Epistemological Custard Pies from Functional Brain Imaging

Jim Bogen†
University of Pittsburgh

This paper discusses features of an epistemically valuable form of evidence that raise troubles for received and new epistemological treatments of experimental evidence.

1. Introduction. On good days experiments yield evidence investigators can rely on in developing and evaluating theories. But we don't know enough about the virtues that make experimental evidence epistemically valuable. Until recently, most philosophers of science approached this issue under the empiricist assumption that the way to learn about the epistemic value of good evidence is attend almost exclusively to human observers, their sensory experiences, and the language by which they report what they observe. This assumption's shortcomings become increasingly obvious as scientists rely more on computer generated evidence derived from inputs from experimental equipment. Human sensory systems and observation reports play too minor a role in the production of such evidence for empiricist theories to tell us much about its virtues. I will illustrate this from some epistemic peculiarities of functional brain images. Deborah Mayo and Jim Woodward have developed more satisfactory approaches to experimental evidence by attending to reliability and error analysis instead of perceptual experiences and their descriptions. But I will argue that functional imaging (FI) also raises difficulties for them. My purpose is not to raise doubts about FI. Where it conflicts with received epistemological doctrines, I think the thing to doubt is the epistemology.

In functional imaging, things are arranged so that radio or radiation signals are emitted from blood circulating through the brains of subjects

†Send requests for reprints to the author, Center for Philosophy of Science, 817 Cathedral of Learning, University of Pittsburgh, Pittsburgh, PA 15260; e-mail: jbogen + @pitt.edu, or rtjbog@aol.com.

Philosophy of Science, 69 (September 2002) pp. S59–S71. 0031-8248/2002/69supp-0006$10.00
Copyright 2002 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/341768 Published online by Cambridge University Press

at rest or performing mental tasks. The signals are used to calculate the locations of elevated blood or oxygenated hemoglobin levels (I will call these biological indicators, or BI) indicative of cognitively significant levels of neuronal activity (CSNA). Neuroscientists do not observe or base their conclusions on the radio or radiation signals any more than physicians observe or base their conclusions about fractured bones on the radiation X-ray technicians use to image them. What functional images exhibit are estimates of BI calculated from the radio or radiation signals and adjusted according to assumptions about their relations to CSNA.

Twentieth century empiricists construed experimental data as reports of what experimenters perceive, and believed their epistemic value depends upon their truth, or acceptability by other investigators. But the epistemic value of functional images depends upon relations between imaged and actual BI, and between BI and CSNA. Since no one perceives the signals, the BI, or the CSNA, functional images are not constructed from, and do not function as reports of what human observers perceive. Thus epistemological theories based upon the empiricist assumption do not apply to them.

Deborah Mayo suggests that the epistemic value of noisy data from a real world experiment depends upon whether the investigator can specify the conditions (unsatisfied by the actual experiment) whose satisfaction would define an ideal, error free version of the same experiment. To the extent that this is known, they can use error statistics to "find a way (through corrections or whatever) of arriving at a desired [i.e., epistemically satisfactory] data statement that effectively subtracts out (or renders harmless) any violation of the" relevant conditions (Mayo 1996, 136). FI statistics can do pretty much what Mayo suggests with regard to certain kinds of error (e.g., in signal detection). But up to now FI methodology has been unable to deal with other sorts of error well enough to enable investigators to make reliable quantitative estimates of how well the BI and CSNA magnitudes they calculate conform to real BI and CSNA magnitudes.[1] Furthermore, some FI computations can be expected to decrease rather than increase accuracy. For example, functional images are typically computed to exhibit the anatomy of significantly atypical brains. Thus, the statistical manipulations involved in functional image production do not always "effectively subtract . . . out (or render harmless) any violation" of the conditions on an ideal experiment in a way that can be counted on to improve accuracy. Finally, in order to produce functional images, experimentalists must rely on stipulations and assumptions from

1. But methods for dealing with noise have improved since I completed this paper. See for example, Christopher Genovese's papers on fMRI statistics at http://www.stat.cmu.edu/~genovese/papers/fmri/mrmodel.ps.

