
Experimentation in Cognitive Neuroscience and Cognitive Neurobiology

3

Jacqueline Sullivan

Contents

Introduction	32
Neuroscience and the Experimental Process	32
Basic Structure of Experiments: Cognitive Neuroscience	32
Basic Structure of Experiments: Cognitive Neurobiology	34
The Experimental Process	35
Reliability and Validity	40
Epistemic Challenges	42
Conclusion	45
Cross-References	45
References	46

Abstract

Neuroscience is a laboratory-based science that spans multiple levels of analysis from molecular genetics to behavior. At every level of analysis, experiments are designed in order to answer empirical questions about phenomena of interest. Understanding the nature and structure of experimentation in neuroscience is fundamental for assessing the quality of the evidence produced by such experiments and the kinds of claims that are warranted by the data. This chapter provides a general conceptual framework for thinking about evidence and experimentation in neuroscience with a particular focus on two research areas: cognitive neuroscience and cognitive neurobiology.

J. Sullivan

Department of Philosophy and Rotman Institute of Philosophy, University of Western Ontario,
London, ON, Canada
e-mail: jsulli29@uwo.ca

Introduction

Neuroscience advances our understanding of the brain and behavior primarily by means of experimentation. Findings from neuroscience shape how we think about the nature of cognition, behavior, diseases and disorders of the mind and brain, consciousness, moral responsibility, and free will. Interpretations of data obtained from neuroscience have the potential to inform diagnostic and treatment decisions and impact assessments of moral culpability and legal responsibility. If the interpretations that neuroscientists make on the basis of data are not warranted, then any claims put forward or decisions made on the basis of that data will lack justification.

Understanding the nature and structure of experimentation in neuroscience and evaluating the explanatory/interpretive claims of neuroscience is crucial for avoiding such epistemological pitfalls. By bringing together insights from the philosophy of neuroscience, the philosophy of scientific experimentation, epistemology and theoretical work in neuroscience and psychology, this chapter puts forward a conceptual framework for thinking about evidence and experimentation in contemporary neuroscience. While the primary focus will be representative examples of experiments undertaken in cognitive neuroscience and cognitive neurobiology, some of the basic lessons are also relevant to experiments conducted in other laboratory-based areas of neuroscience.

Neuroscience and the Experimental Process

One way to think about experimentation in neuroscience and science more generally is as a process, which we may refer to simply as “the experimental process” (See Sullivan 2009). What are the aims of this process? In the simplest terms, the aim of experimentation is to produce data to discriminate among competing hypotheses about a phenomenon of interest. The data from an experiment will serve this function only to the extent that the process of producing those data was *reliable* and the claims made upon the basis of those data are *valid* (See section “[Reliability and Validity](#)” below). A worthwhile place to begin to think about evidence and experimentation in cognitive neuroscience and cognitive neurobiology is to say something about the nature of the kinds of claims about phenomena these two areas of science are interested in supporting and the basic types of experiments we find there.

Basic Structure of Experiments: Cognitive Neuroscience

The aims of *cognitive neuroscience* are, roughly, to locate regions of the brain that subserve cognitive functions, to identify patterns of connectivity between different brain regions, and to understand the processing of information through the brain. Cognitive neuroscience combines the conceptual-theoretical framework and

experimental paradigms of cognitive psychology with structural and functional neuroimaging and electrophysiological recording techniques. Experimentation in cognitive neuroscience is based on several basic assumptions. First, organisms have specific kinds of cognitive capacities. Second, these cognitive capacities may be individuated by appropriately designed experimental tasks. Third, for any given cognitive capacity that can be delineated experimentally, it is possible to locate the neural basis of that cognitive capacity in the brain. Identifying the neural basis of a cognitive capacity is assumed to be achievable by correlating (a) subjects' behavioral performance on experimental tasks or their subjective reports with (b) measurable brain activity.

Experiments in cognitive neuroscience combine the use of experimental paradigms/cognitive tasks of cognitive psychology with computational models, neuroimaging, and electrophysiological techniques. A typical experiment in cognitive neuroscience often begins by pointing to previous findings in the empirical literature pertaining to the nature of a specific cognitive function (e.g., face recognition) and the brain area(s) thought to subservise it. These findings are then used as a basis to make testable predictions that are often formulated as competing correlational claims. An example of an empirical question might be: Is the perirhinal (PrC) cortex involved in face recognition? To address this question, an investigator will make predictions about what brain activity in the PrC during a face recognition task ought to look like if it is indeed the case that the PrC is involved. For example, three competing hypotheses may prevent themselves: h_1 : Activity in the PrC is increased compared to baseline activity (or activity on a different cognitive task), h_2 : Activity in the PrC is decreased compared to baseline activity (or activity on a different cognitive task), and h_3 : There is no change in PrC activity compared to baseline activity (or activity on a different cognitive task) (null). In the best-case scenario, the data will adjudicate between these three competing hypotheses and point to the one that is best supported by the data.

While different procedures exist for correlating the behavioral expression of a cognitive function with neural activity (e.g., evoked response potentials (ERPs), positron emission tomography (PET)), by far the most widely employed technology in contemporary cognitive neuroscience, and the one that both philosophers and neuroscientists themselves have questioned the reliability of, is functional magnetic resonance imaging (fMRI). In a typical fMRI experiment, a subject is placed into a magnetic resonance imaging (MRI) scanner and trained in an experimental paradigm. An experimental paradigm is roughly a standard set of procedures for producing, measuring, and detecting a cognitive capacity in the laboratory that specifies how to produce that capacity, identifies the response variables to be measured during pre-training, training, and post-training/testing, and includes instructions on how to measure those response variables using appropriate equipment. It also specifies how to detect a cognitive capacity when it occurs by identifying what the comparative measurements of the selected response variables have to equal in order to ascribe that capacity to a subject (Sullivan 2009). Given that a subject placed in an MRI scanner is physically constrained, the experimental paradigms used in conjunction with fMRI have historically been computer-based

tasks in which the stimuli are presented to subjects on a flat-screen monitor. Subjects elicit behavioral responses to these stimuli or answer questions about them, depending on the instructions provided to them, typically by means of pressing a button.

During task performance or response elicitation, the investigator “scans” the subject’s brain, focusing on one or several regions of interest (ROIs). The investigator assumes that when a subject performs a task capable of individuating a discrete cognitive capacity, there will be an increase in neural activity compared to baseline activity in those brain regions involved in task performance. To detect such increases in activity, cognitive neuroscientists rely on the blood-oxygen level dependent (BOLD) response signal. The basic idea is that an increase in neural firing in a given region of the brain triggers a hemodynamic response such that blood is delivered to that area at a more rapid rate than blood that nourishes less active neurons. This increase is accompanied by an increase in oxygen utilization and thus an increase in the amount of deoxygenated blood in the region activated compared to oxygenated blood in the surrounding regions. This difference is used as a contrast for distinguishing areas of heightened activity from areas of less heightened activity. While the subject is in the scanner, sample scans of the brain are taken across a selected time course. Time points of sampling are intended to be coordinated as closely as possible with features of the experimental paradigm such as stimulus presentation and the relevant response output (e.g., button pressing). Once enough data has been collected, investigators pre-process the data in order to eliminate experimental artifacts (e.g., motion of subject while in scanner) and increase signal-to-noise. The data are then processed using statistical analysis techniques (e.g., univariate analysis). The statistically analyzed data is then used as a basis for discriminating among competing functional hypotheses about the brain areas under investigation.

Basic Structure of Experiments: Cognitive Neurobiology

A primary aim of cognitive neurobiology is to discover the cellular and molecular mechanisms of learning and memory (e.g., Sweatt 2009). Cognitive neurobiology combines the behavioral techniques of experimental psychology and electrophysiological, pharmacological, genetic, and protein analysis techniques. A basic set of assumptions informs experimentation in cognitive neurobiology. First, all organisms, from the most simple to the most complex, learn and remember. Second, different forms of learning and memory are detectable by appeal to observable changes in behavior, and may be individuated by appropriately designed experimental learning paradigms. Third, many if not all forms of learning and memory require changes in synaptic strength. Fourth, the changes in synaptic strength that underlie learning and memory are mediated by changes in protein activity and gene expression.

Cognitive neurobiological experiments typically test both correlational and causal or mechanistic claims. Oftentimes an experiment will begin with a question of

whether synaptic, cellular, or molecular activity in the brain *is implicated in* changes in the behavior of an organism or a synapse. For example, in a typical behavioral experiment, an investigator will make predictions about what the measurable changes in cellular and molecular activity and in behavior as a result of training organisms in an experimental paradigm ought to look like in order to establish a correlation between the two factors. Once a correlation between two measurable factors has been established, an investigator typically undertakes intervention experiments. Intervention experiments test predictions about the impact of blocking cellular and molecular activity (with either pharmacological or genetic techniques) on the changes in behavior observed from training organisms in the experimental paradigm. Examples of representative competing hypotheses usually take the following form: h_1 : Blocking molecular activity is accompanied by measurable changes in behavior that indicate the learning has been blocked, h_2 : Blocking molecular activity is accompanied by measurable changes in behavior that indicate that learning is not blocked, or h_3 : Blocking molecular activity results in other measurable changes in behavior that are unexpected (e.g., partial blockade). Again, ideally, the data will adjudicate between the competing hypotheses, discriminating the one that is best supported by the data.

The experimental process in both cognitive neuroscience and cognitive neurobiology is heavily informed and shaped by evidence emanating from other scientific disciplines, including cellular and molecular neuroscience, genetics, psychology, physiology, biochemistry, neuroanatomy, and systems neuroscience. Evidence from these areas serves as a basis for making predictions, formulating testable hypotheses, designing experiments to test those hypotheses, and interpreting the data obtained from such tests. Generally speaking, cognitive neuroscientists and cognitive neurobiologists aim to design experiments to test their predictions and to adjudicate between competing claims about phenomena of interest. When we ask whether neuroscientists succeed at this goal, we are asking whether experimentation in neuroscience is sufficient to yield the data requisite to achieve it.

The Experimental Process

If we want to determine if cognitive neuroscience and cognitive neurobiology are knowledge-generating, it is insufficient to look exclusively at textbook descriptions or reviews of neuroscientific findings. The best unit of analysis is the individual research paper, because it is as close as we can get to the experimental process (without visiting the lab). More specifically, evaluating the merits of already published research papers, which in neuroscience is the aim of lab meetings and journal clubs, is the first step toward answering the question of whether the interpretive or explanatory claims being made in a given research paper are warranted by the data.

Although we often take it for granted that each and every scientist ensures the integrity and reliability of the experimental process himself/herself, it is important to remember that the peer-review process exists in part, because scientists are fallible.

A scientist may believe he/she has adequately tested a hypothesis in instances in which he/she has overlooked potential confounding variables or has failed to exclude or neglected to consider alternative explanations for the results. While peer-review is intended to catch such errors, it offers no full-proof guarantee that science produces knowledge, because individuals on peer-review boards are human and thus fallible, too.

However, even with our unit of analysis being an individual research paper, our access to the experimental process is limited. The introduction provides us with insight into the assumptions that informed a given research study and the origin(s) of the empirical question(s) that the study aims to answer. The methods section provides details about the subjects or objects of the study (e.g., college-age human beings, adult male Wistar rats) and the materials, tools, and techniques used. The results section simply conveys the outcomes of using the methods to answer the empirical question(s). These outcomes are not raw data; they are statistically analyzed data that are typically represented in pictures, charts, and diagrams. The discussion section teases out the implications of the study and attempts to situate the findings within the relevant literature. However, oftentimes the kinds of things that may compromise the knowledge-producing capacity of the experimental process of a given research study are hidden from view. The task is upon us to make the aspects of the process that may potentially compromise the knowledge outcomes of the study (e.g., problematic or unwarranted assumptions, investigator errors, equipment malfunctions, mathematical errors, errors of reasoning and analysis) explicit and to do as thorough a probe of the state-space as possible in order to rule out the possibility that errors were made – to make certain the data can be used to support the interpretative or explanatory claims that the investigators aim to make on the basis of the study.

An appropriate set of analytic tools and a strategy for their application may guide the way. While such tools are not antidotes to error, they at least point us in the direction of where to look for problems in the experimental process that may compromise the ability to use data to substantiate claims about phenomena of interest. What follows is one such set of tools that incorporates insights from philosophy of science, philosophy of experimentation and theoretical work in psychology and the social sciences.

The experimental process has discrete stages (Fig. 3.1). It is set in motion when an investigator or research team poses an empirical question about a phenomenon of interest. Examples of empirical questions in cognitive neuroscience may include: What area(s) of the brain are involved in attention? What kinds of information do dopamine neurons in the ventral striatum encode, represent, or process? What brain areas receive information from mirror neurons? Examples in cognitive neurobiology include: What is the role of protein kinase A in spatial memory? Is activation of cyclic-AMP response element binding protein necessary for long-term potentiation in area CA1 of the hippocampus in vivo? Are synaptic changes that accompany learning and memory similar to those that underlie addiction?

The phrase “phenomenon of interest” is intended to only very loosely capture the idea that *something* prompts an investigator to conduct an experiment – *some*

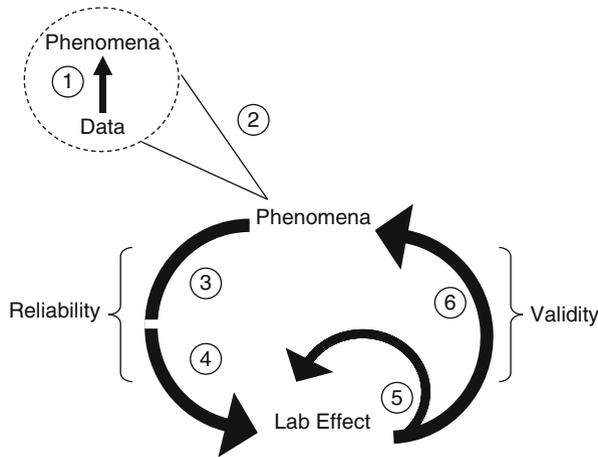


Fig. 3.1 The experimental process. (1) An investigator begins with an empirical question about a phenomenon of interest. This question is then redirected at an effect to be produced in the laboratory, thus initiating the (3) design and (4) implementation stages of data production. If the data production process is reliable, it results in the discrimination of one hypothesis from a set of competing hypotheses about the effect produced in the laboratory. This initiates the stage of data interpretation, in which the discriminated hypothesis is treated as a claim and is taken as true of (5) the effect produced in the laboratory and (6) the original phenomenon of interest in the world. If the claim was produced by a reliable data production process and it is true of the effect produced in the lab, it is valid (internal validity). If it was produced by a reliable data production process and it is true of the effect in the world, it is valid (external validity) (Sullivan 2009)

phenomenon of interest to him/her. One question that is relevant is how an investigator identifies, detects, or conceives of that phenomenon of interest. An obvious answer, if we look at modern neuroscience, is that in the history of cognitive neuroscience and cognitive neurobiology, some terms have been and continue to be widely deployed, although there is no consensus about how generally to define them. Despite such disagreements, constructs such as attention, working memory, face recognition, and spatial learning, are put forward as starting points for empirical inquiry – investigators pose questions that are directed at shedding light on at least as subset of those phenomena picked out by the concept. One question that remains, though, is how those phenomena that come to be designated by a general construct are identified or detected in the first place. One answer to this question is, if we consider cognitive phenomena more generally, that investigators notice changes in the behavior of organisms from some baseline, which serve as data points for their detection (Step 1 in Fig. 1). Bogen and Woodward (1988), for example, introduce a distinction between “data” and “phenomena” and commit themselves to the idea that phenomena are not observable, but only detectable by means of reference to data, which are observable. However, given that investigators begin experiments with questions about something that *is* detectable, whatever that phenomenon is, it is best understood as detectable derivatively, by means of reference to “data points.” For example, most human beings (and non-human

animals) can recognize, after one or more encounters, the faces of conspecifics. This is something that can be detected by noting that on the second or third encounter with a face that was originally novel, an individual's behavior will reflect such recognition. Experiments in cognitive neuroscience and cognitive neurobiology may be said to have their original starting point in such changes in behaviors as exhibited by organisms "in the world." They may also begin with a phenomenon that has been detected by means of data points in the controlled environment of the laboratory (Step 2 in Fig. 3.1). There is most likely a complicated story that could be told as to how an investigator arrived at a particular empirical question. Teasing out this story – looking across review papers and conducting an historical study of the construct/phenomenon in question can be revealing when one attempts to assess the kinds of interpretive claims about a phenomenon that the data obtained from a given research study may be used to support (See, for example, Sullivan 2010).

The experimental process in neuroscience may be regarded as involving two stages: (1) *data production* and (2) *data interpretation* (Woodward 2000). Once an investigator poses an empirical question about a phenomenon of interest, the process of data production begins. Data production may be divided into two discrete stages: (1.1) *design* and (1.2) *implementation*.

The design stage, in basic terms, involves the development of an experimental design and protocol. An experimental design includes the overall set-up of the experiment, in so far as it specifies such things as the experimental context (e.g., how and where objects are to be arranged) and the materials and methods to be used. The experimental protocol is the set of step-by-step instructions that an investigator follows each time he or she runs an experiment. An experimental protocol essentially specifies how each individual experiment is to be run from start to finish. When an investigator is in the middle of an experiment and confused about what to do next – he or she will refer to the experimental protocol (not the experimental design).

Once an investigator has identified a phenomenon of interest, a way to produce that phenomenon of interest in the laboratory must be specified. The phenomenon, whether it is a cognitive function or a form of synaptic plasticity, must be operationally defined. An operational definition is built directly into the design of an experimental paradigm. An experimental paradigm is a standard method or procedure for producing an effect of a specific type. The following features are typically included in the design of experimental paradigms in cognitive neuroscience and cognitive neurobiology: (1) production procedures, namely, a specification of the stimuli (independent or input variables) to be presented, how those stimuli are to be arranged (e.g., spatially, temporally), and how many times they are to be presented during phases of (a) pre-training, (b) training, (c) post-training/testing; (2) measurement procedures that specify the response variables to be measured in the (a) pre-training and (b) post-training/testing phases of the experiment and how to measure them using apparatuses designed for such measurement; (3) detection procedures that specify what the comparative measurements of the response variables from the different phases of the experiment must equal in order to ascribe the cognitive function of interest to the organism, the locus of the function to a given brain area or neuronal population, or a plastic change to a synapse or set

of synapses. This detection procedure is simply an operational definition that specifies the measurable “change” in response variables that must be observed in order to say that the relevant phenomena have occurred.

Investigators in both cognitive neuroscience and cognitive neurobiology have freedom to design experiments – to vary the features of experimental paradigms in ways that they deem most appropriate for their explanatory aims. A “sub-protocol” is a production procedure, written up step-by-step, which corresponds to an experimental learning or electrophysiological stimulation paradigm. It will, for example, specify: (1) the duration of time of the presentation of each stimulus to be used in an experiment, (2) the duration of time that is to elapse between presentation of the stimuli used in an experiment, or the inter-stimulus interval (ISI), (3) the amount of time that is to elapse between individual trials, or the inter-trial interval (ITI), and (4) the amount of time that is to elapse before testing (or biochemical analysis). From the reader’s perspective, the multiplicity of experimental protocols and its implications (Sullivan 2009) are aspects of experimentation that we ought to be privy to when comparing results across different laboratories. This is because subtle changes in experimental paradigms and subprotocols may yield different and sometimes inconsistent results with respect to the phenomenon of interest, rendering it unclear which results should be taken seriously or how to fit the results into a coherent picture or model.

In cognitive neuroscience, the design stage also involves the selection of a subject population, a brain area of interest, experimental techniques (e.g., fMRI, EEG), and statistical analysis procedures (e.g., multivariate pattern analysis (MVPA)). In cognitive neurobiology, the design stage involves the selection of a model organism, a neuronal population or set of synapses, experimental technologies (electrophysiology, biochemistry, immunohistochemistry), and the statistical analysis procedure.

The design stage of data production typically proceeds in discrete stages: Questions are posed and suggestions about how to address them are provided; projections are then made about potential problems that might be encountered in the course of implementing the design and tentative solutions to these problems are offered; and finally, the combined considerations are worked into the design and protocol. Essentially, at this stage, the empirical question of interest is directed at some effect to be produced in the lab.

The implementation stage of data production (Step 4 in Fig. 3.1) begins at some point after an experimental design and protocol has been completed. It involves individual instantiations of the experimental design by means of the systematic following of the experimental protocol using the equipment, materials, and techniques assembled during the design stage. At this point, an investigator takes an individual subject or a group of subjects, and runs them through the steps of the protocol, following those steps as precisely as possible. The immediate output of each individual implementation of the design is an individual data point or set of data points.

Once enough data points for each type of experimental manipulation have been collected, the data points are combined and each complete data set is analyzed statistically. The statistically analyzed data is then used to discriminate one hypothesis from the set of competing hypotheses about the phenomenon of interest

produced in the laboratory. The process of data interpretation then begins. In the first phase of data interpretation (Step 5 in Fig. 1), the hypothesis discriminated by the data is taken as true with respect to the effect produced in the laboratory. That same claim may then be extended back to the original phenomenon of interest in the world that prompted the empirical question in the first place (Step 6 in Fig. 1).

Reliability and Validity

Individual researchers working in laboratories are interested in producing the data requisite to discriminate among competing (correlational, causal, or mechanistic) claims about a single phenomenon of interest. In the ideal case, they aim to design an experiment or set of experiments to produce a set of data e in order to adjudicate between a set of competing hypotheses, h_1 , h_2 , and h_3 , about a phenomenon of interest. To do so, the evidence has to be adequate to this purpose. First, the data has to be the outcome of a reliable data production process. What does it mean for a data production process to be reliable? Mayo's (1991) "severity criterion" offers one understanding. In order for a test of a hypothesis to be reliable, it must pass a *severe* test – it must be *highly probable* that the data arising out of a test of a hypothesis would not yield evidence in support of that hypothesis if that hypothesis were in fact false. A related way of understanding reliability is that the process of producing data may be deemed *reliable* if and only if it results in statistically analyzed data that can be used to discriminate one hypothesis from a set of competing hypotheses about an effect produced in the laboratory (See also Bogen and Woodward 1988; Cartwright 1999; Franklin 1986, 1999; Mayo 1991, 1996, 2000; Woodward 1989, 2000). Reliability ought to operate as a constraint on the experimental process. When assessments are made about whether an experiment is reliable given the hypotheses it was designed to discriminate among, how the hypotheses are formulated is fundamental for assessing if the data may serve as adequate evidence.

A second desirable feature of the experimental process, which differs from reliability, is *validity*. Scientific accounts traditionally make use of a general notion of validity, which is taken to be a feature ascribed to experiments or tests. According to these accounts, an experiment is regarded as *valid* if it supports the conclusion that is drawn from its results (e.g., Campbell and Stanley 1963). Scientists and philosophers draw a distinction between external and internal validity (e.g., Cook and Campbell 1979; Guala 2003, 2005). Investigators not only wish to have the conclusions of their results apply to the effects under study in the laboratory (internal validity), they also hope that these conclusions apply to the phenomena of interest at which their empirical questions were originally directed (external validity).

For example, on Francesco Guala's account (2003, 2005), the *internal validity* of an experimental result is established when that result captures a causal relationship that operates in the context of the laboratory. That experimental result is *externally* valid when it captures a causal relationship that operates in "a set of circumstances of interest," outside the laboratory. However, validity may also be understood as a feature of interpretive claims rather than of experimental results. Whereas

experimental results are statistically analyzed sets of data, interpretive claims are what arises when a hypothesis that has been discriminated from a set of competing hypotheses by a set of data is taken as true of an effect produced in the laboratory as well as the original phenomenon of interest outside the laboratory. On this understanding of validity, an interpretive claim about an effect produced in a laboratory, is *internally* valid if and only if that claim is true about the effect produced in the laboratory. A claim about a phenomenon of interest outside the laboratory is *externally* valid if and only if that claim is true about that phenomenon.

One way to understand the relationship between reliability and validity is that they operate as normative constraints on the experimental process, yet give rise to conflicting prescriptions. Reliability prescribes simplifying measures in the context of the laboratory in order to narrow down a set of competing hypotheses about the effect produced in the laboratory. Insofar as it operates to constrain the process of data production, it inevitably restricts the extension of interpretive claims to the laboratory. Validity, however, pulls in the opposite direction. It prescribes that an investigator build into an experimental design those dimensions of complexity that accompany the phenomenon of interest in the world about which an investigator would like to say something. Adhering to the normative prescriptions of validity will inevitably lead to a decrease in the simplicity of the effect produced in the laboratory and an expansion of the set of competing hypotheses that pertain to that effect. In other words, it will lead to a decrease in reliability. However, without reliability, nothing is gained – for if control is lost in the laboratory, nothing true can even be said about the effect produced there – internal validity will be lost as well.

Although not represented explicitly in Fig. 3.1, it is relevant to mention two other types of validity that also may function as constraints on the experimental process. The first constraint is *ecological validity* (See, for example, Bronfenbrenner 1979; Schmuckler 2001). In contrast to external validity, which is concerned with whether an interpretive claim arrived at in a given study may be extended to the real world, ecological validity is concerned with whether the context, stimuli employed, and responses elicited in the experimental context are similar to those that would be found in the world. For example, performing a cognitive task in an fMRI scanner is different than engaging in a cognitive activity in a less restricted environment, so we might say that experiments using fMRI are not ecologically valid. The second type of validity that may constrain the experimental process is *construct validity* (See, for example, Cronbach and Meehl 1955; Shadish et al. 2002). The basic idea here is that investigators in cognitive neuroscience and cognitive neurobiology are interested in developing experimental paradigms that individuate specific cognitive capacities, because they want to be able to make structure-function or mechanistic claims about those capacities. This means that it ought to be the case that the effect under study in the laboratory is an actual instance of the phenomena picked out by a given construct (e.g., “attention”). Notice that if the constraint of construct validity is not met, this poses a problem for reliability – since an investigator may only use data to adjudicate between competing claims about a cognitive capacity produced in the laboratory if it is actually the case that the effect produced by means of an experimental paradigm is an actual instance of that capacity (See Sullivan 2010 for further discussion).

Failure to meet the criterion of construct validity should ideally prompt investigators to look for or develop an experimental paradigm that does a better job at individuating the capacity of interest.

Epistemic Challenges

The conceptual framework offered in the sections above entitled “[The Experimental Process](#)” and “[Reliability and Validity](#)” may be applied to research papers in cognitive neuroscience and cognitive neurobiology in order to illuminate the steps of the process and identify the various points that decisions are made or courses of action are taken that may impact the reliability of the data production process, the internal and external validity of the interpretive/correlational/causal claims made on the basis of the data, and ecological and construct validity. The framework may also serve as a basis for comparing the experimental process across research studies and determining what kind of interpretive claims are supported by a given body of data. Finally, using this conceptual framework as a backdrop, we can group together epistemic challenges for experimentation in cognitive neuroscience and cognitive neurobiology that have already been identified in the philosophical literature. This is the primary aim of this section.

For example, many philosophers have urged caution with respect to determining the kinds of structure-function claims that fMRI data may be used to support (e.g., Bechtel and Stufflebeam 2001; