the present, which “requires stretching toward the horizon of what cannot be seen with ordinary clarity yet” (Gordon 2008, 195). It involves phenomena within the psychological, cognitive, and brain sciences that are considered oddities, puzzles, and anomalies if one starts from the position that subjectivity is singular, bounded, and circumscribed by the exercise of human reason and self-control. Although the controversy might appear to be of interest only to psychologists or scientists, it carries other controversies that are of interest to arts and humanities scholars. These include: the politics and ethics of open-access publishing; the nature of communication within digital environments and across social media platforms; the politics of science communication and the public communication of science; the question of agency within digital environments; the status of digital archives; the problem with infographics; and the nature of the relationship between narrative and data.

My aim in this chapter is to reanimate the data trails that are obscured or omitted in the narrative framings of this controversy. The meaning and significance of the event, which has been constituted within particular apparatuses of knowledge production and interpretation, is based on giving a “single authoritative representation of the social” usually bound up with linear conceptions of time (Ruppert, Law, and Savage 2013, 39).¹ These forms of “bundled time” (ibid.) obscure, close down, and cover over the spread and circulation of data across different sites, sources, and temporalities afforded by digital devices. These enable new forms of traceability and knowing which allow contagion, flow, and nonlinear temporalities to be followed, tracked, and made more visible. The potential traceability and different temporalities of the digital information that is submerged or displaced by particular practices of knowing—human and nonhuman—is the subject of this chapter. As such, this intervention is strategically designed to show what exceeds data metrics when digital methods are linked predominantly to volume and scale (big data) and opens up a genealogical method which can mine, work with, and attend to the specificities of particular data sets and digital archives. These types of data remain as a persistent and ghostly presence, and can be followed, listened to, and brought to life. This is a political project, one that I hope will be of wider interest to all cultural theorists seeking a rapprochement with the sciences. The data disclose or reveal what exceeds positivism and present the potential dynamism of science—the possibility for science to listen to its own pasts in the present and shape futures yet-to-come.

The Bargh Controversy

John Bargh is a prominent cognitive scientist who received the American Psychological Association's 2014 Award for Distinguished Scientific Contributions. He was also made the James Angell Professor in Psychology at Yale University in January of that year. Two years earlier, in 2012, he became the subject of a controversy after he unleashed what was described on Twitter as a tantrum and “scathing personal attack.” These purported attacks, posted on his blog *The Natural Unconscious* under the titles “Angry Birds” and “Nothing in Their Heads” were directed at a young Belgian post-doc researcher, Stephane Doyen, and his team who had failed to replicate Bargh's 1996 study of priming. Priming studies use a range of different techniques and experimental apparatuses to attempt to influence behavior,

thought, and action in ways not consciously registered (Wetherell 2012). This area of cognitive science and psychology is already controversial with various claims being made, including those made in Bargh's 1996 study, which concluded that participants could be made to walk more slowly to an elevator after being shown words associated with aging. These words were presented in the form of a scrambled language task and were taken to have primed the experimental subjects in ways that they were not consciously aware of. Bargh's blog posts were also directed at the award-winning science writer Ed Yong who had commented upon Doyen's study and its failure to replicate Bargh's study in an entry on his blog *Not Exactly Rocket Science*, published in *Discover*.ⁱⁱ This post compared Bargh to Mr. von Osten the owner of Clever Hans, who was considered to be an unwitting participant in the performance of Clever Hans's prodigious feats and talents. Why this comparison might have upset Bargh is central to the unfolding controversy and opens up questions about experimenter subjectivity, which, in psychology, is considered to be circumscribed by detachment, relationality and to play a more executive role. The failed replication study and Yong's blog post motivated Bargh to rail against the experimental conditions of the Doyen study, the business model of the open access journal *PLoS ONE*, and the practices of science journalists such as Ed Yong.

The responses in the comments section to the blog framed Bargh's rebuttal as a rant, as defensive, inaccurate, and exemplary of some of the problems that accompany science debate in the era of PPPR (post publication peer review). PPPR refers to the afterlives of journal articles as they circulate within and across social media with the potential to extend review by contributing to comments sections of blogs and websites, for example. Although it is recognized that PPPR can extend the article's afterlife allowing a bigger readership and more publicity, for many, including Bargh, this comes at the expense of the integrity of science. Bargh's response, which appears to be an attempt to close down discussion, courted further controversy and placed Bargh at the center. His posts gained traction across social media and were amplified by the comments posted on numerous other blogs, on comments sections of the blogs, and discussion forums that picked up on the controversy. As you can see from the visualization below, the traction across social media, including Facebook, Google + posts, and Twitter is rather small when compared to the volume, scale, and quantity of big data sets. However, by following these data I hope to show why studying small data is important for cultural researchers and what this can reveal about the social life and politics of data and where data becomes compromised.

[Insert Figure 12.1 Here]

Figure 12.1 – The Bargh Controversy

I will reconstruct the data trails related to what has come to be known as the “priming controversy” by initially presenting the reader with four URLs (see Figure 12.1). Three of the URLs relate to the blog posts by Bargh published in *Psychology Today*. The other relates to the initial blog post written by the science journalist, Ed Yong, published in *Discover* magazine. As you will see from the above visualization, all four posts received minor traction across social media. For reasons that will become clearer, the first two URLs relating to two of John Bargh’s blog posts are now “offline,” that is, they have been removed from the scene but the traces of the original event remain. The interested reader can find hyperlinks to Bargh’s post, commentaries on the post, and responses to the responses, but the actual posts are now elusive. One is taken to *Psychology Today* and met with the automated response, “Page Not Found.” The subsequent erasure of these posts has been met with humor, puzzlement, and parody and the erasures have been substituted by various images. As a result of potential copyright issues I am unable to reproduce a particular meme, which included an image of Bargh overwritten with the caption, “Didn’t Replicate My Study? Must be Self-Published Crap.” These images draw attention to the micro-dynamics of scientific debate and the legacy of what were considered Bargh’s hostile blog comments.