https://doi.org/10.1086/341768 Published online by Cambridge University Press

biology, physics, and elsewhere, along with practices, conventions, and seat of the pants decisions required to design and use the equipment, and to solve technical problems as they come up. What is known about the relevant physical, psychological, and physiological mechanisms provides only partial constraints on these assumptions, practices, etc. As we will see, there are no independent checks available for quantitatively determining how inaccurate they are, or for improving their accuracy to any quantitatively measurable extent. In view of all of this, although error analysis is essential to the production of FI images, it does not accomplish enough to account for their epistemic value.

Functional images differ from the data whose use to detect phenomena Jim Woodward and I have discussed (Bogen and Woodward 1988; Woodward 2002). Such data (e.g., thermometer or pressure gauge readings, photomicrographs, astrographs, bubble chamber photographs, drawings, seismic or EKG tracings) owe their epistemic value to their having been produced in such a way that the item they are used to study exerts a detectable causal influence on them. To identify, measure, or draw conclusions about the item of interest, investigators must reduce and interpret the data to eliminate or mitigate obscurities due to the effects of extraneous causal influences. Woodward's and my account of this applies to cases in which what makes good evidence good is analogous to the clarity of parts of a radio broadcast that convey information the audience needs to pick up. By that I mean just that the easier it is for investigators to discern the features of the data that are indicative of the item of interest and discriminate them from features due to obscuring factors, the better the data. (That's all I mean; no endorsement of information-theoretic accounts of data interpretation intended!) Because the reduction and interpretation of the radiation and radio signals from which the BI are calculated is part of the production of the functional image, these images are more like graphic representations of interpretations of data than what Woodward and I meant by data. The virtues of an epistemically good functional image are more analogous to those of an accurate account of what could be learned from a radio broadcast than to those of the broadcast itself.

But suppose we ignore this and apply what Woodward says about data in Woodward 2000 to a functional image **e,** used to argue for a claim **c,** about levels and anatomical locations of CSNA. If **c** is used to evaluate a neuro-cognitive hypothesis **h,** then, according to Woodward, the epistemic value of **e** would depend on the reliability of the process which produced it. This in turn would depend upon whether the causal influence of CSNA of interest on the production of **e** was such that a relation of counterfactual dependency obtains between **c** and **e.** Ideally, it should be counterfactually the case that **e** has the features it has iff **c** is true. Less ideally, it should counterfactually be the case that **e** has the relevant features iff **c** is highly

https://doi.org/10.1086/341768 Published online by Cambridge University Press

probable, or more probable than competing claims about CSNA levels and locations. At the very least, **e** should not be as it is *regardless* of whether **c** is true, highly probable, etc. According to this quasi-Woodwardian story,[2] to evaluate the epistemic value of a functional image we must find out whether any such counterfactual dependencies obtain between **e** and **c.** To do this, we must learn enough about the workings of the physical and computational mechanisms which generate **e** to determine what kinds of errors and inaccuracies they produce.

Section 2 of this paper sketches some of the custard pies FI throws at epistemologists. Section 3 illustrates difficulties for Woodward and Mayo by contrasting FI with optical coherence microscopy (OCM). OCM images are also computer generated from inputs—unperceived by human observers—from experimental equipment. But unlike FI, OCM images are congenial to Woodward's and Mayo's accounts. In the conclusion (Section 4) I will make a positive suggestion about what the epistemic value of functional images depends upon.

