



**HAL**  
open science

## “Voodoo” science in neuroimaging. How a controversy transformed into a crisis

Romy Sauvayre

► **To cite this version:**

Romy Sauvayre. “Voodoo” science in neuroimaging. How a controversy transformed into a crisis. Social Sciences, 2023, 12 (1), pp.15. hal-03909050

**HAL Id: hal-03909050**

**<https://hal.science/hal-03909050v1>**

Submitted on 30 Apr 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



*social sciences*



Article

---

# “Voodoo” Science in Neuroimaging: How a Controversy Transformed into a Crisis

---

Romy Sauvayre



<https://doi.org/10.3390/socsci12010015>



Article

# “Voodoo” Science in Neuroimaging: How a Controversy Transformed into a Crisis

Romy Sauvayre <sup>1,2</sup>

<sup>1</sup> Laboratoire de Psychologie Sociale et Cognitive, CNRS, Université Clermont Auvergne, F-63000 Clermont-Ferrand, France; romy.sauvayre@uca.fr

<sup>2</sup> Polytech Clermont, Clermont Auvergne INP, F-63178 Aubière, France

**Abstract:** Since the 1990s, functional magnetic resonance imaging (fMRI) techniques have continued to advance, which has led researchers and non specialists alike to regard this technique as infallible. However, at the end of 2008, a scientific controversy and the related media coverage called functional neuroimaging practices into question and cast doubt on the capacity of fMRI studies to produce reliable results. The purpose of this article is to retrace the history of this contemporary controversy and its treatment in the media. Then, the study stands at the intersection of the history of science, the epistemology of statistics, and the epistemology of science. Arguments involving actors (researchers, the media) and the chronology of events are presented. Finally, the article reveals that three groups fought through different arguments (false positives, statistical power, sample size, etc.), reaffirming the current scientific norms that separate the true from the false. Replication, forming this boundary, takes the place of the most persuasive argument. This is how the voodoo controversy joined the replication crisis.

**Keywords:** controversy; neuroscience; voodoo; methodology; corrections for multiple comparisons; statistics; scientific norms; replication crisis; fMRI



**Citation:** Sauvayre, Romy. 2023. “Voodoo” Science in Neuroimaging: How a Controversy Transformed into a Crisis. *Social Sciences* 12: 15. <https://doi.org/10.3390/socsci12010015>

Academic Editor: Nik Brown

Received: 19 November 2022

Revised: 16 December 2022

Accepted: 21 December 2022

Published: 27 December 2022



**Copyright:** © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction: The History of Neurosciences

Neuroscience fascinates epistemologists and methodologists as much as it challenges them. A number of “neuro-disciplines” (Vidal 2009, p. 9) have emerged, such as neurophilosophy (Churchland 1989), neuromarketing (Ariely and Berns 2010), and neuroeconomics (Volk and Köhler 2012). Emerging initially in the fields of neurology, neuroanatomy, and neurophysiology, knowledge of the brain has since been incorporated into many different disciplines, thus, creating a vibrant academic and social dynamic (Littlefield and Johnson 2012a). As a whole, these advances have become known as the “neuro-turn”. Neuroscience has become omnipresent (Littlefield and Johnson 2012b), particularly in the United States, Great Britain, and Germany. In France, the publication of *L’homme neuronal* in 1986 by Jean-Pierre Changeux aroused the curiosity of the French public (Chamak and Moutaud 2014). This popularization increased due to the diffusion of brain imaging technology from the 1990s onward (Vidal 2009), causing that era to be described as the “Decade of the Brain” (ibid., p. 7).

Through the “neuro-turn”, different approaches became specialized and institutionalized in the form of particular disciplines—the so-called “neuro-disciplines” (ibid., p. 9). Such was the case for social neuroscience, which was initially called the “neurobiology of social behavior” (Adolphs 2010, p. 752); this discipline aimed to study “the relationship between neural and social processes” (Cacioppo and Berntson 2002, p. 3). For Cacioppo et al. (2006, p. xi), “the notion that humans are inherently social creatures is no longer contestable, either”. Social neuroscience is, therefore, based on the understanding of brain functions, and it includes social dimensions such as social recognition, attachment, attitude changes, and stereotypes and takes into consideration the interaction between emotion and cognition. In short, this discipline is at the border between neuroscience and social psychology.

Neuroimaging, also called brain imaging, measures brain activity using a variety of ever-evolving techniques. Technical innovations, particularly functional magnetic resonance imaging (fMRI), have been favorable with respect to the emergence of this discipline (Cacioppo et al. 2006). fMRI is a dynamic measurement in that it tries to shed light on brain activity by indirectly collecting the oxygenation of different areas of the brain. This collection of data is particularly complex and requires numerous precautions, corrections, and analyses. The principle is to compare two brain activities (active and at rest). For example, if the researcher wishes to measure brain activity during the recognition of a color, it is necessary to keep in mind that the brain is likely to react to the shape of the colored object that will be submitted to it. Thus, the researcher will choose two tasks with the same object with the same shape, but one will be colored and the other will be black and white.

In the 1990s, the use of fMRI became widespread, and studies of brain damage in monkeys and humans had a major impact on the development of social neuroscience (Adolphs 2010, p. 17). Although the term “social neuroscience” was coined in 1992 by psychologists John T. Cacioppo and Gary G. Berntson (Cacioppo and Berntson 1992), a variant conceptualized by psychologists Kevin N. Ochsner and Matthew D. Lieberman (Ochsner and Lieberman 2001) emerged in 2001 under the name “social cognitive neuroscience”. The latter is situated at the boundary between cognitive neuroscience and social psychology (Ochsner and Lieberman 2001) and includes an interdisciplinary combination of neuronal mechanisms, the cognition and processing of information, personal and social contexts, and individual behavior. This approach links cognitive neuroscience tools such as neuroimaging with questions and theories drawn from the social sciences (Lieberman 2007).

Interest in social neuroscience intensified in the early 2000s. Journals of psychology, psychiatry, neuroimaging, and cognitive neuroscience, such as *NeuroImage*, *Journal of Cognitive Neuroscience*, *Neuropsychologia* and *Biological Psychiatry*, devoted special editions to this emerging discipline. The first collective work on the subject was published in 2002 (Cacioppo et al. 2002). Two related journals were published in 2006. The journal *Social Neuroscience* was created “to provide a place to publish empirical articles that intend to further our understanding of the neural mechanisms contributing to the development and maintenance of social behaviors, or to understanding how these mechanisms are disrupted in clinical disorders” (Social Neuroscience. *Aims & Scope* n.d.). In contrast, the journal *Social Cognitive and Affective Neuroscience* (SCAN) aimed to build a bridge between neuroscience and social science (Lieberman 2006), ranging from the use of neuroimaging on humans or animals in neuropsychology to the study of “mental and physical health problems related to social and emotional processes” (About the Journal n.d.).

In this emerging disciplinary context, a lively scientific and media-fueled controversy—better known as the “Voodoo controversy” (Bruder 2019, p. 43)—notably highlighted the methodological flaws of social neuroscience. In the present article, this controversy will be retraced by presenting, step by step and statistical argument by statistical argument, the terms of the debate that set neuroscientists in opposition to statisticians, with the exclusive support of written evidence collected during the scientometric and documentary research conducted during this study. To accomplish this goal, after discussing the initiator of this controversy and the reactions of the scientific community by examining their statements, the main statistical and methodological arguments included in eight published contributions to oppose or agree with the findings and accusations of the authors who initiated this “Voodoo controversy” will be exposed. Three groups of researchers opposing each other with statistical arguments were identified. The aim of this article is, therefore, to trace the history of this controversy, bearing in mind the fact that “debates around statistical objects easily take an epistemological turn” (Dodier 1996, p. 416). In other words, this article stands at the intersection of the history of science, the epistemology of statistics and the epistemology of science.

## 2. The “Voodoo” Article: The Terms Provoke Reactions

### 2.1. Birth and Mediatization of a Controversy

The initiator of the “Voodoo controversy”, Edward Vul, has been an Assistant Professor of Cognitive Psychology (Vul 2020) at the University of California, San Diego (UCSD) since 2010. In 2005, already studying at UCSD, he developed an interest in applied neuroimaging methods. A student of psychology, he completed a Bachelor of Science under the direction of the psychologist Hal Pashler at UCSD and a Bachelor of Art in philosophy. Vul continued his academic career at the Massachusetts Institute of Technology (MIT), writing a dissertation in cognitive science under the direction of Nancy Kanwisher, Professor of Cognitive Neuroscience, and Josh Tenenbaum, Associate Professor of Computational Cognitive Science. Vul’s dissertation, which he defended in 2010, focused on uncertainty and decision making at the intersection of psychophysics, neuroimaging, and computational modeling. During this time, he also developed a mastery of neuroimaging techniques (fMRI). He was also made aware of the problems of statistical processing by an article by Chris I. Baker, then a postdoctoral fellow in his laboratory (Baker et al. 2007). He, thus, began to investigate the topic of the statistical independence necessary for voxel sample selection (Vul and Kanwisher 2010). In so doing, he explored the use of fMRI in social neuroscience and, more specifically, the abnormally high correlations found among subjects’ personalities; their emotions, such as fear; and specific brain areas.