However, the tenacious reader will be able to find the posts circulated within small Twitter communities. Ed Yong, the science writer, has posted “Angry Birds” (Bargh’s first blog post) in a Google cache link available for download as of the time of writing. I generated a visualization of my successful attempts, using Topsy.com, an open access Twitter analytic (recently bought by Apple), to find links to the deleted posts. Due to copyright regulations I have been unable to reproduce the screenshot and link for the reader. Topsy was used to recover tweets from Ed Yong, the science writer, who says that he has a copy of “Angry Birds” and wonders why Bargh has deleted the blog posts. Ed Yong, in a subsequent Tweet (eight months ago at the time of writing), has also recovered the erased post “Nothing in Their Heads” (Bargh’s follow-up blog post), which can be followed via the link: web.archive.org/web/20120307100648/http://www.psychologytoday.com/blog/the-natural-unconscious/201203/nothing-in-their-heads. Again due to copyright regulations I am unable

to provide a screenshot showing the link for the reader, but if an interested reader follows the link above they should be able to recover the post for themselves.

The recovered post was accessed via an Internet archive Wayback Machine, made available to the interested reader by Ed Yong. So that the reader can understand some of the motivations that may have led Bargh to remove the original blog posts, I will stage and reanimate some of the controversy as it unfolded, leading to the erasure of the posts. This erasure has left a series of digital trails and traces that reveal the original event's absent-presence and its hauntological potential, on which I will elaborate below. I will also draw on the recovered blog entries to both reconstruct the **micro**-dynamics of the controversy and to *perform* the hauntological potential of the data. This will involve taking the reader back to the early twentieth century and to a psychological archive related to priming which contemporary psychology has largely written out of its historiography (Valentine 2012). I will argue that it is primarily the splicing of two rather different temporalities within this controversy that reveals its wider significance and afterlives (Fuller 2009). This hauntological method allows the committed researcher to explore how data can be compromised by the entanglement of what Karen Barad (2010) terms different "spacetime configurings," as well as being compromised by software platforms, processes and practices and the intentions and motivations of human actors. All of these entanglements are carried by the data—and the data ghosts or absences—that perform the controversy. All of these modalities of becoming-compromised exceed what can be represented by data visualizations that are based on metrics, frequency, and co-word variation. The haunted life of data works against predictive analytics and discloses what can be done when software tools are integrated into more complex qualitative forms of cultural analysis. And, again, this is better illustrated in the context of small data.

Hans the Horse: Scene **1—Refraction**

[Insert Image 12.1 Here]

Image 12.1: Hans the Horse (Wikimedia Commons)

Ed Yong begins his blog post "Primed by Expectations" by referring to the figure of Clever Hans, a horse at the turn of the twentieth century who was able to tell the time and solve complex multiplication puzzles by stamping his hooves. Hans' prodigious talents were later

linked to his capacity to be moved by minimal unconscious movements unwittingly expressed by his trainer, von Osten (Pfungst 1911), or to what later became known within psychology as the “experimenters expectations.” As Yong suggests, the legend of Clever Hans has largely been forgotten although it has caught the attention of many contemporary cultural theorists interested in embodiment and affect (Blackman 2012; Despret 2004). Yong ends this evocative or perhaps provocative opening with the words: “But history, as we know has a habit of repeating itself.” The revitalization of the Hans controversy is warranted, Yong suggests, by Doyen’s recent failure to replicate Bargh’s aging study (1996). This is one of the most heavily cited classic psychology experiments on priming. Doyen et al.’s (2012) failure was published in an open-access journal *PLoS ONE*. This was after Doyen was unable to publish it within proprietary social psychological journals, which have a history of not publishing “null results” or non-replication studies. However, Yong does not make this point in this particular post. Rather, he uses Doyen’s failure to raise the issue of how we might understand priming effects, and the role experimenter expectation might play in producing these effects and affects (by provocatively comparing Bargh to von Osten). As a respected science journalist,ⁱⁱⁱ Yong carefully considers the range of arguments and parameters of the debates and opens up the problem of replication for an interested scientific community for debate, discussion, and consideration.

The post has twenty-five comments, which range from supportive to more hostile and defensive in relation to this comparison. Out of the twenty-five, the first seven comments are supportive and welcome the debate. There is then a post by Joe that shifts the terrain, which I have coded as “mildly hostile.” It corrects facts and does some work of undermining Ed Yong’s credibility. This is followed by a post from Ed Yong, which clarifies some of the information and openly responds to the comments. Joe responds with an escalation of hostilities and by mildly insulting Yong. The next post by another reader accuses Joe of being “unnecessarily derisive” and defends Yong’s reputation as a science journalist. This reader also does welcome Joe’s contribution to the debate and follows up some of the issues by comparing the different experimental setups and parameters of the original Bargh study and the Doyen failed replication. Joe responds to this by insulting the person who left this post and Yong, again. He brings up Clever Hans and states that Yong is wrong to make this comparison. As Joe argues, “saying history has a habit of repeating itself is wrong.” He claims that he does not mean to be unnecessarily derisive but Ed Yong is “WRONG.” The person Joe is responding to deflates hostilities by welcoming his thoughtful comments and

apologizes for his reading of Joe's posts, which he mistakes for "road rage" and says that he will ponder the arguments between football matches at the weekend. It is not clear whether this comment is sincere or sarcastic.