2. Functional images are brain maps, colored to indicate BI levels estimated at various anatomical regions—reds for high levels, purples for low levels, intermediate colors for intermediate levels. PET scanning typically employs a radioactive oxygen isotope, O^{15}, introduced into the bloodstream as a component of H_2O^{15}, to produce the signals the BI are calculated from. Scanning begins when the isotope has reached the brain. There, decaying O^{15} emits positrons which travel a few millimeters until they encounter and interact with electrons to produce paired gamma rays which travel out of the brain in opposite directions along a straight line (Corbetta 1998). In fMRI, the subject is placed in a magnetic field which is manipulated to induce oxygenated and deoxygenated hemoglobin to emit weak radio waves of measurably different properties (Haxby et al. 1998). Blood levels are known to increase in the neighborhood of elevated neuronal activity. Thus PET assumes more gamma rays will be emitted from decaying O^{15} near regions of elevated CSNA than less active brain regions. The significance for fMRI of the differences between radio signals from oxygenated and deoxygenated blood depends on the fact that more oxygenated hemoglobin circulates to active neurorons than their activity consumes so that the level of oxygenated hemoglobin is elevated just downstream from regions of elevated neuronal activity.

To produce a functional image, what I will call a BI image must be computed from the radiation or radio signals the scanners have detected, and projected onto what I will call an anatomical image. That's a sloppy

2. "Quasi" because I don't know whether or how Woodward (2000) was meant to apply to the likes of functional images.

https://doi.org/10.1086/341768 Published online by Cambridge University Press

way of saying that a mathematical representation computed from the signals must be used to assign BI level colors, pixel by pixel, to pixels in a map of the brain. Here are a few epistemologically remarkable features of this process.

Computed as it is from signals that fluctuate while anatomical structures remain the same, the BI image contains little anatomical information. However, the locations of the posterior and anterior commissures, along with the left, right, top, and bottom extremes of the brain are indicated clearly enough for use in projecting the BI onto an anatomical image. The BI image is oriented on a coordinate system whose axes intersect one another at the anterior commissure. Numerical specifications of the colors to be placed as indications of BI levels along these axes are used to project the BI image onto the anatomical image. The most commonly used anatomical image has been the Talairach-Tournoux atlas, a sequence of detailed maps derived from post mortem photographs of thin slices of a single brain. Some investigators project BI images directly onto the Talairach-Tournoux atlas. Others project each BI image onto MRI images[3] of the brain from which it was derived and project the results onto the atlas. Still others project BI images onto a template produced by computationally averaging MRI images of a number of different brains. The result may then be warped onto the atlas. Because the geography of the standardized atlas differs observably from that of any individual brain, some investigators use estimates of the probability that a given location on the Talairach-Tournoux atlas corresponds to any given structure in an actual brain in reasoning from functional images (cp. Kötter 2001). One reason commonly given for using standardized anatomical images like the Talairach-Tournoux atlas is that claims about functionally significant anatomical features cannot be satisfactorily generalized if they are too closely tied to the geographical idiosyncrasies of one, or a very few brains. Many important landmarks are fissures, furrows, or bulges of neuronal tissue that differ from brain to brain according to the way cortical tissue folds during embryological and later development. Furthermore, even if brains could be flattened to correct for these variations, observably large differences would remain[4] (Caviness Jr. et al. 1996, 566). Another reason is that it is easier to compare BI patterns from different experiments if they are projected onto the same anatomical image (Bavelier et al. 1997, 671).

3. MRI images are computed from radio signals that vary with anatomical features rather than BI.

4. For example Watson et al. identified visual area, V5 by "determining the region of maximal rCBF change" in "a region of the occipital lobe in which cortical activation by visual motion might be expected in each individual". They found that the geographical locations V5, so defined, differed observably from brain to brain (Watson et al. (1993), 83).

https://doi.org/10.1086/341768 Published online by Cambridge University Press

The Talairach-Tournoux atlas is remarkable for its unfaithfulness to the geographies of individual brains. Produced as it was from photographs of sections of the brain of a 60 year old woman, it departs in detail from male brains of all ages and from female brains of many ages. Furthermore, despite functionally significant anatomical differences between the hemispheres, photographs of the left brain were reversed and used to represent the right brain as well! (Collins 1998, 1) This, together with observable geographical differences between individual brains is reason to wonder whether putatively significant details in the functional image are really artifacts of the simplifications and idealizations involved in their production.

