Brain Images, Babies, and Bathwater:
Critiquing Critiques of Functional Neuroimaging

BY MARTHA J. FARAH

Since the mid-1980s, psychologists and neuroscientists have used brain imaging to test hypotheses about human thought processes and their neural instantiation.[1] In just three decades, functional neuroimaging has been transformed from a crude clinical tool to a widely used research method for understanding the human brain and mind. Such rapidly achieved success is bound to evoke skepticism. A degree of skepticism toward new methods and ideas is both inevitable and useful in any field. It is especially valuable in a science as young as cognitive neuroscience and its even younger siblings, social and affective neuroscience. Healthy skepticism encourages us to check our assumptions, recognize the limitations of our methods, and proceed thoughtfully. Skepticism itself, however, also must be examined.

Functional neuroimaging has attracted a substantial amount of skepticism from inside and outside the fields of psychology and neuroscience. In this article, I review the most commonly voiced criticisms of functional neuroimaging. In the spirit of healthy skepticism, I will critically examine these criticisms themselves. Each contains at least a kernel of truth, although I will argue that in some cases the kernel has been overextended in ways that are inaccurate or misleading. In other cases, the criticisms are valid as presented and deserve the careful attention of imaging researchers. The goals of this article are to distinguish between reasonable and unreasonable criticisms and to identify some general characteristics of the two categories. In this way I hope to encourage the fair and realistic evaluation of functional neuroimaging as a scientific method and to foster an understanding of imaging's limitations without "throwing the baby out with the bathwater."

The criticisms I review here can be grouped into four general categories, although there is a degree of overlap among them and many have been combined in the literature despite their conceptual distinctness. The first broad category concerns the nonidentity of, or the gap between, the neural events being studied and the images that purportedly represent them.

Inferential Distance and the Objects of Imaging

As Adina Roskies has pointed out, functional brain images are easily misunderstood as photographs of brain function.[2] There are, however, numerous types of mismatch in the relationship between published functional brain images and the brain activity they represent. Two of these mismatches form the basis of many of the common criticisms reviewed here.
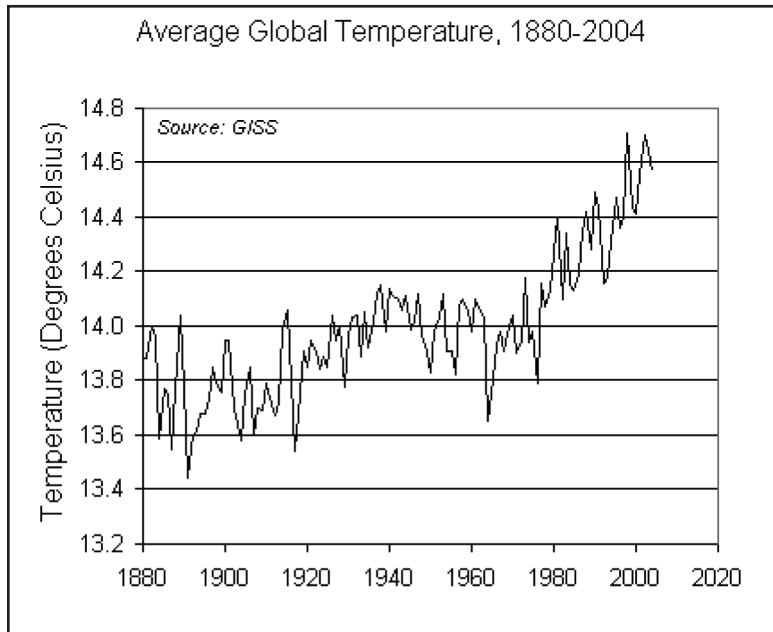
Blood versus brain. The signal measured in functional magnetic resonance imaging, as explained by Geoffrey Aguirre in his essay in this collection, is a characteristic of blood rather than brain tissue. The dependence on hemodynamic proxies for brain activity strikes some commentators as a fundamental flaw in the functional neuroimaging enterprise. I offer three examples of this critique, from philosophers, a science writer, and a psychologist.

Philosophers Valerie Hardcastle and Matthew Stewart question "the excited hoopla over fMRI and other imaging techniques" by pointing out that fMRI informs us about activity only in a relatively large area of brain tissue (on the order of millimeters) and can inform us only at relatively long time intervals (on the order of seconds).[3] They suggest that this poor spatial and temporal

Figure 1.
Average global temperature, 1880-2004

(Source: Goddard Institute for Space Studies)

ably with neural activity. Indeed, despite considerable research on the subject, we do not yet know how fMRI corresponds to specific aspects of neural activity, be they the summed synaptic activity of small populations of neurons, spiking of neurons, or other physiological phenomena.[6]

But let us put these criticisms in perspective. The concern that fMRI shows us blood oxygenation rather than neural activity directly should be weighed alongside the fact that little of what we call science involves direct observations of the subject matter of interest. Cosmologists make inferences about earlier states of the universe by measuring background microwave radiation. Chemists determine the composition of samples by heating or shining light on them and measuring emission or absorption spectra. Climate scientists measure tree rings to study climate trends over previous centuries. Complaints that functional neuroimages do not "show" brain activity appear to be based on a naïve view of science and its methods. Of course, compared to the cosmologists' or chemists' spectroscopy, we have only a weak grasp on the nature of the linkage between fMRI and neural activity. Nevertheless, blood-oxygen-level-dependent (BOLD) fMRI detects a relationship between neural activity and oxygenation that is strong enough to make it a useful research tool.

Nikos Logothetis, an influential voice of caution on the interpretation of fMRI data, has pointed out that similar problems regarding indirectness apply to other neuroscience methods: "Electrical measurements of brain activity, including invasive techniques with single or multiple electrodes, also fall short of affording real answers about network activity. Single-unit recordings and firing rates are better suited to the study of cellular properties than of neuronal assemblies, and field potentials share much of the ambiguity discussed in the context of the fMRI sign." He also affirms that "despite its shortcomings, fMRI is currently the best tool we have for gaining insights into brain function."[7]

In sum, it is not the case that inferences based on functional brain imaging are, in the words of the critics quoted earlier, "fudges" or "stretches." There is indeed a "cause and effect" relationship between neural activity and blood oxygenation levels, even if the nature of that relation is imperfectly understood.