The next three posts by different readers respond to Joe, asking him for further evidence (that Bargh's study has been replicated by labs across the world), accusing him of being mean-spirited and undignified in his responses, and posting a link to <http://psychefiledrawer.org>, an online repository where researchers can post failed replication studies for discussion. The archive is based on what Robert Rosenthal, who coined the term "experimenter effect" within psychology, terms the "file drawer problem." This relates to the publishing bias toward positive results purportedly by many proprietary journals that only publish those experiments that replicate "classic" studies. Interestingly, when I followed this link on June 23, 2014, a non-replication of Bargh's 1996 study was the single most viewed experiment (with 15,105 views). The second most viewed was 7,718 by comparison. The next comment on March 7, 2012 posts a link to John Bargh's response to Doyen and Yong.^{iv} If readers follow the link they will be taken to *Psychology Today* and the automated response, "Page Not Found." Ed Yong responds with a link to a response he has made to Bargh's post.^v The remaining five posts (which include an acknowledgment of the issues by Yong) include an exchange between one reader and Yong, which corrects some facts and suggests that Yong responds to Bargh in the comments section below Bargh's post on *Psychology Today*, which has since erased the post and comments.

My account of this series of exchanges by a small readership might be considered a good example of how scientific debate is transforming within digital environments. As Ashley et al. (2013) have shown, such discussions are not always rational and can become defensive, sometimes offensive and easily polarized. They can threaten what Papacharissi (2004) has termed the foundation of so-called democratic discussion inherited from classical liberal notion of the public sphere; that debate should be deliberative and not imbued with emotion and affect. As Ashley et al. argue, online debate increasingly leads to what they term a "nasty effect" and threatens the public communication of science and public acceptance of science as primarily being based on truth, fact, and objectivity. The anxieties related to this by scientists are captured by the problems increasingly associated with social media and open access journals and forums. These anxieties surround the concept of post-publication peer review (PPPR), where many scientists argue that comments are available to the public as forms of scientific communication. These contribute to the public dissemination of scientific

knowledge and therefore are potentially damaging because of the unprofessional and often nasty tone of comments. The tone of discussion, as we have seen, is not governed exclusively by the norms and conventions of scientific debate and dialogue. Exchange often exceeds the norms of positivist scientific convention and remediates what might previously have been expressed via “closed doors,” including private conversations, in closed conference debates and in private hesitations expressed between relevant actors and agencies.

The Bargh controversy illustrates the assumption made by many media theorists studying social media that such forms of communication are *affective*, spreading emotion and feeling, rather than primarily being informational. This perceived problem and how it negatively impacts on scientific debate, exchange, and the public perception of science was summed up recently by *Popular Science* editor,^{vi} Suzanne Le Barre and her remarks following the decision to close the comment section of the website. She concludes that “comments are bad for science,” reasoning that

<Quote> commentators shape public opinion; public opinion shapes public policy; public policy shapes how and whether and what research gets funded—you start to see why we feel compelled to hit the “off” switch. A politically motivated, decades-long war on expertise has eroded the popular consensus on a wide variety of scientifically validated topics.<Quote>

My restaging of this scene is designed to give a sense of how personal dynamics are entangled with and spliced through different historical temporalities or “spacetime configurings” (Barad 2010). The present scene is shaped by the publishing of a non-replication study and the ad hominem exchange of sentiment and emotion that followed. However, the revitalization of Clever Hans by Ed Yong frames the controversy as also or even primarily being hauntological in nature. That is, the way in which personal sentiments and dynamics governing PPPR revitalize, refract, and bring back some of the lost futures of science, and particularly what Yong calls “the Legend of Clever Hans,” a legend that many psychologists will not have heard of, which takes psychology back to its intimate connection to psychic research and experimental practices (Blackman 2012). Ed Yong is intrigued by the way that psychology has actively suppressed and forgotten Hans and how Doyen’s restaging arguably takes us back to unresolved issues surrounding how a horse can be moved by a human (and vice versa). What questions and issues this might open up and present to studies

of priming in terms of the nature of conscious and unconscious communication is one that Yong suggests is important, even if and although it is disavowed by Bargh's posts. Yong suggests that this is the more controversial aspect of Doyen's study carried by Bargh's comments and the subsequent commentary and discussion. This joining together of a lost historical controversy with a current online social media controversy shows how a single event is more-than-one, and in the rest of the chapter we will go on to explore its more-than-one nature and how it diffracts these different temporalities.