Functional image colors are computed pixel by pixel. PET imaging is so noisy that for a typical run, "differences between pixel values [over the threshold level] of two scans of the same subject performing the same sensorimotor control task have a distribution with a standard deviation of 15–16%" (Haxby et al. 1991, 542). To correct for this, BI images taken from different subjects (to save time and reduce the radiation damage a single subject might incur under repeated imaging) are statistically averaged. PET imagers consider six to eight a small number of subjects to average over (Watson et al. 1993, 91). But averaging washes out details of individual scans to bias PET images toward high BI levels, and anatomical structures that are relatively large, or don't vary much from subject to subject (Steinmetz and Seitz 1991, 1158).

Inter-subject averaging is not required for fMRI noise reduction. But intra-subject averaging is required to correct for BI magnitudes that are idiosyncratic to one or a few performances of the same task, and inter-subject averaging is used to correct for the anatomical, psychological and physiological idiosyncrasies of individual subjects.

In assigning colors to functional image pixels, thresholds are set to correct for the influence of baseline and other extraneous neuronal activity, cognitively irrelevant blood circulation effects, etc. A color is assigned to a pixel only when calculated BI levels meet the threshold for that color. In trying to set the thresholds high enough to screen out significantly inflated BI estimates without filtering out enough to defeat the purpose of making the image, investigators must make do without precise estimates of the errors the investigator seeks to avoid.

Correlations between detected radiation or radio signals, BI, and CSNA are imperfectly understood—particularly with regard to their time courses. BI and CSNA magnitudes are known to begin changing at different times and to change at different rates. But the exact magnitudes of the differences are not known. Furthermore, while it takes 40–60 seconds for PET, and 2–6 seconds for fMRI to collect enough signals to produce an image, the CSNA that supports a given cognitive function may last only a fraction of a second (Corbetta 1998, 98; Haxby 1998, 125). The

https://doi.org/10.1086/341768 Published online by Cambridge University Press

color of a pixel must be adjusted to take account of all of this if it is to be indicative not just of the average of signals detected to produce the image, but of the magnitudes of the CSNA of interest. To this end, investigators choose a reference wave form to represent the development of CSNA beginning at the time of the stimulus in response to which the subject is to initiate a mental task. Once pixel colors have been corrected for various artifacts (e.g., head motion during fMRI scanning) they are adjusted against the reference wave form. Sometimes a square waveform is used for mathematical convenience. Sometimes waveforms are adjusted to the results of human and animal experiments involving detections of neuronal activity associated with optical cortex and motor responses to visual stimuli. The adjusted wave forms can be expected to be more realistic than square waveforms, but too little is known about the relevant cognitive and neurophysiological phenomena for reliable quantitative measurements of the improvement (Clare 2000 § 6.1, 6.3.2; Haxby 1998, 129).

3. To recapitulate, ♦the geography of the anatomical image is highly unrealistic. ♦Because it is unclear what level the BI must reach to reliably indicate CSNA, investigators cannot be confident they have avoided mistaking baseline for significant activity levels unless they set pixel thresholds high enough to risk filtering out some BI indicative of CSNA. ♦Although it is plausible that adjusted reference wave forms introduce less error than square wave functions, their error characteristics are not well known. ♦Image averaging and other computations take imaged BI levels and locations farther from, rather than closer to, their levels and locations in any real brain.[5] And for the most part, ♦there are no reliable quantitative methods for calculating the extent of the many inaccuracies produced by these and related factors. Therefore, what is epistemically good about functional images is *not* that they are highly accurate with regard to BI levels or locations of individual brains. They aren't. Nor is it that the net effect of the statistical operations that go into their production is to bring error-ridden, inaccurate BI estimates recognizably closer to what would have resulted from ideal experiments shielded from significant sources of error.