Functional brain images as fabrications. Another way in which the relation between functional brain images and their objects has been questioned concerns the importance of the decisions that researchers must make (described by

resolution creates an imprecise representation of neural activity and write, "This imprecision forecloses the possibility of directly connecting single-cell activity—which operates three to four orders of magnitude smaller and faster—with larger brain activation patterns. What are we to do? The answer given all too often by neuroscientists is to fudge."

In a popular book on the brain, Judith Horstman likewise argues that "an fMRI is not a direct image of your brain showing mental activity. It's an image created through indirectly measuring the flow of oxygenated blood and then correlating that information to something you are doing or thinking at the time, and the blood flow response takes time. So it's a stretch to say that there is cause and effect and to relate this blood surge to a specific activity."[4]

Perhaps the strongest statement of concern over the nonequivalence of blood flow and brain activity can be found in the writings of psychologist William Uttal, a vocal critic of functional neuroimaging research, who recently wrote, "fMRI is as distant as the galvanic skin response or pulse rate from cognitive processes."[5]

The undeniable truth in these criticisms is that blood flow and oxygenation are not the same as brain activity. Blood flow and oxygenation occur at different spatial scales and over different time courses than brain activity. More fundamentally, even if we allow for those spatial and temporal limitations, there is no guarantee that cerebral blood flow or blood oxygenation will correlate precisely or invari-

Aguirre in his essay) regarding what aspects of brain activity to represent and how to represent the selected activity. The worry, in this case, is that the images are more researcher inventions than researcher observations.

For example, scientists and nonscientists alike have regarded the use of color-coding with suspicion. Writer Michael Shermer states that "[c]olors exaggerate the effects in the brain . . . . The coloring is artificial, and the process of coloring the regions is even more misleading."[8] Horstman echoes these criticisms, warning that despite significant differences in the colors used, "differences in activity levels are tiny."[9] In an article entitled "Some Cautions about Jumping on the Brain-Scan Bandwagon," psychologist and textbook author Carole Wade warns that "decisions about color scales . . . can accentuate or minimize the contrasts among different brains or brain areas. Such decisions can affect whether the gorgeous images we see at conferences, in articles and textbooks, and in the popular press will be striking, ho-hum—or even misleading."[10]

The element of truth in this criticism is that color-coding is arbitrary. The choice of which color represents which numerical value is dictated by both convention and convenience. Conventionally, warmer and brighter colors represent higher activation. But is this in some way prejudicial? No more so than plotting numerical data on axes where higher numbers appear higher on the page. For functional images or for the coordinates on a traditional Cartesian graph, one could reverse the scale to remind viewers of the conventional nature of these data representations, but the value of doing so seems small compared to the inconvenience to readers.

A related criticism concerns the calibration of color scales in images. The criticism is that large differences in color can mislead the naïve viewer into thinking that the differences in neural activity are also large, whereas, to use Horstman's word, they are in fact "tiny." But do these efforts at making different levels of brain activation visible make the images, again using Horstman's word, "misleading"? Not at all. Again, the analogy with Cartesian graphs can help. Consider the graph of average global temperatures shown in figure 1. The calibration of the Y-axis spans less than two degrees Celsius because this makes the relevant relationships among data points salient and enables readers to glean an accurate sense of both the variability and the trend of the data points. These data could have been plotted on a graph showing a fuller range of tempera-tures—for example, the range from the lowest to highest naturally occurring temperatures. Doing so would put the data into a different context, which might be useful for some purposes, but for most purposes it would be counterproductive; it would simply make the relationships of interest hard to see. Similarly, the color scale used for functional brain images spans a small range of BOLD signal change in order to maximize the visibility of the distribution of relevant activations.

Could one fabricate results, in the sense of changing the pattern of activation, by changing color scales? It is not possible to boost or minimize the activation of one region by changing scales without boosting or minimizing other equally activated regions. Choices of color scale or contrast affect the salience of certain features, but this is equally true of other biological images, including micrographs, gels, and Western blots.[11] The issues here are not unique to brain imaging, nor do they indicate that colored images are inherently or even typically deceptive.

Metatheoretical Assumptions and Goals of Neuroimaging

Scientific theories and methods inevitably influence one another. In the case of functional brain imaging, it has been suggested that the method seems particularly compatible with certain assumptions about the mind-brain relation and encourages certain types of theories while preventing others from being tested. Specifically, functional neuroimaging has been criticized for encouraging research aimed merely at localizing psychological functions, for being incapable of testing psychological theories, for assuming a modular relation between mental and neural systems, and even for assuming a one-to-one correlation between these systems.

Localization versus explanation. Neuroimaging has been caricatured as a form of phrenology, with the research goal being simply to associate a psychological function with a specific part of the brain. It is easy to see how this misunderstanding could occur, given the archetypal functional brain image bearing color splotches affixed to various parts of a structural MRI.

The philosopher Jerry Fodor has suggested that the question of where in the brain a given psychological activity occurs is scientifically trivial. He writes, "It isn't, after all, seriously in doubt that talking (or riding a bicycle

or building a bridge) depends on things that go on in the brain somewhere or other. If the mind happens in space at all, it happens somewhere north of the neck. What exactly turns on knowing how far north?"[12] Elsewhere in the same essay he explains why our alleged obsession with localization is not just silly but wasteful: "Science is expensive…. If you put your money (which is to say, our money) into the elaborate technology required to establish localizations of mental functions by imaging techniques, you almost certainly take it out of other kinds of psychological research. Likewise in respect of the time and money that is required to train people to do science; graduate students, too, are a limited resource."[13]

Uttal, whose 2001 book-length critique of neuroimaging was titled *The New Phrenology*, suggests that the quest to localize psychological functions in the brain is a primary, and misguided, goal of neuroimaging research. He notes that "a considerable amount of PET [positron emission tomography] and fMRI localization has simply confirmed some things that we have long known,"[14] and he contrasts such research with research that actually tries to discover "how the brain computes, represents, encodes, or instantiates psychological processes."[15]

It is hard to disagree with Fodor and Uttal's observation that localization is, in itself, a questionable scientific goal. However, most contemporary neuroimaging research has other goals. In the early years of PET and fMRI, each method was used to image processes whose brain localizations were already well known on the basis of lesion studies or single-cell recording in animals. This exercise helped to validate the new methods, testing them in order to see whether they produced the expected localizations. Once these localizations were confirmed, confidence in the ability of functional imaging to detect regional brain activity increased. Researchers then began to use the method to investigate cases in which it was not known which brain areas were likely to be recruited for a given psychological function, as well as to address questions that were not primarily questions of localization. Of course, studies of localization for localization's sake are still occasionally published, just as purely descriptive studies with other methods can be found in any field of science, but the bulk of functional neuroimaging research in the twenty-first century is not motivated by localization per se.