Bargh's Retort: Scene 2—Resurrection and the Work of Interference

Prior to Bargh's "Nothing in Their Heads" post, he had not posted anything on his blog *The Natural Unconscious* for two years. He starts the blog by using the language of resurrection and preempting why he felt it necessary to use his authoritative position to debunk the Doyen study. His rebuke, he informs the reader, is directed at the business model of the open access journal *PLoS ONE*, (which published the Doyen study), and at what he calls "superficial science journalism." These are both threatening the integrity of science, he insists. *PLoS ONE*, he argues, uses a "pay-as-you-go" model of publishing, which does not demand rigorous standards of peer review or editorial direction. Bargh equates this to self-publication; it lets studies through which should not be published, he argues. Bargh then uses his own blog to engage in PPPR and to offer the expert review and editorial scrutiny, which he insists the article did not receive. He then argues that the self-published nature of the study and its misleading conclusions were exacerbated by the fact that it was commented upon by Yong's commentary on his blog *It's Not Exactly Rocket Science* and specifically the post "Primed by Expectations," which, as mentioned earlier, had some traction across social media. Bargh argues that he was not involved in peer review for this article and that, as the leading researcher in this area (with the most highly cited study) his input was needed to assess the significance of the Doyen non-replication study. Unsolicited, Bargh goes on to do this by critiquing the experimental conditions and set-up of the Doyen study and arguing that Doyen's study was a "bad replication study." He ends the post by drawing a link between "self-published studies" and online media sources, which are the actual problem in Bargh's view. They mislead, distort, interfere, and only work to skew debate in ways that are

damaging to scientific integrity. Bargh's post is written with the confidence of a Yale scientist who has "set the record straight."

There are forty-five comments on this post, also accessible via the web archive link posted by Ed Yong.^{vii} These posts perhaps disclose what is at stake in post publication peer review and how social media is transforming scientific debate. Eleven of the posts respond to what are framed as Bargh's mischaracterizations of the business model and editorial practices of the open-access journal *PLoS ONE*. They accuse Bargh of being defensive, of engaging in a rant, of being a bully, and of abusing his own authority and power as a "star academic" (particularly for having a platform on *Psychology Today*, which is considered by some of the commentators to be a "glam mag" with its own agenda). Bargh is also considered incredibly influential in his position as one of the reviewing editors on the board of *Science Magazine*,^{viii} for example. Bargh's response is also deemed to be "littered with inaccuracies and misrepresentations of his (own) work." At least eleven of the comment posts are directed at Bargh's characterization of the business model of *PLoS ONE* and include responses by *PLoS ONE* editor Peter Binfield. He accuses Bargh of factual errors, including that the journal is commercial (it is not-for-profit) and that it does not engage rigorous peer review and editorial direction. Bargh is directed toward the comments section of the Doyen article and encouraged to make his rebuttal there. The subsequent posts become more heated and make potentially slanderous comments about Bargh, engaging in what could be considered a defamation of character. There is a discussion of cyber-bullying and the aggressive practices of anonymous bloggers, as well as insinuations that Bargh's influence on what gets published is threatened by open-access publishers. Further there are suggestions that open access not-for-profit publishers could adopt a more "open" review and editorial policy. The key issue framing this debate is anonymity, both in terms of what anonymous peer review encourages (aggressive and negative reviews, for example) and also how anonymous PPPR (i.e., anonymous blog posting) affords the opportunity for bullying, slander, and defamation. I do not know why *Psychology Today* took down the blog post and the subsequent comments but clearly as a "data source" the traces disclose the micro-dynamics of scientific debate and how social media are transforming science debate and communication. They might be considered an interesting repository of compromised data.

Letting the Horse Lie: Scene 3—Resurrecting Ghosts and Diffractive Reading

Wow, this is going to come back to haunt him.

—Anonymous comment on Bargh’s *Nothing in Their Heads* blog post.

Although Bargh’s blog posts “Nothing in Their Heads” and “Angry Birds” have been taken down from *Psychology Today*, the posts are hyperlinked in complex ways to a distributed network of comments, responses, counter-responses, and links to actors, agencies, sites, and practices that have become part of the controversy’s extended life and afterlives. To that extent, this assemblage of relations is performative, extending the controversy in unforeseen directions, with unanticipated consequences. The data that carry these relations are visible and invisible, material and immaterial, covering the Internet like a spider web of present and ghostly traces. The four URLs in [Figure 12.1](#) act as attractors or mediators for these data traces, attaching themselves and becoming attached to complex networks of actors whose diffused relations are difficult to see and map. It is a network and therefore defies the logic of network analysis, which is based on being able to see and map relations of influence and association between actors. Networks often inadvertently include relations that are present and visible and those that are anonymous and invisible. Anonymity has plagued network analyses based on mathematical analyses, which produce cartographies that include human actors as well as automated agencies. The inclusion of nonhuman agencies, such as bots, can often skew analyses and include ghostly phenomenon such as deleted accounts, reduced to metrics and read as statistical analyses of influence. As Gerlitz and Rieder (2012) have argued, once these problems are recognized—that the quantitative always-already involves qualitative elements—the quantitative/qualitative distinction makes less sense.

What is behind numbers? As I have tried to show in this chapter, small data provides a useful lens to explore the complex and fascinating data relations that can emerge from even a minor online happening such as the Bargh controversy. As the reader will have become aware, these relations cannot be contained or sampled according to the affordances of software platforms and practices, such as Twitter and Facebook, whose digital devices generate their own data

Comment [A1]: AQ: Figure 1 has been changed to Figure 12.1. Please check.

sets. The data I have followed are cross-platform, are distributed and extend across different temporalities, and are set in motion by minor associations (Bargh with von Osten), which reanimate forgotten entities, anomalies, and controversies. These submerged relations exist as outliers to particular regimes of visibility, making it difficult to generalize about the influence, traction, and politics of data without engaging in some form of *interpretation*. As boyd and Crawford (2012) have argued, numbers do not speak for themselves. In this last section, or scene, I will focus on Ed Yong's response to Bargh's posts. My conclusion to this section will be that infographics and data visualizations that use, adapt, and modify digital methods derived from software platforms are constrained or compromised by all of the above issues. In some ways, I follow Bruno Latour's recent invitation to software developers and programmers in his keynote at the prestigious SIGCHI 2013 conference on Human Computer Interaction. In this lecture he challenges the research community to approach big data in ways that foreground the experience of data's connectedness and the relations these connections form through *time*. As I have found with this case study, the tools that might help embed hauntological modes of analysis within digital methods are still to be configured. The question then becomes what is and is not quantifiable within computational terms?