The difficulties FI raises for Woodward and Mayo become obvious if we consider it in comparison to optical coherence microscopy (OCM), another non-invasive imaging technique that relies heavily on physical, biological, and statistical principles, and involves a good deal of of ad hoc engineering and other decision making.[6] OCM was developed to image

5. For more, see Bogen (2001).

6. The following sketch is based on an NSF application for OCM development functioning (1996) and a progress report (1997) by Richard Haskell et al. I am greatly indebted to Haskell for providing this material and answering questions about it.

https://doi.org/10.1086/341768 Published online by Cambridge University Press

cells in embryonic tissue that are located too far beneath the tissue surface to be imaged clearly in vivo with an ordinary light microscope. Its images are computer generated from signals derived from interference fringes. The fringes do not visually resemble the structures they are used to image. Like the signals from which FI images are generated, they are not, and need not be, directly observed in order to produce OCM images. But unlike FI, OCM imaging is susceptible to the kinds of error correction Mayo emphasizes, and the error characteristics of its images can be determined in a way that is congenial to Woodward's counterfactual dependency account of reliability.

Ordinary light microscopes can produce clear enough images to answer some questions about structures near the surfaces of very thin slices of prepared tissue. But when an ordinary light microscope is focussed on structures one or more millimeters beneath the surface of a specimen, the image's resolution is ruined by interference from light scattered back from behind and in front of the focal plain. To see them with an ordinary light microscope, one must cut the specimen into thin enough slices to bring the structures of interest closer to the surface. By contrast, OCM noninvasively delivers well resolved images of structures up to 2 mm below the surface of the living tissue. This is accomplished by arranging a Michelson interferometer so that light scattered back from the focal plane along a fiber optic sample path encounters a reference beam whose length is adjusted as required to produce interference fringes. The fringes are produced by interference which only occurs when the sample and reference paths are nearly the same in length. Thus the equipment can be adjusted so that only light scattered back from structures of interest can have a measurable effect on the fringes. OCM images are computed from measurements of the fringes.

The physical, mathematical, engineering, and other assumptions, as well as the decisions involved in producing the interference fringes and computing OCM images are far more heavily constrained by well established theory than their FI counterparts. Furthermore, in marked contrast to FI, much of what is left open by theory can be checked and adjusted against independently obtained results. For example, the developers of OCM were able to use Monte Carlo simulations to determine how long the sample beam must remain focussed on a given point to provide enough photons for a well resolved image, to correct for the effects of multiple scattering, and to estimate differences in the way light would be scattered e.g., by organelles as opposed to interfaces between intracellular and extracellular surfaces. The conclusions drawn from Monte Carlo simulations could be checked by incorporating them into predictions of images of "tissue phantoms"—artificially produced objects whose sizes, shapes, and optical features could be independently determined. To correct the Monte Carlo

https://doi.org/10.1086/341768 Published online by Cambridge University Press

conclusions, the images they predicted were compared to real images and checked for accuracy against already known features of the phantoms. The reliability of OCM images could also be assessed by comparing them to ordinary light photomicrographs of similar structures at the surface of dead tissue preparations. As these examples illustrate, OCM production can be checked detail by detail, and calibrated against well established theoretical principles and independently ascertainable features of natural and artificial structures. Thus, error statistics can be applied in the way Mayo suggests. And in keeping with Woodward's ideas about reliability, the investigation of OCM image production can deliver a great deal of information about what could reasonably be expected of images produced under various counterfactual conditions. For the most part, no such thing is possible for functional imaging. Here the relations between the quantities of interest and the signals from which they are computed are not well enough understood to check the assumptions, etc. involved in image production against well established theory. FI investigators cannot produce artificial BI or CSNA whose levels are well enough known for use in calibrating functional images. Nor can they independently measure naturally occurring brain activity as required for calibration.[7]

4. How then can the evidential value of a functional image be established? Instead of attempting the impossible, and trying to calculate the error characteristics of functional image production, or trying to determine what ideal, error free or highly accurate images of real brains would look like, I suggest we think of functional images as perfectly accurate, error free representations of the anatomy, and the cognitively significant BI of highly idealized, imaginary brains. Just as the epistemic value of an OCM image depends not just on how well it represents the relevant features of the cells it depicts, but also on whether unexamined cells of the same kind typically have those features, I think the epistemic value of a functional image depends upon whether the idealized brain it portrays is representative of real brains with regard to the anatomical, physiological, and

7. In discussions of an earlier version of this paper, Sandra D. Mitchell asked why fMRI can't provide independent evidence against which to check or calibrate PET images—a reasonable question, since the imaging literature sometimes talks of using fMRI to confirm or disconfirm PET findings. But independence is not an all or nothing feature of evidence or its sources. PET and fMRI differ so much with regard to their signals and signal detection methods, that there might be ways to use one to correct for errors in the other due to just those factors. But fMRI relies on assumptions, etc. too similar to those of PET to allow for epistemically independent checks or corrections of error those factors produce. Both methods use image averaging. Both use the Talairach-Tournoux atlas. They use similar wave forms to model CSNA development. They both employ threshold setting. And so on. Unfortunately these are among the most problematical features of PET imaging.

https://doi.org/10.1086/341768 Published online by Cambridge University Press

psychological factors relevant to the question under investigation. Accordingly, the evidential value of a functional image would depend on what sorts and degrees of resemblance between real brains and the idealized brain the image portrays are required to answer the question under investigation, and whether it is safe to assume those resemblances obtain.

Here is an example of what this would look like. In 1998 Rugg used fMRI results obtained by Brewer and Wagner to argue for the hypothesis that (**h**) an event is more likely to be remembered if the subject attends to "its meaning rather than to more superficial attributes (such as physical appearance)" (Rugg 1998, 1151). Brewer and Wagner made fMRI images of subjects performing tasks involving pictorial or verbal displays that they were not told they would have to remember. Call this experimental condition **C.** Some time later, surprise recognition tests were administered. The functional images for condition **C** indicated that neuronal activity increased in the right parahippocampal cortex and a small region of the right prefrontal cortex when certain pictures were displayed. In the recognition test, subjects tended to remember that they had seen those pictures and not to confuse them with items they had not seen in **C.** Performance on the recognition test was significantly worse for pictures whose display appeared not to have been followed by elevated neuronal activity in the parahippocampal and prefrontal regions (Rugg 1988, 1151; Brewer et al. 1998, 1186). Similar results were obtained for the left parahippocampal and prefrontal areas in connection with verbal displays (Rugg 1998, 1151; Wagner 1998, 1189). In another experiment, different subjects used different criteria to sort images. Here, functional images indicated higher levels of CSNA in the parahippocampal and prefrontal regions when the sorting criteria involved semantic meaning than when images were sorted by such non-semantic features as color (Rugg 1998, 1151; Wagner 1998, 1189–90).

To produce his evidence, Brewer averaged BI images from 6 subjects and assigned colors to an averaged anatomical map to indicate above threshold BI levels. He then "normalized" the results "to the stereotactic space of Talairach and Tournoux" (Brewer et al. 1998, n.13, 1187). I propose we think of Brewer's and Wagner's FI images as accurate, error free measures of CSNA in precisely defined regions of an imaginary, idealized brain. So construed, their epistemic value depends in part upon [a] correspondences between the parahippocampal and prefrontal regions of the idealized brain the images portray and functionally significant anatomical units of real brains. It also depends in part upon [b] correspondences between the BI of the idealized brain and features (e.g., locations, levels, or time courses) of the activities of the brains of Brewer's and Wagner's subjects. There is no general, context independent standard by which to determine what kinds and degrees of similarity with regard to which features

https://doi.org/10.1086/341768 Published online by Cambridge University Press

of the BI or the CSNA are required for testing any given hypothesis. The requirements differ from case to case. What they are in any given case depends upon the content of the hypothesis of interest, the nature of the experiment conducted to test it, and the kinds of inferences by which the experimental results are brought to bear on the hypothesis.

As said above, averaged images projected onto the Talairach-Tournoux atlas cannot be expected to very accurately represent BI or CSNA levels or locations of real brains. And in contrast to OCM, there is no well established theory of detailed dependencies between signal detections, BI levels, and CSNA by which to assess reliability with regard to [a] or [b]. Nor can reliability be evaluated by comparing levels and anatomical locations of BI displayed in the functional image to independently established levels or locations in real brains. Nor can reliability be improved by calibration. Lacking the means (available in the case of OCM) to evaluate the images as quasi-Woodward suggests, or to correct for error as Mayo suggests, what can be done? In Rugg's case, the answer is as follows.

With regard to [a], consider the parahippocampal cortex. Postmortem examinations of humans and dissections of primates establish that this is an anatomically discrete part of the brain located in the medial region of the temporal lobe. With regard to its functional significance, clinical observations and the results of neuro-cognitive psychological studies argue that "bilateral damage to the medial temporal lobe . . . yields . . . a pervasive memory deficit for all new events and facts," that left side lesions impair verbal memory, and that right side lesions impair nonverbal memory (Brewer et al. 1998, 1185). Most of the subjects that neuro-cognitive psychologists study have sustained brain damage that is too extensive or diffuse to enable investigators to establish the precise locations of structures whose function is required for memory—let alone the locations of those structures in relation to a Talairach-Tournoux or averaged brain image. But since the mediotemporal region contains the parahippocampal cortex, the neuro-cognitive evidence does argue that in real brains the mediotemporal region is (or contains structures which are) functionally significant in the sense that memory requires its (their) normal functioning. Animal experiments provide further support for this. Parahippocampal lesions in monkeys cause poor performance in memory tasks analogous to those which humans with mediotemporal lesions have trouble with (Bauer et al. 1993, 552). This argues that real brains resemble the idealized brain Brewer imaged in that their medial temporal lobes are or contain anatomical structures that support memory. There is no doubt that real and idealized anatomical structures differ from one another with regard to their shapes, relative sizes, and locations. These differences cannot be precisely determined. But it is plausible that however they may differ, the parahippocampal cortices of real and idealized brains are, or are located

https://doi.org/10.1086/341768 Published online by Cambridge University Press

in, or are near, anatomical structures whose activity contributes to memory. Similar arguments apply to the prefrontal region.

With regard to [b], if the neuronal activity of the parahippocampal and the relevant prefrontal area in real brains (or the regions which contain them) contributes to memory, does it resemble that of their idealized FI counterpart as needed to argue for **h?** Rough, qualitative evaluations of the images relative to the question Brewer, Rugg, and Wagner were interested in suffice to make it plausible that the required resemblance holds. Their question was whether the brain regions whose normal functioning appears (from the clinical and animal experimental evidence just mentioned) to be necessary for memory are the same as those whose activity above baseline levels in **C** correlated with good performance on recognition tests, and which were active above baseline levels in subjects who attended to the meanings of the displayed words and figures they sorted in **C.** This question can be answered in the affirmative without precise, quantitative measures of the activity, or accurate specifications of the shapes, locations, and relative sizes of the relevant brain regions. Highly accurate quantitative estimates of FI error characteristics are unnecessary. Even though image averaging and projection onto the Talairach-Tournoux atlas obscure the geography of individual brains, there is enough independently obtained anatomical evidence to show that individual brains contain mediotemporal and prefrontal regions which correspond to the parahippocampal cortex and the relevant prefrontal area of the idealized brain the FI images represent. Inaccuracies and imprecisions stemming from unrealistic assumptions, stipulations, etc., used to produce the images cannot be measured quantitatively. But thresholds can be set high enough to ensure that above baseline BI and CSNA levels in the idealized parahippocampal and prefrontal cortices correspond to above baseline levels of BI and CSNA in the mediotemporal and prefrontal areas of real, individual brains. This comes at the price of failure to detect cognitively significant sub-threshold activity. But that inaccuracy is irrelevant to the assessment of Rugg's arguments for **h.**

To summarize, this example illustrates that the epistemic value of a functional image depends upon [a] whether regions of interest in the idealized FI brain correspond, *as required for the investigation of the question of interest,* to functionally significant regions of real brains, and [b] what particular sorts of inaccuracies and imprecisions due to image production techniques would make it impossible in practice to reason cogently from the image to an acceptable answer of the question of interest. The evaluation of a functional image along these lines might involve some applications of error statistics. It might involve answering counterfactual questions about what the image would have or might have shown under other circumstances. But at least as things stand now, an analysis of image pro-

https://doi.org/10.1086/341768 Published online by Cambridge University Press

duction will not establish quantitative counterfactual dependencies, or deliver quantitative measures of error characteristics, or degrees of imprecision and inaccuracy like those available for the evaluation of OCM and some other kinds of computer generated images. But as Rugg's research illustrates, functional images may nevertheless provide epistemically good evidence whose epistemic value can be assessed.

REFERENCES

Bauer, R. M., B. Tobias, and E. Valenstein (1993), "Amnesic Disorders" in K. Heilman and E. Valenstein (eds.), *Clinical Neuropsychology.* New York: Oxford University Press, 215–263.

Bavelier, D. et al. (1997), "Sentence Reading: A Functional MRI Study at 4 Tesla", *Journal of Cognitive Neuroscience,* 9:5.

Bogen, J. (2001) "Functional Imaging—Some Epistemic Hot Spots" in R. Grush et al., *Limits of Science: Philosophical Problems in the Neurosciences,* Pittsburgh: University of Pittsburgh Press, 173–199.

Bogen, J. and J. Woodward (1988), "Saving the Phenomena", *Philosophical Review* 97: 303–357.

Brewer, J. B., Z. Zhao, J. E. Desmond, G. H. Glover, and J. D. E. Gabrieli (1998), "Making Memories: Brain Activity that Predicts How Well Visual Experience Will Be Remembered", *Science* 281: 1185–1187.

Caviness, V. S. Jr. et al. (1996), "MRI-Based Topographic Parcellation of Human Neorcortex: An Anatomically Specified Method with Estimate of Reliability", *Journal of Cognitive Neuroscience* 8:6.

Clare, S. (2000), *Functional MRI Methods and Application.* PhD. Dissertation, htt5p://www.fmrib.ox.ac.uk/~Stuart.

Collins, J. (1998), "Probabilistic Atlases of Cerebral Anatomy", at http://www.sop.inria.fr/epidaure/pe . . . dain/neuro98/abstracts/collins.html.

Corbetta, M. (1998), "Functional Anatomy of Visual Attention in the Human Brain", in R. Parasuraman (ed.), *The Attentive Brain.* Cambridge, Mass.: MIT Press, 95–122.

Haxby, J. V. et al. (1998), "Functional Magnetic Resonance Imaging and the Study of Attention", in R. Parasuraman (ed.), *The Attentive Brain.* Cambridge, Mass.: MIT Press, 123–142.

Mayo, D. (1996), *Error and the Growth of Experimental Knowledge.* Chicago: University of Chicago Press.

Kötter, R. (2001), "Neuroscience Databases—Tools for Exploring Brain Structure-Function Relationships", *Philosophical Transactions of the Royal Society* (B) 356: 1111–1120.

Rugg, M. D. (1998), "Memories Are Made of This", *Science* 281: 1151–1152.

Steinmetz, H. and R. J. Seitz (1991), "Functional Anatomy of Language Processing: Neuroimaging and the Problem of Individual Variability", *Neuropsychologia* 29: 1149–1161.

Wagner, A.D. et al. (1998), "Building Memories: Remembering and Forgetting of Verbal Experiences as Predicted by Brain Activity", *Science* 281: 1188–1191.

Watson, J. D. G, G. A. Meyers, R. S. J. Frackiowski, J. V. Hajnal, R. P. Wood, J. C. Mazziota, S. Ship, and S. Zeki (1993), "Area V5 of the Human Brain: Evidence from a Combined Study Using Positron Emission Tomography and Magnetic Resonance Imaging", *Cerebral Cortex* 3: 79–94.

Woodward, J. (2000), "Data, Phenomena, and Reliability", *Philosophy of Science* 67 (Supplement): S163–S179.

https://doi.org/10.1086/341768 Published online by Cambridge University Press