Further reason to reject the "neophrenology" charge comes from neuroimaging methods that exploit other features of activation in addition to location. For example, "adaptation paradigms" make use of changes in activity during the course of perception or cognition, specifically the diminution in response to a repeated stimulus or operation when the same set of neurons is reactivated.[16] This is because neurons "adapt" to stimulation, reacting less strongly if they have recently been active. By carefully choosing and arranging the order of stimuli and then measuring the effects of preceding one stimulus with another, it is possible to determine the proportion of neurons that the two stimuli activate in common, in effect, the similarity of neural coding between the stimuli. When one knows the nature of representational similarity, one knows much about the nature of representations.[17] Using this approach, researchers have been able to distinguish between viewpoint-dependent representations of an environmental location (akin to a photograph taken from a particular vantage point) and viewpoint-invariant representations (similar to a map representation of location).[18]

Other decidedly nonphrenological methods of analysis take advantage of neuroimaging's ability to give information about the state of the entire brain, which contrasts with the necessarily piecewise approach of nonimaging methods such as single-cell recording and lesion methods. Analyses of functional connectivity reveal which subset of areas show correlated activity, suggesting that they are working together.[19] Such analyses pick out functional networks of areas, which may change depending on the task conditions, and hence put strong constraints on the nature of the organization of cognitive subsystems.

Localization is also generally beside the point in multivoxel pattern analysis methods, described elsewhere in this report by Aguirre. Indeed, the results of such studies can often be stated without any reference to anatomical localization. These methods enable researchers to detect specific mental states and decode how specific stimuli or events are represented in the brain. None of the foregoing uses of neuroimaging fits the description of "neophrenology."

Relevance to psychological theory. While anatomical location is an intrinsic property of neuroimaging data, the use of these data is not confined to the study of localization any more than the use of reaction time, the primary methodology of cognitive psychology, is confined to the study of cognitive speed. Both localized activations and response latencies are used by researchers to test psychological hypotheses. Some critics acknowledge this use of functional imaging but question whether it has been effective as a means of testing psychological theories.[20] As with the preceding objections, there is validity to this one. Let us consider the sense in which it is valid as well as the sense in which it misses the mark.

In response to examples of how neuroimaging can confirm or disconfirm psychological theories, Max Coltheart has provided alternative explanations of specific findings to show that they are not, in fact, decisive.[21] What gets lost in the debate is the fact that decisive experiments are not generally possible in psychology, and it would be equally difficult to offer examples of traditional cognitive psychology's success by this criterion.[22] The phenomena under study within cognitive psychology and neuroscience are

Can functional brain imaging contribute to confirming psychological hypotheses in roughly the way behavioral studies do?

too complex, and the associated theories have too many degrees of freedom, to hope for decisive single findings.

A fairer and more realistic question is this: Can functional brain imaging contribute to confirming psychological hypotheses in roughly the way behavioral studies do? Can functional imaging, experiment by experiment, rule out the more straightforward alternative hypotheses and leave progressively more complex or strained alternatives in relation to the supported hypothesis? Many fruitful research programs answer this question in the affirmative. Different examples of issues in cognitive psychology that imaging has *helped* to resolve have been offered.[23] My own favorite example concerns the relation between mental imagery and perception, a long-standing issue in cognitive psychology.[24] Does the visual system do "double duty" for perceptual processes and mental images generated from memory, or is the system used to generate mental images from memory independent of the visual system? Measures of localized brain activity as rudimentary as electroencephalogram (EEG) and single-photon emission computed tomography (SPECT),[25] later confirmed with PET and fMRI, ruled out a set of alternative explanations that had plagued the behavioral approach to this question, with the evidence indicating at least some shared representations.[26] Of course, the original question was replaced by new questions, inquiring, for example, more specifically as to which visual representations are shared with imagery, but this was progress. Recently, Mara Mather and colleagues surveyed a variety of research programs and identified four distinct ways in which localized brain activity can be used to test cognitive theories.[27]

Biasing hypothesis generation. It has also been suggested that the use of imaging constrains the kinds of theories of mind-brain relations that will be devised and tested. The concern is not that functional brain imaging is merely an exercise in localization or that it is incapable of testing psychological hypotheses. Rather, the concern is that it directs us to think about brain function in the wrong way. First, it invites us to focus on a subset of the relevant data. The problem, writes Uttal, is "the mistaken idea that when all lesser peaks are reduced to invisibility by arbitrary scaling, the largest remaining peak represents the sole locale of a particular cognitive process." Second, imaging invites us to frame hypotheses "at the wrong (macroscopic) level of analysis rather than the (correct) microscopic level."[28] The result is "misdirected attention and effort": "we are doing what we can do when we cannot do what we should do."[29]

The kernel of truth here is that early approaches to the design and analysis of functional neuroimaging experiments were best suited to studying relatively small numbers of macroscopic regions of activation, ignoring questions of representation within areas or complex interactions among areas. The method of subtraction (again, see Aguirre), which was first used to show how functional brain imaging can illuminate cognitive processes, assumes that a cognitive process A will have the same neural instantiation whether it is accompanied by cognitive process B, C, D, or E.[30] In this way, the subtraction method assumes fixed, context-independent modules. Of course, analytic approaches such as adaptation designs and network modeling, discussed earlier, show how fMRI can transcend this limitation.

As Erik Parens pointed out in the course of the meetings that gave rise to this collection of essays, the history of neuroimaging may be not unlike that of behavioral genetics in this regard. In the 1980s and '90s, encouraged by the discovery of several examples of rare medical diseases that were caused by single genes, some researchers set out to discover single genes responsible for common psychiatric illnesses such as depression, schizophrenia, and autism. In effect, these behavioral geneticists were also assuming a simple and modular theoretical framework, with individual genes responsible for the psychological phenomena of interest. With time and the accumulation of research results showing the inadequacy of this assumption, behavioral genetics moved to more complex models, in which genes exert their effects on behavior through complex networks of interaction with one another and with the environment. It wasn't the case that the gene as a unit of analysis was useless, but that genes had to be seen in still more complex terms.

Finally, the concern that brain imaging puts scientific blinders on researchers, allowing them to see only simple modular systems, is assuaged by the concurrent use of other methods. Hypotheses to be tested are selected in part based on research with other methods of neuroscience and psychology, and results are interpreted in part based on research with these other methods.[31]

Wanton reverse inference. If there were a one-to-one relationship between brain regions and psychological processes, as assumed in strict modular hypotheses of brain function, then it would be easy to infer what psychological processes are under way at any given moment simply by observing which brain regions are active. Given that the brain-mind relation is far more complex than that, with a single psychological process of interest typically engaging multiple regions and a single region typically involved

in the implementation of multiple psychological processes, inferring a psychological process from an observed brain activity is not straightforward. This has not stopped researchers from attempting to make such inferences, ignoring the intrinsic ambiguity of a single region or pattern of activation in a single experimental context, considered in isolation.

This type of inference, going from an observation of brain activation to an inference about the psychological process that caused it, was called "reverse inference" by Russell Poldrack. The name highlights a difference between this research practice and the more common, and less problematic, "forward inference" practice of manipulating psychological processes and observing resulting brain activation.[32]

Although reverse inference has often been criticized, it is not, in itself, invalid. If one has done due diligence to ascertain the range of psychological processes that can activate a region under a given set of circumstances, then appropriately framed reverse inferences will be possible.[33] Unfortunately, it has often been used wantonly, interpreting a pattern of activation without knowing, or acknowledging, the variety of psychological processes that could produce that pattern. A celebrated example of such wanton reverse inference appeared in a 2007 New York Times op-ed written by neuroscientists Marco Iacoboni, Joshua Freedman, and Jonas Kaplan during that year's presidential primaries. They presented their findings from an fMRI study of undecided voters viewing still photos and videos of the leading candidates. On the basis of activity in the anterior cingulate cortex (ACC), known from other studies to become active when subjects feel conflicting response tendencies, the authors concluded that voters felt ambivalence toward Hilary Clinton and were "battling unacknowledged impulses to like" the candidate.[34]

Several cognitive neuroscientists harshly criticized the op-ed, pointing out numerous flaws in the research design and data analysis[35]—and emphasizing the wanton use of reverse inference. The problem with the reverse inference in this case was that many other psychological states also activate the ACC. Activity in this area has been elicited by processes as diverse as attention to one's own heartbeat and emotional regulation.[36]

Many of the commercial applications of functional brain imaging involve reverse inference. For example, neuromarketing relies on the ability to infer liking and wanting from patterns of brain activation.[37] Similarly, the use of functional brain images for diagnosis in some for-profit psychiatric clinics is also based on the premise that certain patterns of activation can be used to infer the presence of certain disorders or the suitability of certain treatments. Brain-based lie detection promises to distinguish true from deceptive responses by determining whether a person's

pattern of brain activation when giving a response more closely matches that previously associated with true or deceptive responses.[38] In principle, and with the right base of evidence, these reverse inferences could lead to valid conclusions along the lines of "there is a 75 percent chance that the subject" either wants the product, would respond to Prozac, or is telling a lie. In most cases, however, the empirical groundwork needed for these claims is far from complete. For example, for all we know now, psychological processes other than lying could produce the "lying pattern" with high probability.[39]

The multivoxel pattern analysis (MVPA) technique (see Aguirre) is in effect a combination of reverse and forward inference approaches. Reverse inferences can be made with confidence only when one knows the full range of psychological processes that could produce a given pattern of activation under the circumstances of the study. In MVPA studies, forward inference is first performed on all of the psychological states that reverse inference will later be used to select among. When the time comes for reverse inference, it is in effect a very large multiple-choice among the psychological states that were the subject of forward inference. More open-ended uses of reverse inference, in situations where one does not know all of the subject's possible psychological states, would require more exhaustive programs of forward inference research to have been completed and might yield an answer taking the form of "state A with probability X, state B with probability Y, and so on."

As with many of the other concerns reviewed here, unwarranted reverse inference is not unique to functional neuroimaging. It will affect any technology in which a specific cause is inferred from the presence of an effect that could have resulted from other causes. For example, in the use of conventional polygraphy for lie detection, responses can be accompanied by autonomic reactions for reasons other than deception.[40] In prostate-specific antigen testing, elevations of the protein prostate-specific antigen can result from causes other than a growing tumor.[41] In sum, the problem with reverse inference is not unique to brain imaging, and reverse inference is not inherently problematic. The problem is with making a reverse inference without the relevant knowledge from forward inference.

Neuroimaging's Slippery Statistics

Virtually all scientific research depends on statistical analysis at some critical juncture or other, but functional brain imaging research is particularly dependent on statistics. As described earlier by Aguirre, extensive signal processing and statistical analysis intervene between the data acquired from the scanner and the published image. A number of criticisms of functional brain imaging have focused on the statistics involved.

When carried out properly, statistical analyses deepen our understanding of the data and the larger reality from which they were sampled.

Statistical inference versus direct observation. Some criticisms concern the sheer amount of statistical analysis involved in producing a functional brain image. The use of statistics often involves substituting estimated values for raw data. When the very earliest stages of image processing consist of replacing measured signal values with estimated values, the result can be viewed as a fabrication, with all the negative connotations of that term discussed earlier.

In addition, the extensive statistical processing involved in neuroimaging offers many opportunities to distort scientific evidence and therefore makes some critics suspicious. For example, neuropsychologist Carlo Umilta is quoted as questioning, "Would coloured images be so convincing even if readers knew those images are the result of an elaborate sequence of 'cleaning actions,' each one characterized by a not insignificant error probability?"[42] More extreme skepticism comes from neuroscientist Steven Rose. Commenting on Fodor's critique of imaging, he wrote that the "images are marvelously seductive, but by the time you see them they have been so massaged as to risk being thoroughly misleading."[43] Michael Shermer observes that "[b]rain images are statistical compilations," and he advises his readers to "keep all [these many stages of data analysis] in mind next time you see one of those colorful brain scans. . . . [Such images are] highly misleading."[44] Even the widely employed and basic statistical process of averaging can seem problematic in the context of imaging: Carole Wade offers that the "vexing problem of individual differences in brain anatomy," which may make the "uniqueness of fingerprints or facial features seem simple by comparison," is problematic "when scans from a number of individuals are averaged to produce a single image."[45]

The problem with this criticism is that, while the incorrect use of statistics is indeed misleading, there is nothing inherently misleading about using statistics. The idea that each statistical operation on data from the scanner is a step away from reality and toward artifice is a misunderstanding. The reality of interest is brain function, not the raw data collected by the scanner, and most of the statistical analyses used are tools to learn more about that reality. When carried out properly, statistical analyses deepen our understanding of the data and the larger reality from which they were sampled. This is the whole point of using statistical methods in any field, from cognitive neuroscience to demography. In other words, statistical methods are used not to mask reality but to *better approximate it* based on

necessarily limited measurements. Like every other aspect of scientific research, statistical analyses can be done well or poorly, with objectivity or bias, but the use of statistics per se is not a problem.

The kernel of truth in the criticisms just discussed is that the extensive use of statistics provides equally extensive opportunities for error. Two such errors are reviewed in the next two sections.

Multiple comparisons. Functional brain imaging is susceptible to a particular kind of statistical problem that can exaggerate the reliability of findings. The problem arises because of the enormous number of statistical tests that can be carried out with image data. A functional MRI may contain 50,000 voxels (see the essay by Aguirre for explanation of "voxel"), and each of those voxels could be the site of an independent statistical test comparing the value of the BOLD response measured in that small bit of brain between the conditions of the experiment. Statistical tests yield a "significance level," which is the probability that the observed difference between two conditions was due to chance variation alone. Given that experiments are designed so that the hypothesis of interest predicts a "real" difference—a difference, that is, that would be expected any time the experiment was run and not just on occasions when it happens by chance—one should employ a fairly low, and thus stringent, significance level. The conventional cut-off for considering a finding reliable or "real" is a probability of less than one in twenty that the difference between conditions is due to chance, usually written as "$p<0.05$." Of course, researchers feel more confident about their conclusions when $p<0.01$ or $p<0.001$.

When significance testing is carried out with brain imaging data, the following problem arises: if we test all 50,000 voxels separately, then by chance alone, 2,500 would be expected to cross the threshold of significance at the $p<0.05$ level, and even if we were to use the more conservative $p<0.001$ level, we would expect 50 to cross the threshold by chance alone. This is known as the problem of multiple comparisons, and there is no simple solution to it. For example, if we were to consider as activated only those voxels where the difference between conditions achieves a significance level of $p<0.00001$, then we would also be setting the bar too high for most real differences to be found, given realistic limits on the power of our experiments to detect effects.

One vivid demonstration of the multiple comparison problem was published by Craig Bennett and colleagues in an article titled "Neural Correlates of Interspecies Perspective Taking in the Post-mortem Atlantic Salmon," in which a dead fish was placed in an MRI scanner and instructed to think about the emotions being experienced by people depicted in photographs.[46] Thanks to the large number of independent tests carried out on the data, and with the help of some entertainingly good luck concerning the location of the spurious differences, the dead salmon's brain was found to have regions engaged in perspective-taking at the p<0.001 level of significance.

Statisticians have developed solutions to the problem of multiple comparisons. These include limiting the so-

Table 1.
Criticisms of functional brain imaging

Criticism	*Intended Scope*	*Evaluation*
Objects of Imaging		
Blood not brain	All fMRI	*Kernel of truth:* Neural activity per se is not imaged; the relation of BOLD to neural activity is not well understood.
Fabrications	All fMRI	*Kernel of truth:* Data representation requires arbitrary design decisions about, for example, scales and colors.
Metatheoretical Assumptions, Goals		
Localization vs. explanation	All fMRI	*Kernel of truth:* Some studies are aimed at localization per se, often for validation of method.
Cannot test psychological hypotheses	All fMRI	*Kernel of truth:* There is no single decisive imaging finding.
Bias for modular hypotheses	All fMRI	*Kernel of truth:* Early approaches to image analysis were best suited to modular theories.
Wanton reverse inference	Particular studies	*Unqualified truth:* Without appropriate (often extensive) research, one cannot infer a psychological process from brain activation. Many studies have drawn invalid conclusions.
Statistics		
Statistics not reality	All fMRI	*Kernel of truth:* "Reality" is represented only through the filter of multiple statistical operations.
Multiple comparisons	Particular studies	*Unqualified truth:* Images lend themselves to large numbers of independent statistical tests, inviting spurious results if the tests are not appropriately constrained or corrected.
Circularity	Particular studies	*Unqualified truth:* Researchers must select regions or subsets of data by means independent of the subsequent test.
Influence		
Overly convincing	All fMRI	*Kernel of truth:* This phenomenon has been observed but does not appear robust; the aura of science more generally may be persuasive.
Overly appealing	All fMRI	*Kernel of truth:* This phenomenon is intuitively plausible, but there is little evidence to support it.

Functional neuroimages have been called "convincing" and "gorgeous." Two prominent worries are that the images are too convincing and too appealing.

called family-wise error rate and false discovery rate, both of which were tried by Bennett and colleagues and which resulted in a finding of no activation in the dead salmon's brain (this part of their article received less attention than the part reporting perspective-taking activity in the dead salmon's brain.) Another legitimate tack around this problem is to use a priori regions of interest—that is, to simply limit the number of comparisons by specifying in advance the regions relevant to the research hypothesis. Still other methods include reduction of the number of independent tests, in light of the dependence among voxels, and permutation analyses.[47] In short, functional neuroimaging research is not doomed to produce spurious results because of the problem of multiple comparisons. Most neuroimaging articles from the last decade avoid the error highlighted by the dead salmon study, but not all do, which is why Bennett and colleagues published their study. Finally, the problem of multiple comparisons is not unique to neuroimaging. For example, epidemiology has its own reductio ad absurdum demonstration of the need to properly manage multiple comparisons, analogous to the dead fish study: a study relating health to astrological signs.[48]

Circularity. A related problem was pointed out by Ed Vul and colleagues in a paper titled "Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition,"[49] which was widely discussed online and in the print media under its original and more pointed title, "Voodoo Correlations in Social Neuroscience."[50] These critics identified a circularity in the way some published findings had been analyzed. Some researchers first identified the voxels most activated by their experimental task and then—with the same data set—carried out analyses only on those voxels to estimate the strength of the effect.

Just as differences due to chance alone inflate the uncorrected significance levels in the dead fish experiment, differences due to chance alone contribute to the choice of voxels selected for the second analysis step. The result is that the second round of analyses is performed on data that have been "enriched" by the addition of chance effects that are consistent with the hypothesis being tested. In their survey of the social neuroscience literature, Vul and colleagues found many articles reporting significant and sizeable correlations with proper analyses, but they also found a large number of articles with circular methods that inflated the correlation values and accompanying significance levels.

As with the problem of multiple comparisons, the problem of circularity is not unique to functional neuroimaging. In the words of Vul and Harold Pashler, "Variants of this problem seem to arise in every field that takes on the considerable challenge of identifying and quantifying signals found in massively multivariate data, where one cannot ascertain in advance where the signals of interest may lie."[51] They cite psychometrics, epidemiology, genetics, and finance as examples of fields in which circular analyses have distorted the results of research.

The Undue Influence of Brain Images

Functional neuroimaging has also been criticized as unduly persuasive or appealing. Although this criticism is aimed at the public's lack of scientific literacy rather than at imaging per se, it figures in many of the criticisms of neuroimaging cited earlier. Recall the earlier references to images as "colorful," "gorgeous," and "convincing." Two worries in particular have been prominent: that images are too convincing and that they are too appealing.

Overly convincing. Matthew Crawford refers to brain imaging as "that fast-acting solvent of critical faculties,"[52] and Steven Poole writes that brain images, "like religious icons, inspire uncritical devotion."[53] Others have expressed concern about the persuasive power of images in applied contexts. One group of researchers, for example, has noted the social harms that could ensue from "the mistaken impression that fMRI, in particular, is an infallible mind-reading technique that can be used to establish guilt or innocence, infer 'true intentions,' detect lies, or establish competency to drive, vote, or consent to marriage."[54] Others warn of "the potential for brain scan images to create biases in the laboratory, the clinic and the courtroom."[55]

Evidence of the outsized persuasive power of brain imaging was obtained by David McCabe and Alan Castel, who assessed the effects of functional brain images on perceptions of the quality of cognitive neuroscience research. Using both fictional research descriptions and a real BBC science news article, they documented higher ratings of credibility when the texts were accompanied by functional brain images compared to bar charts or topographical maps of scalp-recorded EEGs.[56] Although heavily cited in the years since it was published, the findings have not been replicated, and indeed the recent failure to replicate them

at four different laboratories casts doubt on the phenomenon.[57]

Studies of the role of brain images in legal decision-making have also failed to demonstrate special influence, although clinical and neuroscience evidence that gives jurors relevant information does of course have an effect. Mock jurors in one study were more likely to render a verdict of "not guilty by reason of insanity" if defendants had a prior history of psychiatric disorder or neurological damage and the jury was shown brain images.[58] However, the brain images were always accompanied by additional written testimony in this study, so the influential factor may have been the testimony and not the brain images per se. Another study, assessing the effects of various types of lie detection evidence, found that participants rendered more guilty verdicts when fMRI evidence was described than when polygraphs, thermal face imaging, or no lie detection method was described (although the effect disappeared when the lie detection methods were criticized in a cross examination).[59] Note that in this case, brain images were not shown, and there was differing information associated with the lie detection conditions (for example, activation of frontal lobes for fMRI, rise in facial temperature for thermal imaging). The finding speaks to jurors' views of neuroscience evidence, but its relevance to brain images more specifically remains unclear. Two other recent studies of juror decision-making and brain images offer additional evidence,[60] although neither study found an effect of brain images over and above information delivered verbally.

Overly appealing. Even if brain images do not routinely persuade viewers of conclusions they might otherwise question, some have worried that brain images are so attractive or fascinating that they garner more than their fair share of attention and resources and so crowd out other, more worthy science. As Paul Bloom put it, "Psychologists can be heard grousing that the only way to publish in *Science* or *Nature* is with pretty color pictures of the brain. The media, critical funding decisions, precious column inches, tenure posts, science credibility and the popular imagination have all been influenced by fMRI's seductive but deceptive grasp on our attentions."[61]

Little evidence has been brought to bear on this claim, however. In one experiment, brain images led laypersons to rate newspaper-style research descriptions as being more interesting, compared to descriptions accompanied by photographs, but brain images have not been found to perform better than photographs at making the descriptions seem more worthy of funding.[62]

In sum, although brain imaging seems to lack the dangerous allure attributed to it by some critics, neuroscience more generally may command credibility and interest, possibly more than is warranted under some circumstances.[63] This is not unique to neuroscience. For example, the presence of equations also enhances evaluations of scientific work.[64]

The Baby and the Bathwater

Functional brain imaging has been subject to many criticisms in its first three decades as a method of psychology research, summarized in table 1. This is appropriate, given how thoroughly the use of imaging has transformed the field of psychology. As functional brain imaging is taken up by various applied disciplines outside of psychology research—for example, clinical psychiatry diagnosis or lie detection in legal and other contexts—the stakes grow even higher. Inferences based on functional brain imaging, whether for basic science or applications, require scrutiny.

As we apply such scrutiny, it is important to distinguish between specific criticisms of particular applications or specific studies and wholesale criticisms of the entire enterprise of functional neuroimaging. In the first category are criticisms aimed at improving the ways in which imaging experiments are designed and the ways in which their results are interpreted. Uncontrolled multiple comparisons, circular analyses and unconstrained reverse inferences are serious problems that undermine the inferences made from brain imaging data. Although the majority of research is not compromised by any of these errors, a substantial minority of published research is, making such criticisms both valid and useful.

In contrast, the more sweeping criticisms of functional imaging concern the method itself and therefore cast doubt on the conclusions of any research carried out with imaging, no matter how well designed and carefully executed. These more wholesale criticisms invoke the hemodynamic nature of the signal being measured, the association of neuroimaging with modular theories of the mind, the statistical nature of brain images, and the color schemes used to make those images seductively alluring. As mentioned earlier, each of these criticisms contains an element of truth, but overextends that element to mistakenly cast doubt on the validity or utility of functional neuroimaging research as a whole. None of the criticisms reviewed here constitute reasons to reject or even drastically curtail the use of neuroimaging. Rather, they remind us that neuroimaging, like other scientific methods, is subject to various specific errors that the self-correcting process of science continues to address.

References

1. M. J. Posner and M. E. Raichle, *Images of the Mind*, rev. ed. (New York: W.H. Freeman & Company, 1997).
2. A. L. Roskies, "Are Neuroimages Like Photographs of the Brain?" *Philosophy of Science* 74 (2007): 860-72.
3. V. G. Hardcastle and C. M. Stewart, "What Do Brain Data Really Show?" *Philosophy of Science* 69 (2002): 572-85.

4. J. Horstman, *The Scientific American Brave New Brain* (San Francisco CA: Jossey-Bass, 2010), p. 81.

5. W. R. Uttal, *Mind and Brain: A Critical Appraisal of Cognitive Neuroscience* (Cambridge, MA: MIT Press 2011), p. 31.

6. N. K. Logothetis, "What We Can and Cannot Do with fMRI," *Nature* 453 (2008): 869-77; A. D. Ekstrom, "How and When the fMRI BOLD Signal Relates to Underlying Neural Activity: The Danger in Dissociation," *Brain Research Reviews* 62 (2010): 233-44.

7. Logothetis, "What We Can and Cannot Do with fMRI," p. 877.

8. M. Shermer, "Why You Should Be Skeptical of Brain Scans," *Scientific American Mind* 19, no. 5 (2008): 66-71, p. 69.

9. Horstman, *The Scientific American Brave New Brain*, p. 81.

10. C. Wade, "Some Cautions about Jumping on the Brain Imaging Bandwagon," *American Psychological Society Observer* 19, no. 9 (2006): 23-24.

11. M. Rossner and K. M. Yamada, "What's in a Picture? The Temptation of Image Manipulation," *Journal of Cell Biology* 166 (2004):11-15.

12. J. Fodor, "Diary: Why the Brain?" *London Review of Books* 21, no. 19 (1999): 68-69.

13. Ibid.

14. W. R. Uttal, *The New Phrenology: The Limits of Localizing Cognitive Processes in the Brain* (Cambridge, MA: MIT Press, 2001), at p. 208.

15. Ibid, p. 217.

16. See K. Grill-Spector and R. Malach, "fMR-adaptation: A Tool for Studying the Functional Properties of Human Cortical Neurons," *Acta Psychologica* 107 (2001): 293-321.

17. For example, if "bear" and "pear" are more similar than "pear" and "grape," then phonological or orthographic representations are likely in use; if "pear" and "grape" show the greater similarity, then semantic or conceptual representations are likely in use.

18. R. A. Epstein, K. S. Graham, and P. E. Downing, "Viewpoint: Specific Scene Representations in Human Parahippocampal Cortex," *Neuron* 37 (2003): 865-76.

19. See M. D. Greicius et al., "Functional Connectivity in the Resting Brain: A Network Analysis of the Default Mode Hypothesis," *Proceedings of the National Academy of Sciences* 100 (2003): 253-58. See J. Zhuang et al., "Connectivity Exploration with Structural Equation Modeling: An fMRI Study of Bimanual Motor Coordination," *NeuroImage* (2005): 462-70; A. Roebroeck, E. Formisano, and R. Goebel, "Mapping Directed Influence over the Brain Using Granger Causality and fMRI," *Neuroimage* 25 (2005): 230-42; K. J. Friston, L. Harrison, and W. Penny, "Dynamic Causal Modeling," *Neuroimage* 19 (2003): 1273-302; S. A. Huettel, A. W. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*, 2nd ed. (Sunderland, MA: Sinauer Associates, 2009), chapter 11.

20. See M. Coltheart, "What Has Functional Neuroimaging Told Us about the Mind (So Far)?," *Cortex* 42 (2006): 323-33, and G. van Orden and K. R. Papp, "Functional Neuroimages Fail to Discover Pieces of Mind in Parts of the Brain," *Philosophy of Science Proceedings* 64 (1997): S85-S94.

21. Coltheart, "What Has Functional Neuroimaging Told Us?" For a thoughtful explication of the ways in which neuroimaging can inform cognitive theory, including the point that no single finding in psychology will ever be a "Popperian 'falsification bullet,'" see also R. Henson, "What Can Functional Neuroimaging Tell the Experimental Psychologist?," *Quarterly Journal of Experimental Psychology* 58A (2005): 193-233.

22. Henson, "What Can Functional Neuroimaging Tell the Experimental Psychologist?"; C. Klein, "Philosophical Issues in Neuroimaging," *Philosophy Compass* 5 (2010): 186-98.

23. See Henson, "What Can Functional Neuroimaging Tell the Experimental Psychologist?"; J. Jonides, D. E. Nee, and M. G. Berman, "What Has Functional Neuroimaging Told Us about the Mind? So Many Examples, So Little Space," *Cortex* 42 (2006): 414-17; M. Mather, N. Kanwisher, and J. Cacioppo, "How fMRI Can Inform Cognitive Theories," *Perspectives on Psychological Science* 8 (2013): 108-13; E. E. Smith and J. Jonides, "Working Memory: A View from Neuroimaging," *Cognitive Psychology* 33 (1977): 5-42; C. Umilta, "Localization of Cognitive Functions in the Brain Does Allow One to Distinguish between Psychological Theories," *Cortex* 42 (2006): 399-401.

24. R. A. Finke, "Levels of Equivalence in Imagery and Perception," *Psychological Review* 87, no. 2 (1980): 113-32.

25. M. J. Farah, "Is Visual Imagery Really Visual? Overlooked Evidence from Neuropsychology," *Psychological Review* 95 (1988): 307-17.

26. Farah, "Is Visual Imagery Really Visual?"; S. M. Kosslyn et al., "Visual Mental Imagery Activates Topographically Organized Visual Cortex: PET Investigations," *Journal of Cognitive Neuroscience* 5 (1993): 263-87; M. D'Esposito et al., "A Functional MRI Study of Mental Image Generation," *Neuropsychologia* 35 (1997): 725-30.

27. Mather et al., "How Can FMRI Inform Cognitive Theories?"

28. Uttal, *Mind and Brain*, p. 366.

29. Ibid., p. 369.

30. Posner and Raichle, *Images of Mind*.

31. To obtain an indication of the methodological segregation or integration of functional neuroimaging research, the references cited by the last twenty fMRI articles published in *Science* as of May 21, 2013, were obtained and the literature cited therein was examined for research methods other than imaging. In all cases, the articles cited work outside of functional neuroimaging.

32. R. A. Poldrack, "Can Cognitive Processes Be Inferred from Neuroimaging Data?" *Trends in Cognitive Sciences* 10 (2006): 59-63.

33. Ibid.

34. M. Iacoboni, J. Freedman, and J. Kapla, "This Is Your Brain on Politics," *New York Times*, November 11, 2007.

35. See A. Aron et al, "Politics and the Brain," *New York Times*, November 14, 2007, http://www.nytimes.com/2007/11/14/opinion/lweb14brain.html; M. J. Farah, "This Is Your Brain on Politics," *Neuroethics & Law Blog*, November 12, 2007, accessed February 6, 2014, http://kolber.typepad.com/ethics_law_blog/2007/11/this-is-your-br.html; G. K. Aguirre, "The Political Brain," *Cerebrum* (blog), September 12, 2008, http://www.dana.org/Cerebrum/2008/The_Political_Brain/.

36. See H. D. Critchley et al., "Human Cingulate Cortex and Autonomic Control: Converging Neuroimaging and Clinical Evidence," *Brain* 126 (2003): 2139-52; G. Bush, P. Luu, and M. I. Posner, "Cognitive and Emotional Influences in Anterior Cingulate Cortex," *Trends in Cognitive Sciences* 4, no. 6 (2000): 215-22.

37. D. Ariely and G. S. Berns, "Neuromarketing: The Hope and Hype of Neuroimaging in Business," *Nature Reviews Neuroscience* 11 (2010): 284-92.

38. M. J. Farah, B. Hutchinson, E. A. Phelps, and A. D. Wagner, "fMRI Lie Detection: Scientific and Societal Challenges," *Nature Reviews Neuroscience* 15 (2014): 123-31, doi:10.1038/nrn3665.

39. Ibid.

40. L. Saxe and G. Ben-Shakhar. "Admissibility of Polygraph Tests: The Application of Scientific Standards Post-Daubert." *Psychology, Public Policy, and Law* 5, no. 1 (1999): 203-23.

41. I. M. Thompson et al., "Prevalence of Prostate Cancer among Men with a Prostate-Specific Antigen Level ≤4.0 ng per Milliliter," *New England Journal of Medicine* 351 (2004): 2239-46.

42. As quoted in D. Ramani, "The Brain Seduction: The Public Perception of Neuroscience," *Journal of Science Communication* 84, no. 4 (2009): 1-8.

43. Fodor, "Why the Brain?"

44. Shermer, "Why You Should Be Skeptical," pp. 70-71.

45. Wade, "Some Cautions about Jumping on the Brain Imaging Bandwagon."

46. C. M. Bennett et al., "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 (2010): 1-5.

47. See Huettel, Song, and McCarthy, *Functional Magnetic Resonance Imaging.*

48. P. C. Austin et al., "Testing Multiple Statistical Hypotheses Resulted in Spurious Associations: A Study of Astrological Signs and Health," *Journal of Clinical Epidemiology* 59 (2006): 964-69.

49. E. Vul et al., "Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition," *Perspectives on Psychological Science* 4 (2009): 274-90.

50. See S. Begley, "Of Voodoo and the Brain," *Newsweek*, January 30, 2009.

51. E. Vul and H. Pashler, "Voodoo and Circularity Errors," *Neuroimage* 62 (2012): 945-48, at 946.

52. M. B Crawford, "The Limits of Neuro-talk," *The New Atlantis* 19 (2008): 65-78, at 65.

53. S. Poole, "Your Brain on Pseudoscience: The Rise of Popular Neurobollocks," *New Statesman* (September 6, 2012), at paragraph 18.

54. S. B. Johnson, R. W. Blum, and J. N. Giedd, "Adolescent Maturity and the Brain: The Promise and Pitfalls of Neuroscience Research in Adolescent Health Policy," *Journal of Adolescent Health* 45, no. 3 (2009): 216-21, at 219.

55. J. Illes et al., "Neurotalk: Improving the Communication of Neuroscience Research," *Nature Reviews Neuroscience* 11 (2009): 61-69, at 61.

56. D. P. McCabe and A. D. Castel, "Seeing Is Believing: The Effect of Brain Images on Judgments of Scientific Reasoning," *Cognition* 107 (2008): 343-52.

57. D. Gruber and J. A. Dickerson, "Persuasive Images in Popular Science: Testing Judgments of Scientific Reasoning and Credibility," *Public Understanding of Science* 21 (2012): 938-48; C. J. Hook and M. J. Farah, "Look Again: Effects of Brain Images and Mind-Brain Dualism on Lay Evaluations of Research," *Journal of Cognitive Neuroscience* (early access, April 22, 2013, doi:10.1162/jocn_a_00407); R. B. Michael et al., "On the (Non)Persuasive Power of a Brain Image," *Psychonomic Bulletin & Review* 20, no. 4 (2013): 720-25; see also M. J. Farah and C. J. Hook, "The Seductive Allure of 'Seductive Allure,'" *Perspectives on Psychological Science* 8 (2013): 88-90, for a discussion of the McCabe and Castel study.

58. J. R. Gurley and D. K. Marcus, "The Effects of Neuroimaging and Brain Injury on Insanity Defenses," *Behavioral Sciences & The Law* 26, no. 1 (2008): 85-97.

59. D. P. McCabe, A. D. Castel, and M. G. Rhodes, "The Influence of fMRI Lie Detection Evidence on Juror Decision-Making," *Behavioral Sciences & the Law* 29, no. 4 (2011): 566-77.

60. Ibid.

61. P. Bloom, "Seduced by the Flickering Lights of the Brain," *Seed*, June 26, 2006, http://seedmagazine.com/content/article/seduced_by_the_flickering_lights_of_the_brain/.

62. Hook and Farah, "The Seductive Allure."

63. D. S. Weisberg et al., "The Seductive Allure of Neuroscience Explanations," *Journal of Cognitive Neuroscience* 20 (2008): 470-77; A. K. Lindell and E. Kidd, "Consumers Favor 'Right Brain' Training: The Dangerous Lure of Neuromarketing," *Mind, Brain, and Education* 7 (2013): 35-39.

64. A. D. Sokal, "A Physicist Experiments with Cultural Studies," accessed July 5, 2013, http://www.physics.nyu.edu/faculty/sokal/lingua_franca_v4/lingua_franca_v4.html; K. Eriksson, "The Nonsense Math Effect," *Judgment and Decision Making* 7 (2012): 746-49.