Ed Yong begins his rebuttal to Bargh's "Nothing in Their Heads" post with an image of a pram; toys are being thrown out of it. The post is titled "A Failed Replication Draws a Scathing Personal Attack from a Psychology Professor" posted on March 10, 2012. It is clear that although this is a carefully considered response to Bargh and the commentary around the study, it is also designed to be provocative. The entry is rather unsurprising and instigates a relevant and timely discussion surrounding the problem of replication within psychology and science more generally. This is the authorized narrative that surrounds this controversy—that the priming controversy discloses the problems with replication, which is considered the cornerstone of scientific innovation, discovery, and progress. As we have seen, this is the single authoritative representation of the event given by Bargh and many others. This is also confirmed by the sixty comments posted to Yong's entry, which for the most part are reasonable and useful discussions of the problems with replication and how this might be improved. Yong is generally welcomed for his careful consideration of this issue and his contribution to opening it up for extended discussion via social media and the mechanisms of PPPR. There are a few detractors from the parameters of this discussion which redraw what is seen as an unfounded association between John Bargh and the trainer and owner of Clever Hans, von Osten. One commentator puts it in thus: "Frankly, I don't think the tone of his post

was all that surprising given the insinuation in your post that he had fallen prey to Clever Hans.” A subsequent user also picks up on this: “So to compare Bargh’s findings with the Clever Hans story is insulting and misleading.”

These two comments disclose the hauntological potential of the data and act as traces that set in motion a genealogical trail. Although they are not representative of the majority of the commentary surrounding Yong’s entry, they point toward another scene haunting the problem of replication. This problem is overlaid and threaded through with the revitalization of an earlier psychological controversy, one that entangles the discussion with submerged narratives, ghostly figures, displaced actors and agencies, and different temporalities. Time is disjointed, dispersed and diffracted through itself, revealed in data traces that are fleeting and act as outliers to the central discussion and focus. In numerical terms, these traces are inconsequential and insignificant. However, they also point toward the ghosts of Clever Hans and von Osten, who courted controversy at the turn of the twentieth century and are reanimates in this present controversy. It would seem that John Bargh is followed by these ghosts and the area of priming within psychology has continually to be policed in order to prevent these historical associations from surfacing. This is confirmed in the first comment made by Bargh in his later and subsequently erased entry, “Angry Birds” (which I recovered using Wayback Machine as discussed earlier) posted in *Psychology Today* on March 23, 2012. His retort begins:

<Quote>The discussion sparked by my previous post has now far transcended the remarks I made in the post itself, in defense of our lab in the face of the “Clever Hans” charge. That was a slur on our lab that had to be responded to in order to set the record straight. Insults like that typically make people angry, and so a lot of heat was generated, but too much heat produces smoke, and smoke obscures clear vision. Let’s see if we can continue the discussion without anger and hostility clouding the real issues. <Quote>

The blog goes on to engage with some of the issues that have been raised in relation to the study and to put the record straight, again. The comments for this entry have not been recovered and no subsequent entries by Bargh have been made in *Psychology Today* since this post. As a cultural theorist interested in affect and phenomena that disrupt borders and boundaries between the inside and outside, material and immaterial, past and present, public

and private and self and other, have been fascinated by the unfolding of this controversy and what has taken form (see also Blackman 2012). I was led to this controversy by my own research into the phenomenology of will and what has come to be known as the “half second delay” within affect studies (Massumi 2002; Thrift 2007). I knew that the neuroscientific evidence underpinning this statement was controversial and through my own searches I was directed toward this particular social media controversy. Indeed, I became a very small part of it when I gave a talk at the Max Planck Institute in 2012 at an event called “Experimental Entanglements.” I argued that social media science controversies provide an important data-source for examining what Hans-Jorg Rheinberger (1994) has termed, following Derrida, the historical movement of a trace (its haunting, perhaps)—the tension between persistence and transformation, which he argues is not captured by Kuhn’s more totalizing notion of a paradigm and a paradigm shift to understand change and transformation within science. My contribution was tweeted by a participant to the event and now forms part of the corpus of tweets connected to Bargh’s “Nothing in Their Heads” blog post. Rheinberger convincingly shows how science controversies, although considered settled at particular times, have the tendency to resurface in new ways and forms. In his recent work (2010) he has extended this insight to consider what he terms the “economy of the scribble,” those traces of practices which exceed the parameters of recognized scientific practice.