In May 2009, although he had not yet completed his PhD, he published an article in the journal *Perspectives on Psychological Science* (PoPS) initially titled “Voodoo Correlations in Social Neuroscience”, which he coauthored with UCSD psychologists Christine Harris, Piotr Winkielman, and Harold Pashler (Vul et al. 2009a). In this article—hereafter called the “Voodoo’ article” for the sake of convenience—the authors presented an analysis of several published works and found abnormally high correlation rates, half of which they attributed to poor adherence to statistical requirements. Vul and colleagues then questioned articles published several years previously in prestigious journals such as *Science*, *Nature*, *NeuroImage*, and *Neuron*. Vul and colleagues conspicuously named the authors whose methodology they intended to disqualify. One such author was Lieberman, the co-founder of social cognitive neuroscience.

The “Voodoo” article became so widespread prior to its publication in late 2008 that it went “viral”, as Vul explained in an interview with *Scientific American* magazine (Lehrer 2009a). A neuroscientist he knew had received the article from seven different sources in addition to the authors themselves (Lehrer 2009a). The study spread just as quickly throughout the scientific community as it did via blogs and magazines. As early as 28 December 2008, the science-minded public began to become aware of this work via a blog post written by Amy Rogers (2008), which was quickly relayed throughout the blogosphere. On 29 December, the blogger Vaughanbell (2008) reported on the piece, as did the blogger *The Neurocritic* (2008) on 31 December and again on 5 January 2009 (*The Neurocritic* 2009). Many other blogs followed suit. Media coverage increased when, on 9 January, science journalist Sharon Begley (2009a) devoted an article in the generalist magazine *Newsweek* to this topic, which was repeated on 30 January (S. Begley 2009b). With nearly 4 million subscribers, this magazine played a major role in publicizing the study, which described the “voodoo science” practiced by the researchers mentioned by Vul and colleagues as having, as the journalist reported, presented results that were “too good to be true”. Begley (2009a) described the study as a “bombshell” for the incriminated scientists and an exciting development for science bloggers, who stood at the forefront of watching a respected discipline being torn apart. Scientific dissemination then took place beginning on 13 January with the publication of an article in the prestigious journal *Nature* (Abbott 2009). On 17 January, the magazine *New Scientist* devoted an editorial (*The New Scientist Staff* 2009) and an article (Giles 2009) to this topic, presenting, among other things, a mea culpa for having participated in the dissemination of studies that had been highlighted for their lack of methodological rigor (*The New Scientist Staff* 2009). On 29 January, the popular magazine *Scientific American* interviewed Vul, as part of which he discussed the birth of the study

and its conclusions (Lehrer 2009a). On 27 February (Lehrer 2009b), the magazine continued this discussion by featuring an interview with Lieberman, Vul's most vocal opponent. This preliminary dissemination continued until the article's publication in May 2009, with the article being mentioned by blogs such as *Discover Magazine's* Neuroskeptic (Neuroskeptic 2009) and *Stanford Law School* (Greely), as well as by magazines such as *Seed* (Bardin), *Wired* (Saini), and *Sciences Humaines* (Marmion 2009). Finally, in 2012, neuroscientist Daniel S. Margulies (2012), sociologist Svenja Matusall (2012), and historian Cornelius Borck (2013) traced the history of the "Voodoo" article. A total of 12 magazine articles, albeit no print media articles, transmitted the work of Vul and colleagues. This form of media coverage did not occur until 2009, although Vul and Pashler (2012) reiterated the account in 2012 through an article published in the journal *NeuroImage* and entitled "Voodoo and circularity errors". In terms of media coverage, the dissemination of the "Voodoo" article, thus, occurred from January to May 2009; however, it had a profound impact on the neuroscience community.

## 2.2. Protest by Neuroscientists

Indeed, in addition to this media coverage, which was perceived to be disproportionate and alarmist (Lindquist and Gelman 2009), the "Voodoo" article provoked a "firestorm" (Poldrack et al. 2011, p. 184) throughout the neuroscientific community as early as January 2009. This intense diffusion in scientific circles was the sign of a "substantial controversy" (Diener 2009, p. 272) that became particularly heated for several weeks (Matusall 2012).

Beyond the debate concerning the content of the article, scientists complained about the "confrontational and not constructive" tone (Lieberman et al. 2009b, p. 272) employed by the authors (Diener 2009). Vul and colleagues wrote most of the text of the "Voodoo" article in a professional and scientific manner, with the exception of a few sentences that could be perceived as brutal. Indeed, they claimed, for example, that in half of the studies reviewed, "the reported correlation coefficients mean almost nothing, because they are systematically inflated by the biased analysis" (Vul et al. 2009a, p. 281); they "should not be believed" (ibid., p. 285) because "it is quite possible that a considerable number of relationships reported in this literature are entirely illusory" (ibid., p. 274). In addition, they strongly urged these highlighted researchers to reanalyze their data.

Neuroscientists believed that Vul and colleagues had cast the statistical methods used in neuroimaging into disrepute (Margulies 2012; Nichols and Poline 2009) by demanding "that authors retract or restate results" (Nichols and Poline 2009, p. 291). The "Voodoo" article caused such a furor among neuroscientists that Vul apologized in late January 2009 for offending or embarrassing some authors (Lehrer 2009a). Vul and colleagues also changed the title of their paper at the request of the journal *PoPS*. The title ("Voodoo Correlations"), which Vul intended to be "humorous" (ibid.), was seen as "provocative" (Abbott 2009, p. 245). Some neuroscientists perceived the term "voodoo" as a reference to the use of fraud, in particular to the "voodoo science" (Park 2005) that had been accused by Robert L. Park of being intrinsically associated with fraudulent science (Lieberman et al. 2009b). In so doing, Vul directly challenged the character of neuroscientists. As the editor explained in an introductory text, the title change was intended to mitigate the animosity that the term "voodoo" had provoked among specialists by replacing it with "less sensational" terminology (Diener 2009, p. 272). This change was also a way of preventing the stigmatization of one particular discipline, namely, social neuroscience, which was explicitly mentioned in the title and in the first version of the text. The title was, therefore, stripped of the term "voodoo" and this disciplinary reference and reframed to encompass the study of social cognition, personality, and emotion more broadly. The title "Voodoo correlation in social neuroscience", thus, became "Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition".

Despite this change, the tone of the article and its "pointed attack on social neuroscience" (Lieberman et al. 2009b, p. 299) remained unpopular with social neuroscientists (Matusall 2012). These neuroscientists felt so attacked in terms of their practice, character, and reputation (Matusall 2012) that they responded on blogs and via online texts as early

as January 2009 (Gelman 2009; Jabbi et al. 2009; Lieberman et al. 2009a), as well as in an article published by the journal PoPS in May 2009.

This journal published the “Voodoo” article alongside an introduction by the editor, as well as several articles discussing the claims of Vul and colleagues (Barrett 2009; Lazar 2009; Lieberman et al. 2009a; Yarkoni 2009). The issue concluded with a response by Vul and colleagues (Vul et al. 2009b). Subsequently, an article concomitantly published in the journal SCAN also played a role in the controversy (Poldrack and Mumford 2009). Finally, almost a year later, on 23 June 2010, an article coauthored by Kriegeskorte et al. (2010) presented the relevant arguments in a question-and-answer format; however, this piece did not have a substantial impact.

Consequently, apart from the response by Vul and colleagues, as well as the summary of arguments published a year later (Kriegeskorte et al. 2010), eight texts directly contributed to the methodological controversy initiated by the “Voodoo” article from January to June 2009<sup>1</sup>.

### 2.3. The “Voodoo” Article and Its Findings

Beyond the issues raised by its mode of expression, this “Voodoo” article initiated a methodological controversy<sup>2</sup> that led to the awareness desired by Vul (Lehrer 2009a). As Margulies (2012) explained, following the prepublication appearance of the article, neuroscientists became concerned about the methods employed in their discipline, even asking whether the very foundations of their methods should be reviewed. Although neuroscientists criticized Vul and colleagues for their approach, they agreed that the problem highlighted by the “Voodoo” article was significant. According to Bennett and Miller (2010), this article caused neuroscientists to address issues of fMRI reliability, which had been a topic of little concern prior to 2009.

In terms of content, the authors of the “Voodoo” article made an observation: when they looked at studies of emotions, personality, or social cognition that were conducted via fMRI, they found that the results presented exhibited particularly high correlation coefficients, with a value of approximately 0.8<sup>3</sup>.

The coefficient  $r = 0.8$ , which was found by Vul and colleagues to have been reported most often in neuroimaging works, demonstrates a particularly strong link between a region of the brain and an emotion (e.g., fear). Intrigued by these abnormally high coefficients, Vul and colleagues tried to explore the causes of such results by focusing on the methods used by neuroscientists. Vul had already theorized concerning this topic in a book chapter (Vul and Kanwisher 2010) written prior to the publication of the “Voodoo” article. To accomplish their goal, Vul and colleagues selected 55 articles and interviewed their authors via a questionnaire to investigate the methods they had used to process the data they had collected via fMRI. Vul and colleagues found that 29 articles had reported high correlation coefficients (greater than 0.8) and that their authors had used a particular method to produce these results. In the statistical processing of the fMRI data, the neuroscientists identified the areas of the brain that were activated by comparing brain activity recorded during the control task to that recorded during the test task. In this context, it is possible either to conduct this analysis with respect to the whole brain (Friston 2006) or to limit the research to a particular area, which is called a region of interest (ROI). Neuroscientists can, thus, choose to limit their analysis to an ROI to reduce the risk of false positives (Poldrack et al. 2011) and, thus, avoid applying a specific correction to the whole-brain analysis, an approach that may be considered too conservative<sup>4</sup>. The selection of ROIs is done based on anatomical criteria (e.g., the amygdala region), functional criteria (e.g., the activated voxels of a subject who responded better to one task than to another), or both, as explained by Vul and colleagues. Half of the researchers in question, they continued, used the same data to select their ROI and to calculate their correlation coefficient. In short, these researchers calculated a correlation between each voxel and a threshold value (for example, the average activity of subjects during the task). Then, they selected the significant voxels, at which point they once again applied a correlation. In this context, a sort of circular reasoning and practice emerges; before applying their statistics,

the researchers select the areas of the brain that are most likely to be active, and they then discover a strong correlation in these same areas. For Vul and colleagues, this practice artificially increases the relevant correlations and can even lead to a detection of significance in cases where only noise is measured (false positives). They called this type of error the “nonindependence error” (Vul et al. 2009a, p. 279).