Although Rheinberger in this instance focuses on what gets left out of studies once they are written up for publication (scribbles on pieces of paper, workings out, ponderings, etc.) I argue that digital archives represent opportunities to explore such economies within distributed, extended networks of actors, agencies, and practices. These exist as traces, which can be followed, mapped, listened to, and reanimated. Some of these traces are deleted or removed, and as I have tried to show, exist as ghostly presences, requiring ingenuity, tenacity, and some knowledge of software and data practices in order to cover and re-perform. They are lively and disclose some of what is carried by the social and cultural life of data. Data traces can be moved and re-moved, redacted as well as remixed, and require the work of articulation, translation and staging in order to be made visible. This is a different strategy to predictive analytics and is one that mines and focuses upon the potential of compromised data. These traces often become hidden or covered over by representationalism, by narratives and representations that come to stand in for the event. These obscure the more temporal aspects of events, the historicity of time, and the different temporalities that disclose the more-than-one nature of controversies and events. The strategy of panspectric

surveillance^{ix} led and extended by a hauntological mode of inquiry reveals the dynamic nature of archives and their unexcavated potential. It also opens up questions of ethics and ethical entanglements, how our attempts to repurpose and reimagine data are highly situated engagements. The focus here is much more on the performing and re-performing of archives, rather than conserving them as monuments.

Denouement: Setting the Record Straight

I have presented here a very partial restaging of the Bargh controversy. It forms part of a bigger book-length project, *Haunted Data: Social Media, Queer Science and Archives of the Future*, which follows three science social media controversies in the area of anomalous science, including Priming, Feeling the Future, and Hearing the Voice. I am working with digital methods, where useful, and on developing software tools that can remediate the hauntological dimensions of the analysis. This extends the work of what it means to follow a media object within digital environments. As David Beer (2013) has argued digital archives qualify how we approach movement and circulation when we take data as our object; this is also related to what Adrian Mackenzie (2005) has termed the “performativity of circulation.” My analysis extends work in the area of mediation within media studies to consider how social media is transforming science within the context of what Mark Deuze (2012) **terms** **“media life”**^x As a temporary denouement to this controversy, I will end with John Bargh’s Wikipedia page, which illustrates well some of the competing interests and contrary energies enacted within this controversy. This conclusion will illustrate how the distributed nature of this controversy and its extension across both space and time is closed down (again) through John Bargh’s Wikipedia page. I will approach this page as a form of representationalism or “bundled time” (Ruppert, Law and Savage 2013), which erases process, movement, and temporality, replacing these with what Sarah Kember and Joanna Zylińska (2012) term *animations*—images, narratives, and forms of knowledge which erase or occlude those traces or forms of haunted data that exceed particular strategies of *knowingness*.

As I have identified throughout this social media hauntology, one of the refrains or performative statements integral to John Bargh’s blog posts has been the desire to “set the

record straight.” In the process these actions have also threatened to damage his reputation, or have certainly created hostile comments from other users that verge on slander and defamation of character. As many social and digital media scholars have argued, digital archives are “archives in motion” (Beer 2013) and present dynamic and mobile processes which are not fixed in time. However, as we have seen, the capacity of data to move on from the original event and accrue “afterlives” (Fuller 2009) also threatens the performative strategies of individual users. As Zizi Papacharissi (2012) has argued, one of the issues facing social media users is precisely how to engage in forms of self-presentation and self-management in light of the complex digital relations that characterize social media practices. Goffman’s (1990) work on “impression management” is seen by many social media theorists as useful to explore the kinds of deliberative action that users might engage in order to manage these tensions and construct a coherent self-presentation. These tensions Papacharissi (ibid.) argues, are heightened within online social platforms, which intensify, she suggests, practices of self-monitoring and self-awareness. She argues that networked performances or performative strategies of subjectivity often resemble “micro-celebrity, personal branding, and strategic self-commodification” (1990).

Now to John Bargh’s Wikipedia page. Many prominent academics have Wikipedia pages that might be considered forms of personal branding. They tell a particular story about the academic’s life and career and establish relations of influence, status, and prestige within Google’s Page Rank algorithm (Rieder 2012). Using a Wikipedia Edits Scraper and IP Localizer,^{xi} it is possible to see the edited and erased content of this page. This includes attempts by other users to include references to the controversy and what has been erased by the removal of Bargh’s blog posts from *Psychology Today*. The edited history discloses the dynamism of this controversy and the work that has to be done in order to remove it from particular regimes of visibility. Although the Wikipedia page “forgets” the erased content, the Wikipedia Edit Scraper is not so forgiving and performs different regimes of remembering and forgetting. An example of this is provided in a screenshot below.

[Insert Figure 12.2 Here]

Figure 12.2. Wikipedia Edit Scraper and IP Localizer,
<https://tools.digitalmethods.net/beta/wikipedia2geo/>, courtesy of the Digital Methods Initiative, New Media and Digital Culture, Media Studies, University of Amsterdam.

Comment [A2]: AQ: Please provide the in-text citation for Figure 12.2.

Attempts to “set the record straight” online and to tell a particular story—bound up with specific **story**telling devices and narratives—is one that illustrates some of the contrary logics that govern data regeneration, movement, and analysis. The hauntological potential of data and the question of how we reanimate what gets erased from particular practices of remembering and forgetting is a crucial one for cultural theorists committed to exploring the politics of data. Science historiographies are written on the basis of particular narrative conventions often based on stories of linear scientific progress. As Michel Foucault cogently taught us, histories of progress are never simply histories of the unfolding of some purist notion of scientific truth. In this **analysis**, I have attempted to bring these **genealogical** insights into dialogue with a hauntology based on what Barad (following Donna Haraway) terms “diffraction as methodology” or rather “reading texts intra-actively through one another” (2010, 243). Within the context of digital and social media, this becomes less about *text* and more about data traces which entangle past, present and future, human and **non**human, affect and deliberation, and space and time in complex ways. This hauntological method allows the committed researcher to explore how data can be compromised by the entanglement of what Karen Barad (2010) terms different spacetime configurings, as well as being compromised by software platforms, processes, and practices, as well as the intentions and motivations of particular human actors. All of these entanglements are carried by the data and the data ghosts or absences that perform the controversy. In this way, I hope that what might be recognized as a distinctly queer or feminist engagement with what it means “to set the record straight” can be brought more explicitly into our understanding of where and how data becomes compromised and how we might animate or restage the haunted life of data.

Acknowledgments

I would like to thank Bernhard Rieder who gave proficient instruction in some of the digital methods I use in this hauntology. Also Lucas Freeman for his excellent **copy editing**.

References

Comment [A3]: AQ: The following references are listed in the reference list but not cited in the text. Please specify if they can be deleted from the reference list or please provide citation for “Anderson et al. (2013), Blackman (Forthcoming), Bargh, Chen, and Burrows (1996), Deleuze (1995) and Reider (2012).”

Comment [A4]: AQ: The following citations are cited in the text but not listed in the reference section. Please provide references for “Beer (2013), Ashley et al. (2013), Rieder (2012) and Deleuze (1990).”

- Anderson, Chris. 2008. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete." *Wired Magazine* 16(7). Accessed June 15, 2014.
http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory.
- Anderson, A. A., D. Brossard, D. A. Scheufele, M. A. Xenos, and P. Ladwig. 2013. "The 'Nasty Effect': Online Incivility and Risk Perceptions of Emerging Technologies." *Journal of Computer-Mediated Communication* 19(3): 373–87.
- Blackman, Lisa. 2012. *Immaterial Bodies: Affect, Embodiment, Mediation*. London: Sage.
- Blackman, Lisa. *Forthcoming*. "Haunted Data and Social Media Controversies: The Politics of Small Data." *Theory, Culture & Society*.
- Barad, Karen. 2010. "Quantum Entanglements and Hauntological Relations of Inheritance: Dis/continuities, SpaceTime Enfoldings, and Justice-to-Come." *Derrida Today* 3(2): 240–68.
- Bargh, J., M. Chen, and L. Burrows. 1996. "Automaticity of Social Behaviour: Direct Effects of Trait Construct and Stereotype Activation on Action." *Journal of Personality and Social Psychology* 71(1): 230–44.
- Beer, David. 2012. *Popular Culture and New Media: The Politics of Circulation*. New York: Palgrave.
- Beer, David and Roger Burrows. 2013. "Popular Culture, Digital Archives and the New Social Life of Data." *Theory, Culture & Society* 30(4): 47–71.
- Bolter, Jay and Richard Grusin. 2000. *Remediation: Understanding New Media*. Cambridge, MA: MIT Press.
- boyd, danah and Kate Crawford. 2012. "Critical Questions for Big Data." *Information, Communication and Society* 15(5): 662–79.
- Cho, Grace. 2008. *Haunting the Korean Diaspora: Shame, Secrecy, Silence and the Forgotten War*. Minneapolis: University of Minnesota Press.

- DeLanda, Manuel. 1991. *War in the Age of Intelligent Machines*. New York: Zone Books.
- Deleuze, Gilles. 1995. "Postscript on Control Societies." *Negotiations*. Trans. Martin Joughin. New York: Columbia University Press.
- Derrida, Jacques. 1994. *Specters of Marx: The State of the Debt, the Work of Mourning and the New International*. Trans. Peggy Kamuf. London: Routledge.
- Despret, Vinciane. 2004. "The Body We Care For: Figures of Anthro-zoo-genesis." *Body & Society* 10(2/3): 111–34.
- Deuze, Mark. 2012. *Media Life*. Cambridge, UK: Polity.
- Doyen, Stephane, O. Klein, C.-L. Pichon, A. Cleeremans. 2012. "Behavioral Priming: It's all in the Mind, but Whose Mind?" *PLoS ONE* 7(1): e29081.
- Fuller, Matthew. 2009. <http://www.spc.org/fuller/texts/active-data-and-its-afterlives>.
- Gerlitz, Caroline and Bernhard Rieder. 2013. "Mining One Percent of Twitter: Collections, Baselines, Sampling." *M/C Journal* 16(2). Accessed June 27, 2014. <http://journal.media-culture.org.au/index.php/mcjournal/article/view/620>.
- Gitelman, Lisa. 2013. *"Raw Data" is an Oxymoron*. Cambridge, MA: MIT Press.
- Goffman, Erving. 1990. *The Presentation of Self in Everyday Life*. London: Penguin.
- Gordon, Avery. 2008. *Ghostly Matters: Haunting and the Sociological Imagination*. Minneapolis: University of Minnesota Press.
- Kember, Sarah and Joanna Zylinska. 2012. *Life After New Media: Mediation as a Vital Process*. Cambridge, MA: MIT Press.

Mackenzie, Adrian. 2005. "The Performativity of Code: Software and Cultures of Circulation." *Theory, Culture & Society* 22(1): 71–92.

Manovich, Lev. 2013. *Software Takes Command: International Texts in Critical Media Aesthetics*. New York: Bloomsbury.

Massumi, Brian. 2002. *Parables for the Virtual: Movement, Affect, Sensation*. Durham, NC: Duke University Press.

Palmas, Kay. 2011. "Predicting What You'll Do Tomorrow: Panspectric Surveillance and the Contemporary Corporation." *Surveillance and Society* 8(3): 338–54.

Papacharissi, Zizi. 2004. "Democracy On-line: Civility, Politeness and the Democratic Potential of On-Line Political Discussion Groups." *New Media & Society* 6(2): 259–84.

Papacharissi, Zizi. 2012. "Without You I'm Nothing: Performances of the Self on Twitter." *International Journal of Communication* 6: 189–206.

Pfungst, Oscar. 1911. *Clever Hans (the Horse of Mr. von Osten): A Contribution to American Animal and Human Psychology*. New York: Henry Holt and Co.

Reider, Bernhard. 2012. "What is in PageRank? A Historical and Conceptual Investigation of a Recursive Status Index." *Computational Culture: A Journal of Software Studies*. http://computationalculture.net/what_is_in_pagerank.

Rheinberger, Hans-Jorg. 1994. "Experimental Systems: Historicity, Narration, and Deconstruction." *Science in Context* 7(1): 65–81.

Rheinberger, Hans-Jorg. 2010. *An Epistemology of the Concrete: Twentieth-Century Histories of Life*. Durham, NC: Duke University Press.

Ruppert, E., J. Law, and M. Savage. 2013. "Reassembling Social Science Methods: The Challenge of Digital Devices." *Theory, Culture & Society* 30(4): 22–46.

- Savage, Mike. 2013. "The Social Life of Methods: A Critical Introduction." *Theory, Culture & Society* 30(4): 3–21.
- Thrift, Nigel. 2007. *Non-Representational Theory: Space, Politics, Affect*. New York: Routledge.
- Valentine, Elizabeth. 2012. "Special issue on the Relations between Psychical Research and Academic Psychology in Europe, the USA and Japan." Editorial. *History of the Human Sciences* 25(2): 68–84.
- Wetherell, Margaret. 2012. *Affect and Emotion: A New Social Science Understanding*. London: Sage.

-
- ⁱ It is interesting to note that although "bundled time" is aligned by the authors primarily with qualitative (non-digital) methods, such as narrative analysis, the remediation of such tendencies—for example, the desire to provide an account or tell a story of an event bound up with particular story-telling devices—is being remediated within particular social media applications. Indeed, the desire for narrative sequencing and the capacity to "tell a story" about an event is remediated within a particular software analytics, known as *Storify*. This tool allows a user to represent a series of Twitter conversations and to create stories using social media. For an interesting account of the controversy see: <http://cedarsdigest.wordpress.com/2012/03/21/put-your-head-up-to-the-meta-a-peer-reviews-post-post-publication-peer-review-a-bargh-full-of-links>.
- ⁱⁱ "Primed by Expectations: Why a Classic Psychology Experiment Isn't What it Seemed," January 18, 2012.
- ⁱⁱⁱ Ed Yong has received numerous awards and honours for his science writing. This includes Winner: NUJ Stephen White Award for best communication of science in a non-science context, Association of British Science Writer Awards 2012.
- ^{iv} <http://www.psychologytoday.com/blog/the-natural-unconscious/201203/nothing-in-their-heads>.
- ^v <http://blogs.discovermagazine.com/notrocketscience/2012/03/10/failed-replication-bargh-psychology-study-doyen>.
- ^{vi} Popular Science is a website that exists to promote scientific innovation and to report on current trends and discoveries in the area of science and technology studies. See: <http://www.popsoci.com/science/article/2013-09/why-were-shutting-our-comments?>
- ^{vii} web.archive.org/web/20120307100648/http://www.psychologytoday.com/blog/the-natural-unconscious/201203/nothing-in-their-heads.
- ^{viii} Interestingly, if the reader uses a search engine such as Google to find a link to Bargh's role as a review editor for *Science Magazine*, one also finds a copy of Bargh's CV ranked number 2 on the page, closely followed in third position by Ed Yong's subsequent blog post, "A Failed Replication draws a Scathing Personal Attack." This will form the subject of Scene 3.
- ^{ix} Panspectric surveillance is a term used to describe forms of regulation, management and control, which actualise new diagrams of power (Palmas 2011). These are seen to be qualitatively different to Foucaudian panoptic forms of surveillance (DeLanda 1991). One version of what Deleuze (1990) termed "abstract machines" is that which brings together new data-mining techniques, the increasing digitalisation of cultures, with new ways of acting upon (human) subjects (including targeting what many refer to as pre-cognitive, or non-cognitive registers of experience). Thus, data analytics and strategies of pre-emption and anticipation that

increasingly organize software cultures are being taken up by business and consumer organizations using “sophisticated techniques that anticipate the propensities of customers to act in certain ways” (Palmas 2011, 339). Palmas’s (ibid) very interesting article in respect of this shift or trend calls for a “political economy of propensity” (352), which can contribute to an excess or “something else” to these debates. The project I partially outline in this chapter is an attempt to enact a “something else” in respect of the cultural politics of data, which mines or exploits where data might become compromised. It is a strategy that performs what lies in excess of current regimes of anticipation and pre-emption.

^x Deuze (2012) argues that there is no external to media life as media are now ubiquitous, immersive, and all-pervasive. As he argues, “media are to us as water is to fish” (x).

^{xi} <https://tools.digitalmethods.net/beta/wikipedia2geo>.