### 3. The Statistical and Methodological Arguments against Vul and Colleagues

We now explore the “Voodoo” controversy and the factors that transformed it into a methodological debate by examining more deeply the arguments made by the eight contributions that played an active role in the exchanges that took place at the beginning of 2009. These contributions can be divided into three types: (1) authors who were challenged directly, (2) supporters of the method of corrections for multiple comparisons, and (3) supporters of an alternative statistical approach.

#### 3.1. Justification and Response by the Challenged Authors

The 55 articles analyzed by Vul and colleagues involved 128 unique authors, but only five authors (or 4% of the total) responded publicly regarding the problem of the overestimation of the correlations reported in their work. These responses were included in two papers, one coauthored by Mbemba Jabbi, Christian Keysers, Tania Singer, and Klaas Enno Stephan (Jabbi et al. 2009) and the other by Matthew D. Lieberman, Elliot T. Berkman, and Tor D. Wager (Lieberman et al. 2009b).

Jabbi and colleagues were the first to respond formally to the “Voodoo” article in the form of a short, unpublished four-page article posted on the Northwestern University website on 13 January 2009 and reproduced by The Neurocritic blog, among others. Jabbi, a graduate of the University of Groningen, was then a postdoctoral fellow at the National Institute of Mental Health (NIH) in Bethesda. Keysers, a graduate of the University of St. Andrews was a professor of Social Brain at the University of Groningen. Singer, a psychology graduate of the Free University of Berlin, held the academic chair of social neuroscience and neuroeconomics at the University of Zurich. Stephan, a medical graduate from RWTH Aachen University, is an Assistant Professor in Computational Neuroeconomics at the University of Zurich.

Jabbi and Keysers were criticized directly by Vul and colleagues and included in the list of articles identified as having presented results from a non independent analysis of their data reported by an article published in 2007 in the journal *NeuroImage*. Singer was also implicated in two articles, one of which was published in the journal *Science* in 2004, while the other, coauthored with Stephan, was published in the journal *Nature* in 2006.

In their response, which was addressed to both Vul and colleagues and especially to the press, Jabbi and colleagues claimed that the analysis by the authors of the “Voodoo” article was mistaken and that it, in their opinion, contained obvious conceptual and methodological errors. Jabbi and colleagues, therefore, set out to demonstrate the invalidity of the claims of the “Voodoo” article in several respects.

First, the authors reminded Vul and colleagues that they had applied the expected statistical correction to avoid false positives, namely, corrections for multiple comparisons. They also noted that their work had been confirmed by other studies. This reminder concerning the replication<sup>5</sup> of results by different studies and in different laboratories was a strong argument, especially since, according to the [National Academies of Sciences, Engineering, and Medicine \(2019\)](#), replication of experimental studies is a way of increasing confidence in and the reliability of scientific results. Finally, the authors criticized the questionnaire that Vul sent to them; the questions asked were perceived to be ambiguous, and the questionnaire was viewed as incomplete.

In simple terms, Jabbi and colleagues mainly claimed that they were following the appropriate scientific standards mandated by the American Psychological Association (APA). The replication argument was the most persuasive in the scientific community, since



the more confirmation a study receives from other studies, the more valid and robust its results are perceived to be.

The most virulent and influential response from the scientific community was that of Matthew D. Lieberman, Elliot T. Berkman, and Tor D. Wager (Lieberman et al. 2009b). Published in May 2009 in the journal *PoPS*, this response was circulated on 27 January of that year on the websites of the universities of Harvard, the University of California, Los Angeles (UCLA), Northwestern, and Penn State, as well as more widely on scientific blogs such as *The Neurocritic*.

In 2009, Lieberman was an Assistant Professor of Social Psychology at UCLA and a graduate of Harvard University in the same field. Similar to Jabbi and colleagues, Lieberman was one of the authors directly implicated in the “Voodoo” article with respect to two articles of which he was the second author. The first article was published in 2003 in the journal *Science* (Eisenberger et al. 2003), and the second appeared in 2005 in the journal *Cognitive, Affective and Behavioral Neuroscience*. These two articles, thus, appeared 2 years and 4 years, respectively, after the publication of the seminal article on social cognitive neuroscience that Lieberman coauthored with Ochsner (Ochsner and Lieberman 2001). Berkman, the 2nd coauthor of the response to Vul and colleagues, was a PhD student in social psychology at UCLA under Lieberman’s supervision and defended his dissertation in April 2010. Finally, Wager, the 3rd coauthor, was a graduate of the University of Michigan, Ann Arbor, in cognitive psychology and was an Associate Professor at Columbia University in 2009.

Lieberman and colleagues began their rich, nine-page response by expressing outrage at the way that Vul and colleagues had criticized work in the field of social neuroscience, especially with respect to their tone. Previously, they explained that they had hardly ever encountered such an aggressive tone in scientific literature. This “Voodoo” article, they claimed, was designed to attract the attention of the press and the media. The response of Lieberman and colleagues was thus intended, as they noted very directly, to “set the record straight” concerning several inaccuracies and errors they claimed to have identified in the “Voodoo” article.

For these authors, the source of the “misunderstanding” and “frustration” (Lieberman et al. 2009b, p. 300) of the researchers mentioned in the “Voodoo” article was the result of the incomplete nature of the questionnaire employed by that study. Lieberman and colleagues thus endorsed the arguments made by Jabbi and colleagues. Vul was alleged to have erred by failing to ask more specific questions concerning the analyses carried out during the relevant stage of the mentioned investigations. In other words, the authors of the “Voodoo” article were so confident of their knowledge of the procedure used by social neuroscientists, despite the fact they did not work in the same field (discipline), that they failed to inquire into the practices of the authors whom they studied. Lieberman and colleagues, thus, challenged their description of the method used in social neuroscience. To support this claim, Lieberman and colleagues interviewed the authors of 23 of the 28 articles mentioned by the “Voodoo” article. They also addressed certain technical dimensions of fMRI data processing.

However, the most virulent attack on Vul and colleagues pertained to the proposed methodology of their meta-analysis. Lieberman and colleagues argued that the statistical reasoning presented in the “Voodoo” article was based on a weak methodology and assumptions. They highlighted the irony of having their research results challenged by such questionable reasoning. Lieberman and colleagues showed that the authors of the “Voodoo” article had not followed one of the most important standards of scientificity in the relevant field, namely, reproducibility. Indeed, the description of the method used to constitute the corpus studied by Vul and colleagues was so imprecise<sup>6</sup>, in the opinion of Lieberman and colleagues, that it prevented the replication of the study and, thus, generated bias. Lieberman and colleagues illustrated this point by conducting a literature search using one of the selected keywords, namely, “altruism”. They obtained articles that had not been included in the corpus, and they were unable to find a justification for this omission in the “Voodoo” article. Furthermore, they reviewed all the correlation

coefficients presented in the articles investigated by Vul and colleagues. The authors found that 54 correlations (25% of the available correlations) were excluded from the analysis of the “Voodoo” article for no apparent reason. In addition, three correlations were unduly added to the corpus. Lieberman and colleagues reanalyzed the newly collected data and concluded that they could obtain drastically different results even when using the same data as Vul and colleagues. Indeed, the means of the non independent and independent correlations were almost identical ( $r = 0.69$  and  $r = 0.70$ , respectively).

This simple comparison of means allowed Lieberman and colleagues to minimize the findings of the “Voodoo” article and, thus, to put the conclusions of that article into question. Finally, the inclusion of missing correlations completed the demonstration, leaving room for doubt concerning the “Voodoo” article. For Margulies (2012, p. 277), this “rigorous methodological rebuttal” by Lieberman and colleagues returned the debate to the proper perspective by the end of January 2009, after the “Voodoo” article had caused cognitive neuroscientists worldwide to worry about the very future of their profession. The arguments of Lieberman and colleagues had a major impact on the controversy, extinguishing it as quickly as it had begun. The controversy then transformed into a methodological debate fueled by the articles by proponents of corrections for multiple comparisons that appeared in May and June 2009.

### 3.2. The Proponents of Corrections for Multiple Comparisons

In the same issue of the journal PoPS, two responses focused more on putting the causes of the problem of overestimated correlations exposed in the “Voodoo” article into question than on the problem itself. These contributions were written by Thomas E. Nichols and Jean-Baptiste Poline (Nichols and Poline 2009) and by Lisa Feldman Barrett (2009). These articles were supplemented by a contribution from Russel A. Poldrack and Jeanette A. Mumford (Poldrack and Mumford 2009), which was published in the journal SCAN. The debate then underwent a strong statistical and methodological turn that led to a greater focus on the problem of multiple corrections.

In their short, three-page contribution, which was published in PoPS, Nichols and Poline (2009) provided a technical response to the “Voodoo” controversy. The first author, Nichols, was a statistics graduate of Carnegie Mellon University, and in 2009, he was a *Senior Research Fellow* in the Department of Clinical Neurology at Oxford University. The second author, Poline, graduated from the University of Paris 7 with a degree in biomathematics and was a specialist in the mathematical processing of MRI data at the NeuroSpin Brain Imaging Innovation Research Center at the Paris Saclay Atomic Energy Commission, University of Paris Sud 11. Accordingly, both authors were specialists in the statistical processing of fMRI data. In particular, Nichols adapted statistical correction methods specifically related to multiple comparisons to fMRI data, namely, the false discovery rate (FDR) (Genovese et al. 2002) and the family-wise error rate (FWE) (Nichols and Hayasaka 2003), which are designed to limit false positives (Type I error) and false negatives (Type II error).

After criticizing Vul and colleagues for the way in which they had approached their subject and the discredit they had cast on neuroimaging, Nichols and Poline agreed that the description of methods in neuroimaging papers was sometimes incomplete, as Vul and colleagues had noted. This well-known problem, according to Nichols and Poline, could be confusing and did a disservice to scientific discourse. However, they continued, the researcher and the reader share an unspoken responsibility: it is the researcher’s responsibility to clearly state his or her data, and it is the reader’s responsibility to know and understand the underlying techniques and methods used to interpret the researcher’s results correctly.

Nichols and Poline noted that Vul and colleagues had addressed what they believed to be a well-known problem in neuroscience, namely, the problem of multiple comparisons. However, they continued, this problem had been solved by an appeal to consensus in the early 2000s. Everyone had to apply corrections to repeated measurements collected by fMRI. The authors invited Vul and colleagues to reread several chapters of the book on methods that Nichols had coedited with Friston (2006). Nichols and Poline, thus, seemed

to be giving a “statistics lesson” to the authors of the “Voodoo” article. After challenging the conclusions of those authors and reiterating the basics of neuroimaging data analysis, Nichols and Poline virulently criticized Vul and colleagues for presenting neuroimaging statistics as flawed. Although they praised the authors of the “Voodoo” article for initiating the controversy, they noted that their tone could have been more “moderate” (Nichols and Poline 2009, p. 292) and used “less alarmist rhetoric” (ibid.).

Nichols and Poline’s argument was based on the premise that corrections for multiple comparisons should be applied to any study that uses fMRI data. Jabbi and Lieberman made the same point. Since it was obvious that these corrections were applied and that they prevented false positives, the argument made by the “Voodoo” article, which stipulated that the significance found by studies in the field of social neuroscience could only be the result of manipulation, was invalidated.

Poldrack and Mumford also focused their argument on the issue of corrections for multiple comparisons. These authors were quick to respond to the “Voodoo” controversy, even though their article appeared after the PoPS articles. Indeed, they submitted their 6-page response to the journal *SCAN* on 28 February, which accepted it on 6 March and published it on 1 June 2009.

Poldrack was a graduate student in cognitive psychology at the University of Illinois at Urbana. In 2009, he was a professor of psychology, psychiatry, and behavioral neuroscience at UCLA, later becoming a professor of psychology and neurobiology at the University of Texas at Austin. Mumford was a graduate of the University of Michigan, Ann Arbor, with a focus on biostatistics as applied to fMRI data, and in 2009, she was an Assistant Research Professor in the Poldrack Lab at UCLA. Her work at the laboratory began in 2006. Since that time, Poldrack and Mumford have regularly co-published articles or books on functional neuroimaging methods.

These authors considered the arguments of the “Voodoo” article to be “simple and statistically incontrovertible” (Poldrack and Mumford 2009, p. 208). They also cited another methodological meta-analysis that had attempted to demonstrate this same problem under the heading of “circular analysis” (Kriegeskorte et al. 2009)<sup>7</sup>. However, Poldrack and Mumford disputed the causes of the problem discussed in the “Voodoo” article and, similar to Nichols and Poline, pointed out the necessity of corrections for multiple comparisons. To illustrate their point, Poldrack and Mumford used an example that they considered to be “instructive”, namely, the neuroimaging of dead salmon by Craig M. Bennett and colleagues (2010). They, thus, reported that by using an uncorrected method with a threshold  $\alpha$  of 0.001, Bennett and colleagues found active areas in the brains of dead salmon. However, when they used corrections for multiple comparisons, these active areas disappeared.

Poldrack and Mumford also challenged the hypothesis of Vul and colleagues with statistics. They concluded their argument with the following statement: “The problem of multiple comparisons is well known but unfortunately many journals still allow publication of results based on uncorrected whole-brain statistics” (Poldrack and Mumford 2009, p. 209). Applying different corrective methods for multiple comparisons would prevent outliers, they added. They also urged researchers to use more robust analyses whenever possible, since it was known that correlations between fMRI signals and personality tests were affected so frequently by outliers that the correlations obtained did not truly reflect the group studied. Finally, they urged researchers to detail their methodology no matter what approach they used because it should not be necessary to send a questionnaire (such as that of Vul and colleagues) to authors to ask about their methods.

The third contribution to this argument focused on corrections for multiple comparisons was made by Barrett (2009). She aimed to build a bridge between psychometrics and neuroimaging. At the time of the “Voodoo” controversy, Barrett, a clinical psychology graduate of the University of Waterloo, was a professor of psychology at Boston College and an Associate Neuroscientist in the Department of Radiology at Harvard University’s Massachusetts General Hospital.

In her five-page contribution to the “Voodoo” controversy, she considered the findings of Vul and colleagues to be important and saw an opportunity to revisit the history of psychology. According to Barrett, the “Voodoo” article highlighted certain problems related to the measurement of brain activity and the translation of this measurement into knowledge concerning the functioning of the mind. Neuroimaging would then have much to learn, or rather to relearn, from the methods developed by Wilhelm Wundt at the end of the 19th century and from the errors encountered in physical measurement and psychometry in general. Barrett then engaged with the theories of measurement elaborated in the framework of the construction of psychometric tests to extract the “three lessons” that seemed to her to be applicable to neuroimaging:

- (1) Reliability (the reproducibility of a measurement) should not be confused with validity (the meaning of the measurement);
- (2) The statistical error should be estimated correctly;
- (3) The statistical dependence of the experimental subjects should be taken into consideration when interpreting the results.

Barrett noted that Vul and colleagues had highlighted the frequent confusion between the reliability of the measurement and its validity. Barrett agreed with the findings of those authors but added that the best way of avoiding these errors was to use corrections for multiple comparisons. Such corrections can lower the magnitude of the correlations obtained. Barrett, similar to Nichols and Poline, as well as Poldrack and Mumford, thus, put the issue of multiple comparisons at the center of the controversy surrounding the question of the reliability of such measurement.

### *3.3. An Alternative Statistical and Methodological Approach*

While the practice of multiple comparisons seemed to be an appropriate response to the errors of nonindependence highlighted by the “Voodoo” article, three contributions approached the problem in a different way and proposed competing paths and explanations. [Yarkoni \(2009\)](#), [Lazar \(2009\)](#), and [Lindquist and Gelman \(2009\)](#) proposed such alternative statistical and methodological arguments.

When the “Voodoo” controversy erupted, Yarkoni was a PhD student in cognitive neuroscience under the supervision of Todd S. Braver, Professor of Neuroscience and Radiology at Washington University in St. Louis; he defended his dissertation in August 2009.

In his five-page paper, Yarkoni roughly paralleled Poldrack and Mumford by disputing the causes of the problem highlighted by Vul and colleagues; however, he began his remarks by expressing his agreement with the authors of the “Voodoo” article. Thus, Yarkoni explicitly agreed with the main conclusions of the “Voodoo” article regarding the overestimation of correlations observed in social neuroscience work. However, he disputed the fact that the overestimation of correlations was primarily attributable to the non independent analyses discussed by Vul and colleagues. Indeed, he explained that the sample size, significance level ( $\alpha$ ), and effect size involved in the relevant studies could overshadow the hypothesis proposed in the “Voodoo” article. Yarkoni ultimately focused on the problem of statistical power<sup>8</sup> in fMRI studies as the most relevant hypothesis, whereas [Poldrack and Mumford \(2009\)](#) focused on the problem of multiple comparisons.

These three dimensions (significance level, sample size, and effect size) were, for Yarkoni, the proper focus of the debate, a point for which his main reference was a book chapter he coauthored with Braver ([Yarkoni and Braver 2010](#)) prior to the publication of this response in the journal *PoPS*. This chapter argued that these dimensions could cause researchers to overestimate the importance of their results and cautioned them to remain wary of such a possibility.

In his demonstration, Yarkoni presented several computer simulations in which he varied the sample size. Yarkoni showed that the smaller the sample size was, the greater the statistical power and effect size. Yarkoni pointed out that for the studies in question to attain the statistical power conventionally attributed to 0.8 ([Cohen 1988](#)), they would have to be carried out on a corpus of 29 subjects for a threshold  $\alpha = 0.05$  and 60 subjects for a

threshold  $\alpha = 0.001$ . However, the sample sizes used in the studies emphasized by Vul and colleagues included fewer than 30 subjects. However, again by reference to simulations, Yarkoni showed that as the sample size increased, the average correlation coefficient found by whole-brain analysis decreased. Although this coefficient could exceed  $r = 0.8$  with fewer than 20 subjects, the coefficient decreased below this limit as soon as the number of subjects increased. Thus, Yarkoni demonstrated the influence of sample size on the correlations obtained in fMRI studies.

Yarkoni concluded his discussion by cautioning researchers and making several proposals for improvement. The problem of overestimated correlations could lead researchers to misinterpret the data obtained from fMRI. Conversely, he agreed that equal attention should be given to the ability of tests to detect between-variable effects (Type II errors) because with 20 subjects, the researcher had a 61% chance of detecting significance at the threshold  $\alpha$  of 0.05 ( $p = 0.61$ ). This rate, which was deemed inadequate by Yarkoni, could drop to 12% at the threshold  $\alpha$  of 0.001. To him, this fact indicated that the majority of fMRI studies detected only a tiny fraction of observable effects. Yarkoni concluded his write-up by echoing Vul and colleagues in claiming that correlations between 0.7 and 0.8 should not be trusted, as they were too good to be true.

The second statistical contribution to the “Voodoo” controversy was made by Lazar (2009), a statistician by training and a specialist in fMRI data analysis. A graduate of the University of Chicago, in 2009, she was a professor of statistics at the University of Georgia. Alongside Christopher R. Genovese and Thomas E. Nichols, she coauthored the founding article on the FDR method of multiple corrections for neuroimaging (Genovese et al. 2002).

In her short, two-page response to the controversy initiated by Vul and colleagues, Lazar showed that she did not share the excessively pessimistic and “bleak picture” (Lazar 2009, p. 309) of the neuroscience community described by the “Voodoo” article. Nevertheless, she agreed with the main findings of the “Voodoo” article. Discussing the problem of the nonindependence error that the article highlighted, Lazar nevertheless preferred to replace this term with “selection bias”, which referred to the field of sampling. She then noted that this problem was neither new nor specific to neuroimaging. Indeed, sampling biases are found particularly in online surveys, in which context people with the strongest convictions are most likely to answer the questionnaire and are not representative of the overall population. Lazar noted that this bias is also found in the sciences in the form of publication bias, according to which articles reporting significant results have a greater tendency to be published. When meta-analyses were conducted to calculate the effect size of the published articles in question, an overestimation bias was found.

For Lazar, the presence and maintenance of this selection bias despite the awareness of the scientific community was the result of two joint effects: (1) the voluminous and complex nature of the data to be analyzed and (2) difficult scientific questions. According to her, these two causes prompt researchers to lose their lucidity to the point of using the same data twice and making mistakes. Opposing the proponents of corrections for multiple comparisons, Lazar argued that while such an approach could limit false positives (Type I errors), it was insufficient.

Lazar agreed with Vul and colleagues by remarking on the effectiveness of the corrections for multiple comparisons method. In so doing, she noted that it was necessary to develop new specific methods, even if achieving this goal entailed a departure from linear models and commonly used correlations. This shift was precisely what Lindquist and Gelman proposed.

Lindquist and Gelman intervened in the “Voodoo” controversy to propose a Bayesian analysis method for fMRI data. Lindquist, who was pursuing a master’s degree in engineering physics, produced a thesis in statistics at Rutgers University that pertained to fMRI. At the time of this controversy, he was an Associate Professor in the Department of Statistics at Columbia University. Gelman, his coauthor, was a statistics graduate of Harvard University and was a professor in the same department as Lindquist as well as a member of the political science department.

In their contribution, which [Gelman \(2009\)](#) outlined and published on his website on 29 January 2009, these two statisticians did not fuel the controversy in terms of its confrontational nature but rather used this sequence of events as an opportunity to address statistical problems specific to neuroimaging. Notably, they spoke of their satisfaction at seeing “such a stimulating discussion” ([Lindquist and Gelman 2009](#), p. 310) of this subject following the publication of the “Voodoo” article. They, thus, initiated the transformation of this controversy into a methodological debate. The structure of their four-page contribution testified to this aim in that they virtually debated with the other contributions in the relevant issue of the journal *PoPS* by addressing the different arguments proposed by the texts that were published concomitantly with their contribution point by point. They did, however, engage in dialog with the contributions of Lieberman, Nichols and Poline, and Vul.

After summarizing the key elements discussed in the “spirited debate” ([Lindquist and Gelman 2009](#), p. 310) initiated by the “Voodoo” article and fueled by the various related contributions, Lindquist and Gelman offered their thoughts and statistical analyses. The two statisticians both agreed with Vul and colleagues’ conclusions regarding the nonindependence error as well as with the relevance of the response by Lieberman and colleagues. By using the appropriate corrections for multiple comparisons, as suggested by Nichols and Poline, the resulting correlation promised to be “radically inflated” (*ibid.*, p. 311). The multiple comparisons method, however, remained insufficient according to Lindquist and Gelman. Similar to Lazar, they deconstructed the idea that this method would solve the problems discussed in the “Voodoo” article.

The authors also addressed the issue of statistical power, noting, as in the case of Yarkoni, that published studies lacked statistical power so frequently that the use of only the magnitude of indicated correlations seemed suspicious. [Lindquist and Gelman \(2009\)](#) added the issue of estimating the standard error of measurement to Yarkoni’s comment, which was focused on statistical power.

Lindquist and Gelman showed that there was a mechanical effect between the standard error and the effect size ensuing from the sample size: the larger the standard error was, the larger the effect size, and conversely, the smaller the standard error was, the smaller the effect size, regardless of the actual size of the effect in question. In short, for these authors, significant results did not necessarily provide strong evidence for the claims of researchers. Lindquist and Gelman, therefore, put into question the contribution of hypothesis testing based on a probabilistic approach; this problem of significance would be highlighted two years later by Joseph P. Simmons and colleagues (2011) to the point that researchers expressly called for the abandonment of the term “statistically significant” ([Wasserstein et al. 2019](#)). As such, the notions of evidence, reliability, and validity were interrogated in the course of the methodological debate initiated by the “Voodoo” controversy.

To overcome the problem of overestimated correlations and standard error of measurement, [Lindquist and Gelman \(2009\)](#) proposed another way forward: the analysis could be based on a relevant model constructed with reference to all relevant studies published on the subject. In short, they invited researchers to start from a better basis rather than to make subsequent corrections (corrections for multiple comparisons). The authors proposed a multilevel Bayesian approach. This method would then allow for the conduction of comparisons that would more likely be statistically valid, thus limiting false positives without reducing the researcher’s ability to “detect true differences, as is often the case in the multiple comparisons framework” (*ibid.*, p. 312).

### *3.4. From the “Voodoo” Controversy to Corrections for Multiple Comparisons*

The “Voodoo” controversy, thus, featured three aspects. The contributions fueling this controversy were initially straightforward and confrontational, beginning in January 2009; namely, they were provided by the neuroscientists who were challenged directly by Vul and colleagues. Other contributions were dispassionate and focused mainly on the statistical and methodological issues highlighted in the “Voodoo” article. A final type of contribution completed the picture, namely, contributions by actors focusing on technique

and methods, i.e., trained and professional statisticians who approached these exchanges with Vul and colleagues as a methodological debate and not as a controversy. However, apart from the opposition of the challenged authors, the contributions agreed that the problem raised by the “Voodoo” article was certainly relevant. Nonetheless, they disputed the explanation proposed in the “Voodoo” article. Some of these contributions, therefore, substituted this hypothesis of nonindependence with that of the importance of applying corrections for multiple comparisons, while others tried to propose other statistical and methodological approaches<sup>9</sup>. The controversy was therefore short-lived<sup>10</sup> and quickly turned into a methodological debate, which in turn fueled a methodological crisis that went beyond the boundaries of the field of social neuroscience, from which it emerged.

The technical elements on which the contributions to the methodological controversy initiated by the “Voodoo” article are based are as diverse as the authors who participated in the controversy. The angles of attack employed by such authors were numerous, including the selection of voxels or clusters, the sample size of the subjects used, the choice of statistical threshold, the statistical power, the measurement error and the correction of noise. For Yarkoni (2009), statistical power is a central dimension of the practices used in psychology, whereas when one moves away from these contributions in attempt to put them into the proper perspective, this evidence seems to play less of a role. Indeed, for Howell (2010), the calculation of statistical power on dependent samples<sup>11</sup> for Student’s *t* test is “often impractical”. For Friston and colleagues (Friston et al. 2006), with respect to the three dimensions that constitute the power calculation, the number of scans performed should be added because the power is slightly different in fMRI studies. Friston explained that 100 scans of 20 subjects are statistically more powerful than 400 scans of five subjects.

However, contributions to the controversy did not focus on these dimensions but rather on the practice of multiple comparisons. While Vul and colleagues indicated that they were not addressing the issue of multiple comparisons but rather emphasizing the selection of regions for analysis, the methodological debate has largely focused on this practice. Indeed, despite the diversity of the relevant contributions, the majority of authors agreed concerning the issue of multiple comparisons and the necessity of applying such correction. Corrections for multiple comparisons appear to some contributors (Nichols and Poline, Vul and colleagues, Poldrack and Mumford, Barrett) to be an obvious point and to constitute a scientific consensus, but other authors view them as not very effective (Lazar) or as insufficient (Lindquist and Gelman). Therefore, although statistics could be imagined to be a univocal scientific field composed of objective, quantifiable elements, this debate demonstrated that it is also permeated by opinions and interpretations. These same beliefs led a majority of neuroscientists to use corrections for multiple comparisons, praised by some, while they encouraged a minority not to take this approach. Thus, the debate did not have a significant effect on the adoption of this correction within the neuroscience community. Indeed, Eklund et al. (2016) claimed in 2016 that a bug in the software package used by neuroscientists and the misuse of parametric statistics cause erroneous results. Six years earlier, Bennet and colleagues (2010) showed that 40% of the works concerning fMRI compiled by these authors did not report having used corrections for multiple comparisons when discussing their methodology.

#### 4. Conclusions: When Controversy Becomes a Methodological Crisis

The strong feelings of outrage generated by the comments of Vul and colleagues soon gave way to a lively methodological debate concerning fMRI data analysis practices. The “Voodoo” article had an earthshaking effect (Poldrack et al. 2011), generating panic in the scientific community (Matusall 2012), which affected neuroscientists so strongly that it generated shockwaves in related areas beyond the issue of overstated correlations that was initially emphasized by Vul and colleagues.

In parallel with this initial emotional reaction, this study shows that three groups fought through different arguments (false positives, statistical power, sample size, etc.), reaffirming the current scientific norms that separate the true from the false. Replication,

forming this boundary, takes the place of the most persuasive argument because replicated studies obtain the highest epistemic value. Indeed, reproducibility is “one of science’s defining features” (Open Science Collaboration 2015) and is “often cited as hallmarks of good science” (National Academies of Sciences, Engineering, and Medicine 2019).

Finally, a technical detail presented in the “Voodoo” article led researchers to question their practices more broadly and generated questions concerning the very foundations of their methods. These reflections echoed in other fields that perceived the same methodological problems. The relevant articles explored the issue of statistical power in further detail, beginning with Yarkoni. They also explored the issues of false positives (Bennett et al. 2010; Eklund et al. 2016), study reliability (Bennett and Miller 2010), and the difficult balance between Type I and Type II errors (Lieberman and Cunningham 2009).

The impact of the “Voodoo” article ranged beyond the confines of social neuroscience to generate a methodological debate within neuroscience as a whole, which was particularly enhanced by the meta-analysis by Kriegeskorte and colleagues (2009) as well as by the article by Fiedler (2011), which generalized the findings of Vul and colleagues, who were convinced that the phenomenon they had observed was likely to affect other areas of research. The shockwave of this controversy continued to spread through psychology (Bird 2021; Yarkoni 2022), joining the problem of reproducibility (Bakker et al. 2012; Colling and Szucs 2021; Flake et al. 2022) and false positives (Simmons et al. 2011) common to other areas of research. In fact, many studies are not confirmed by replication. Indeed, 70% of researchers report failing to reproduce their results (M. Baker 2016). The reproducibility crisis in science (Begley and Ioannidis 2015) is, therefore, spreading. The vast majority (90%) of researchers questioned admit its existence, and half of them (52%) consider this crisis to be “significant” (M. Baker 2016). The debate continues in neuroscience (Eklund et al. 2016; Szucs and Ioannidis 2017, 2020; Yokum et al. 2021), begins in deep learning (Laine et al. 2021), and raises the question of the validity of the research. Indeed, how can the sciences demonstrate and prove their statements if they are not able to confirm their results by reproducing them?

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The author declares no conflict of interest.

## Notes

- <sup>1</sup> The search for articles that contributed to the controversy was conducted across various sources, namely, the citation network of the “Voodoo” article extracted from Web of Science, blog posts, print articles, and magazine articles citing the “Voodoo” article, the diachrony of the scientific and media dissemination of the “Voodoo” article, and finally, by studying the subsequent publications of the relevant authors.
- <sup>2</sup> Sociologist Dominique Raynaud (2018) held that “scientific controversy” is synonymous with a disagreement concerning the nature of the problem, whereas “technological controversy” is synonymous with a disagreement on how to solve the problem. As soon as the dimension of conflict disappears, the use of the term controversy becomes inappropriate. Consequently, this distinction was endorsed in the present study, but the term “technological controversy” was replaced by “methodological controversy”, which seems to more accurately reflect the context being studied. In short, as soon as the actors involved no longer disagree concerning the causes of or the way of solving the problem under consideration, the terms “methodological debate” will be used and no longer “methodological controversy”.
- <sup>3</sup> These correlations are obtained by means of parametric statistical tests (Pearson’s  $\rho$  noted  $r$ ) or nonparametric tests (Spearman’s  $\rho$  noted  $r_s$ ). The aim of these approaches is to determine whether two variables (A and B) are statistically related. The closer the correlation coefficient is to the extremes ( $-1$  or  $1$ ), the more closely these variables are linked or the stronger that link is. Conversely, the closer the coefficient is to zero, the less closely related the variables are or the weaker the relationship is. Researchers then use this correlation coefficient to determine whether or not variable A is likely to explain variable B. To calculate this explained variance, the correlation coefficient is multiplied by itself and is then noted as “ $r^2$ ”. For example, the coefficient



$r = 0.506$  gives an  $r^2$  of 0.256, that is, an explained variance of 26%: twenty-six percent of the variability of variable A is explained by variable B, while the rest is attributable to other variables. For more details, see [Howell \(2010\)](#).

- 4 This correction would increase the risk of false negatives (Type II errors) in order to limit false positives (Type I errors). The chances of finding a real effect would, therefore, decrease. See [Friston \(2006\)](#).
- 5 The term “replicability” indicate that the study contains all the details necessary for potential replication. “Replication” then describes a study that was effectively reproduced either identically or with an adaptation of the original method.
- 6 Indeed, Vul and colleagues did not clearly note the terms that were used in their literature search. The “Voodoo” article only mentioned a few examples: “social terms (e.g., *jealousy, altruism, personality, grief*)” ([Vul et al. 2009a](#), p. 276).
- 7 The article by Kriegeskorte and colleagues (2009) appeared in May 2009 in the journal *Nature Neuroscience* and had a particularly substantial impact on the methodological debate initiated by the “Voodoo” article at the bibliometric level without directly participating in it.
- 8 The power (P) of a test corresponds to the probability of committing a type II error (false negative) (beta  $\beta$ ) and is calculated according to the formula  $P = 1 - \beta$ , whereas the significance level (alpha  $\alpha$ ) represents the probability of committing a type I error (false positive). In hypothesis testing, the researcher is faced with a difficult choice with respect to limiting Type I errors ( $\alpha$ ) without unduly increasing Type II errors ( $\beta$ ). This choice can be quantified by calculating the ratio of alpha to beta ([Cohen 1988](#)). When  $P = 0.8$ ,  $\beta = 1 - P$ , i.e., 0.2 and  $\alpha = 0.05$ , the ratio  $\beta/\alpha = 0.2/0.05 = 4$ . The researcher then notes that generating false positives is four times worse than generating false negatives. This statistical power varies depending on three elements: the significance level  $\alpha$ , the sample size, and the effect size (d) ([Bakker et al. 2012](#)).
- 9 For a condensed summary of these arguments, see [Kriegeskorte et al. \(2010\)](#). This article appeared on 23 June 2010, slightly more than a year after the controversy began, and was presented in a question-and-answer format. However, the article did not have a substantial impact.
- 10 For [Margulies \(2012\)](#), after disrupting the neuroscience community in January 2009, the controversy ended entirely the following month.
- 11 The fMRI data are dependent when several scans are performed on the same subject at different times (time series). All these scans are then paired measurements.

## References

- Abbott, Alison. 2009. Brain Imaging Studies under Fire. *Nature* 457: 245. [[CrossRef](#)] [[PubMed](#)]
- About the Journal. n.d. Social Cognitive and Affective Neuroscience. Available online: <https://academic.oup.com/scan/pages/About> (accessed on 26 July 2020).
- Adolphs, Ralph. 2010. Conceptual Challenges and Directions for Social Neuroscience. *Neuron* 65: 752–67. [[CrossRef](#)] [[PubMed](#)]
- Aims & Scope. n.d. Social Neuroscience. Available online: <https://www.tandfonline.com/action/journalInformation?show=aimsScope&journalCode=psns20> (accessed on 24 May 2020).
- Ariely, Dan, and Gregory S. Berns. 2010. Neuromarketing: The Hope and Hype of Neuroimaging in Business. *Nature Reviews Neuroscience* 11: 284–92. [[CrossRef](#)] [[PubMed](#)]
- Baker, Monya. 2016. Is There a Reproducibility Crisis? *Nature* 533: 452–54. [[CrossRef](#)]
- Baker, Chris I., Tyler L. Hutchison, and Nancy Kanwisher. 2007. Does the Fusiform Face Area Contain Subregions Highly Selective for Nonfaces? *Nature Neuroscience* 10: 3–4. [[CrossRef](#)]
- Bakker, Marjan, Annette Van Dijk, and Jelte M. Wicherts. 2012. The Rules of the Game Called Psychological Science. *Perspectives on Psychological Science* 7: 543–54. [[CrossRef](#)]
- Bardin, Jon. 2009. That Voodoo That Scientists Do. *Seed Magazine*. Available online: [http://seedmagazine.com/news/2009/02/that\\_voodoo\\_that\\_scientists\\_do.php?utm\\_source=seedmag-main&utm\\_medium=rss](http://seedmagazine.com/news/2009/02/that_voodoo_that_scientists_do.php?utm_source=seedmag-main&utm_medium=rss) (accessed on 22 April 2020).
- Barrett, Lisa Feldman. 2009. Understanding the Mind by Measuring the Brain: Lessons From Measuring Behavior (Commentary on Vul et al., 2009). *Perspectives on Psychological Science* 4: 314–18. [[CrossRef](#)]
- Begley, Sharon. 2009a. Sharon Begley: Of Voodoo and the Brain. *Newsweek*. Available online: <http://www.newsweek.com/sharon-begley-voodoo-and-brain-77717> (accessed on 20 April 2020).
- Begley, Sharon. 2009b. The “Voodoo” Science of Brain Imaging. *Newsweek*. Available online: <https://www.newsweek.com/voodoo-science-brain-imaging-221796> (accessed on 22 April 2020).
- Begley, C. Glenn, and John P. A. Ioannidis. 2015. Reproducibility in Science: Improving the Standard for Basic and Preclinical Research. *Circulation Research* 116: 116–26. [[CrossRef](#)]
- Bennett, Craig M., and Michael B. Miller. 2010. How Reliable Are the Results from Functional Magnetic Resonance Imaging? *Annals of the New York Academy of Sciences* 1191: 133–55. [[CrossRef](#)]
- Bennett, Craig M., Michael B. Miller, and George L. Wolford. 2010. Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument For Proper Multiple Comparisons Correction. *Journal of Serendipitous and Unexpected Results* 1: 1–5. [[CrossRef](#)]
- Bird, Alexander. 2021. Understanding the Replication Crisis as a Base Rate Fallacy. *The British Journal for the Philosophy of Science* 72: 965–93. [[CrossRef](#)]

- Borck, Cornelius. 2013. Comment faire du vaudou avec l'imagerie cérébrale fonctionnelle? *Revue D'anthropologie des Connaissances* 7: 571–87. [CrossRef]
- Bruder, Johannes. 2019. *Cognitive Code: Post-Anthropocentric Intelligence and the Infrastructural Brain*. Montréal: McGill-Queen's University Press. ISBN 978-0-7735-5970-7.
- Cacioppo, John T., and Gary G. Berntson. 1992. Social Psychological Contributions to the Decade of the Brain. Doctrine of Multilevel Analysis. *American Psychologist* 47: 1019–28. [CrossRef] [PubMed]
- Cacioppo, John T., and Gary G. Berntson. 2002. Social Neuroscience. In *Foundations in Social Neuroscience*. Edited by John T. Cacioppo, Gary G. Berntson, Ralph Adolphs, C. Sue Carter, Martha K. McClintock, Michael J. Meaney, Daniel L. Schacter, Esther M. Sternberg, Steve Suomi and Shelley E. Taylor. Cambridge: The MIT Press, pp. 3–10. ISBN 978-0-262-53195-5.
- Cacioppo, John T., Gary G. Berntson, Ralph Adolphs, C. Sue Carter, Martha K. McClintock, Michael J. Meaney, Daniel L. Schacter, Esther M. Sternberg, Steve Suomi, and Shelley E. Taylor, eds. 2002. *Foundations in Social Neuroscience*. Cambridge: The MIT Press. ISBN 978-0-262-53195-5.
- Cacioppo, John T., Penny S. Visser, and Cynthia L. Pickett, eds. 2006. *Social Neuroscience: People Thinking about Thinking People*. Cambridge: The MIT Press. ISBN 978-0-262-03335-0.
- Chamak, Brigitte, and Baptiste Moutaud, eds. 2014. *Neurosciences et société: Enjeux des savoirs et pratiques sur le cerveau*. Paris: Armand Colin. ISBN 978-2-200-28764-1.
- Churchland, Patricia Smith. 1989. *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. Cambridge: The MIT Press. ISBN 978-0-262-53085-9.
- Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale: Lawrence Erlbaum Associates. ISBN 978-1-134-74277-6.
- Colling, Lincoln J., and Denes Szucs. 2021. Statistical Inference and the Replication Crisis. *Review of Philosophy and Psychology* 12: 121–47. [CrossRef]
- Diener, Ed. 2009. Editor's Introduction to Vul et al. (2009) and Comments. *Perspectives on Psychological Science* 4: 272–73. [CrossRef] [PubMed]
- Dodier, Nicolas. 1996. Les sciences sociales face à la raison statistique (note critique). *Annales* 51: 409–28. [CrossRef]
- Eisenberger, Naomi I., Matthew D. Lieberman, and Kipling D. Williams. 2003. Does Rejection Hurt? An fMRI Study of Social Exclusion. *Science* 302: 290–92. [CrossRef]
- Eklund, Anders, Thomas E. Nichols, and Hans Knutsson. 2016. Cluster Failure: Why fMRI Inferences for Spatial Extent Have Inflated False-Positive Rates. *Proceedings of the National Academy of Sciences* 113: 7900–5. [CrossRef]
- Fiedler, Klaus. 2011. Voodoo Correlations Are Everywhere—Not Only in Neuroscience. *Perspectives on Psychological Science* 6: 163–71. [CrossRef]
- Flake, Jessica Kay, Ian J. Davidson, Octavia Wong, and Jolynn Pek. 2022. Construct Validity and the Validity of Replication Studies: A Systematic Review. *American Psychologist* 77: 576–88. [CrossRef]
- Friston, Karl J. 2006. Topological Inference. In *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Edited by William D. Penny, Karl J. Friston, John T. Ashburner, Stefan J. Kiebel and Thomas E. Nichols. London: Academic Press, pp. 237–45. ISBN 978-0-12-372560-8.
- Friston, Karl J., John T. Ashburner, Stefan J. Kiebel, Thomas E. Nichols, and William D. Penny, eds. 2006. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. London: Academic Press. ISBN 978-0-12-372560-8.
- Gelman, Andrew. 2009. More on the Voodoo Correlations in Neuroscience. In *Statistical Modeling, Causal Inference, and Social Science*. Available online: [https://statmodeling.stat.columbia.edu/2009/01/29/more\\_on\\_the\\_so/](https://statmodeling.stat.columbia.edu/2009/01/29/more_on_the_so/) (accessed on 22 April 2020).
- Genovese, Christopher R., Nicole A. Lazar, and Thomas Nichols. 2002. Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate. *NeuroImage* 15: 870–78. [CrossRef] [PubMed]
- Giles, Jim. 2009. Doubts Raised over “hot” Neuroscience Results. *New Scientist* 201: 11. [CrossRef]
- Greely, Henry T. 2009. Law and Neuroscience Is More than fMRI. *Stanford Law School*. Available online: <https://law.stanford.edu/2009/02/21/law-and-neuroscience-is-more-than-fmri/> (accessed on 22 April 2020).
- Howell, David C. 2010. *Statistical Methods for Psychology*, 7th ed. Belmont: Wadsworth. ISBN 978-0-495-59784-1.
- Jabbi, Mbemba, Christian Keysers, Tania Singer, and Klaas Enno Stephan. 2009. Rebuttal of “Voodoo Correlations in Social Neuroscience” by Vul et al.—Summary Information for the Press. Available online: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.527.1670&rep=rep1&type=pdf> (accessed on 22 April 2020).
- Kriegeskorte, Nikolaus, W. Kyle Simmons, Patrick S. F. Bellgowan, and Chris I. Baker. 2009. Circular Analysis in Systems Neuroscience: The Dangers of Double Dipping. *Nature Neuroscience* 12: 535–40. [CrossRef] [PubMed]
- Kriegeskorte, Nikolaus, Martin A. Lindquist, Thomas E. Nichols, Russell A. Poldrack, and Edward Vul. 2010. Everything You Never Wanted to Know about Circular Analysis, but Were Afraid to Ask. *Journal of Cerebral Blood Flow & Metabolism* 30: 1551–57. [CrossRef]
- Laine, Romain F., Ignacio Arganda-Carreras, Ricardo Henriques, and Guillaume Jacquemet. 2021. Avoiding a Replication Crisis in Deep-Learning-Based Bioimage Analysis. *Nature Methods* 18: 1136–44. [CrossRef]
- Lazar, Nicole A. 2009. Discussion of “Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition” by Vul et al. (2009). *Perspectives on Psychological Science* 4: 308–9. [CrossRef]
- Lehrer, Jonah. 2009a. In Defense of the Value of Social Neuroscience. *Scientific American*. Available online: <https://www.scientificamerican.com/article/defense-social-neuroscience/> (accessed on 22 April 2020).

- Lehrer, Jonah. 2009b. Voodoo Correlations: Have the Results of Some Brain Scanning Experiments Been Overstated? *Scientific American*. Available online: <https://www.scientificamerican.com/article/brain-scan-results-overstated/> (accessed on 21 April 2020).
- Lieberman, Matthew D. 2006. Social Cognitive and Affective Neuroscience: When Opposites Attract. *Social Cognitive and Affective Neuroscience* 1: 1–2. [CrossRef]
- Lieberman, Matthew D. 2007. Social Cognitive Neuroscience: A Review of Core Processes. *The Annual Review of Psychology* 58: 259–89. [CrossRef]
- Lieberman, Matthew D., and William A. Cunningham. 2009. Type I and Type II Error Concerns in fMRI Research: Re-Balancing the Scale. *Social Cognitive and Affective Neuroscience* 4: 423–28. [CrossRef]
- Lieberman, Matthew D., Elliot T. Berkman, and Tor D. Wager. 2009a. Correlations in Social Neuroscience Aren't Voodoo: Commentary on Vul et al. (2009). *Perspectives on Psychological Science* 4: 299–307. [CrossRef]
- Lieberman, Matthew D., Elliot T. Berkman, and Tor D. Wager. 2009b. Invited Reply Submitted to Perspectives on Psychological Science. Available online: [https://www.scn.ucla.edu/pdf/LiebermanBerkmanWager\(invitedreply\).pdf](https://www.scn.ucla.edu/pdf/LiebermanBerkmanWager(invitedreply).pdf) (accessed on 22 April 2020).
- Lindquist, Martin A., and Andrew Gelman. 2009. Correlations and Multiple Comparisons in Functional Imaging: A Statistical Perspective (Commentary on Vul et al., 2009). *Perspectives on Psychological Science* 4: 310–13. [CrossRef] [PubMed]
- Littlefield, Melissa M., and Jenell Johnson, eds. 2012a. *The Neuroscientific Turn: Transdisciplinarity in the Age of the Brain*. Ann Arbor: The University of Michigan Press. ISBN 978-0-472-11826-7.
- Littlefield, Melissa M., and Jenell Johnson. 2012b. Theorizing the Neuroscientific Turn—Critical Perspectives on a Translational Discipline. In *The Neuroscientific Turn: Transdisciplinarity in the Age of the Brain*. Ann Arbor: The University of Michigan Press, pp. 1–25. ISBN 978-0-472-11826-7.
- Margulies, Daniel S. 2012. The Salmon of Doubt. Six Months of Methodological Controversy within Social Neuroscience. In *Critical Neuroscience. A Handbook of the Social and Cultural Contexts of Neuroscience*. Edited by Suparna Choudhury and Jan Slaby. Oxford: Wiley-Blackwell, pp. 273–85. ISBN 978-1-119-23789-1.
- Marmion, Jean-François. 2009. Imagerie Cérébrale ou Culte Vaudou? *Sciences Humaines*. Available online: [https://www.scienceshumaines.com/imagerie-cerebrale-ou-culte-vaudou\\_fr\\_23511.html](https://www.scienceshumaines.com/imagerie-cerebrale-ou-culte-vaudou_fr_23511.html) (accessed on 22 April 2020).
- Matusall, Svenja. 2012. Searching for the Social in the Brain: The Emergence of Social Neuroscience. Ph.D. thesis, ETH Zurich, Zurich, Switzerland. [CrossRef]
- National Academies of Sciences, Engineering, and Medicine. 2019. *Reproducibility and Replicability in Science*. Washington, DC: The National Academies Press. ISBN 978-0-309-48616-3.
- Neuroskeptic. 2009. “Voodoo Correlations” in fMRI—Whose Voodoo? *Discover Magazine*. Available online: <https://www.discovermagazine.com/mind/voodoo-correlations-in-fmri-whose-voodoo> (accessed on 22 April 2020).
- Nichols, Thomas, and Satoru Hayasaka. 2003. Controlling the Familywise Error Rate in Functional Neuroimaging: A Comparative Review. *Statistical Methods in Medical Research* 12: 419–46. [CrossRef] [PubMed]
- Nichols, Thomas E., and Jean-Baptiste Poline. 2009. Commentary on Vul et al.'s (2009) “Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition”. *Perspectives on Psychological Science* 4: 291–93. [CrossRef] [PubMed]
- Ochsner, Kevin N., and Matthew D. Lieberman. 2001. The Emergence of Social Cognitive Neuroscience. *American Psychologist* 56: 717–34. [CrossRef] [PubMed]
- Open Science Collaboration. 2015. Estimating the Reproducibility of Psychological Science. *Science* 349: aac4716. [CrossRef]
- Park, Robert L. 2005. *Voodoo Science: The Road from Foolishness to Fraud*. New York: Oxford University Press. ISBN 978-0-19-860443-3.
- Poldrack, Russell A., and Jeanette A. Mumford. 2009. Independence in ROI Analysis: Where Is the Voodoo? *Social Cognitive and Affective Neuroscience* 4: 208–13. [CrossRef]
- Poldrack, Russell A., Jeanette A. Mumford, and Thomas E. Nichols. 2011. *Handbook of Functional MRI Data Analysis*. New York: Cambridge University Press. ISBN 978-1-139-49836-4.
- Raynaud, Dominique. 2018. *Sociologie des controverses scientifiques: De la philosophie des sciences*. Paris: Éditions Matériologiques. ISBN 978-2-37361-158-8.
- Rogers, Amy. 2008. Voodoo Correlations in Social Neuroscience. *Seth's Blog. Personal Science, Self-Experimentation, Scientific Method*. Available online: <https://sethroberts.net/2008/12/28/voodoo-correlations-in-social-neuroscience/> (accessed on 22 April 2020).
- Saini, Angela. 2009. The Brain Police: Judging Murder with an MRI. *Wired*. Available online: <https://www.wired.co.uk/article/guilty> (accessed on 22 April 2020).
- Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science* 22: 1359–66. [CrossRef]
- Szucs, Denes, and John P. A. Ioannidis. 2017. Empirical Assessment of Published Effect Sizes and Power in the Recent Cognitive Neuroscience and Psychology Literature. *PLoS Biology* 15: e2000797. [CrossRef]
- Szucs, Denes, and John P. A. Ioannidis. 2020. Sample Size Evolution in Neuroimaging Research: An Evaluation of Highly-Cited Studies (1990–2012) and of Latest Practices (2017–2018) in High-Impact Journals. *NeuroImage* 221: 117164. [CrossRef]
- The Neurocritic. 2008. Scan Scandal Hits Social Neuroscience. *The Neurocritic. Deconstructing the Most Sensationalistic Recent Findings in Human Brain Imaging, Cognitive Neuroscience, and Psychopharmacology*. Available online: <https://neurocritic.blogspot.com/2008/12/scan-scandal-hits-social-neuroscience.html> (accessed on 22 April 2020).

- The Neurocritic. 2009. Voodoo Correlations in Social Neuroscience. *The Neurocritic. Deconstructing the Most Sensationalistic Recent Findings in Human Brain Imaging, Cognitive Neuroscience, and Psychopharmacology*. Available online: <https://neurocritic.blogspot.com/2009/01/voodoo-correlations-in-social.html> (accessed on 22 April 2020).
- The New Scientist Staff. 2009. A Timely Warning about Voodoo Brain Scans. *New Scientist* 201: 3.
- Vaughanbell. 2008. Voodoo Correlations in Social Brain Studies. *Mind Hacks. Neuroscience and Psychology News and Views*. Available online: <https://mindhacks.com/2008/12/29/voodoo-correlations-in-social-brain-studies/> (accessed on 22 April 2020).
- Vidal, Fernando. 2009. Brainhood, Anthropological Figure of Modernity. *History of the Human Sciences* 22: 5–36. [CrossRef] [PubMed]
- Volk, Stefan, and Tine Köhler. 2012. Brains and Games: Applying Neuroeconomics to Organizational Research. *Organizational Research Methods* 15: 522–52. [CrossRef]
- Vul, Ed. 2020. Available online: <https://www.edvul.com/elementor-13/> (accessed on 21 April 2020).
- Vul, Edward, and Nancy Kanwisher. 2010. Begging the Question: The Nonindependence Error in fMRI Data Analysis. In *Foundational Issues in Human Brain Mapping*. Edited by Stephen José Ed Hanson and Martin Ed Bunzl. Cambridge: The MIT Press, pp. 71–91. ISBN 978-0-262-01402-1.
- Vul, Edward, and Hal Pashler. 2012. Voodoo and Circularity Errors. *NeuroImage* 62: 945–48. [CrossRef] [PubMed]
- Vul, Edward, Christine Harris, Piotr Winkielman, and Harold Pashler. 2009a. Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. *Perspectives on Psychological Science* 4: 274–90. [CrossRef]
- Vul, Edward, Christine Harris, Piotr Winkielman, and Harold Pashler. 2009b. Reply to Comments on “Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition”. *Perspectives on Psychological Science* 4: 319–24. [CrossRef]
- Wasserstein, Ronald L., Allen L. Schirm, and Nicole A. Lazar. 2019. Moving to a World Beyond “ $p < 0.05$ ”. *The American Statistician* 73: 1–19. [CrossRef]
- Yarkoni, Tal. 2009. Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power—Commentary on Vul et al. (2009). *Perspectives on Psychological Science* 4: 294–98. [CrossRef]
- Yarkoni, Tal. 2022. The Generalizability Crisis. *Behavioral and Brain Sciences* 45: e1. [CrossRef]
- Yarkoni, Tal, and Todd S. Braver. 2010. Cognitive Neuroscience Approaches to Individual Differences in Working Memory and Executive Control: Conceptual and Methodological Issues. In *Handbook of Individual Differences in Cognition: Attention, Memory, and Executive Control*. Edited by Aleksandra Gruszka, Gerald Matthews and Błażej Szymura. The Springer Series on Human Exceptionality; New York: Springer, pp. 87–107. ISBN 978-1-4419-1210-7.
- Yokum, Sonja, Cara Bohon, Elliot Berkman, and Eric Stice. 2021. Test-Retest Reliability of Functional MRI Food Receipt, Anticipated Receipt, and Picture Tasks. *The American Journal of Clinical Nutrition* 114: 764–79. [CrossRef]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.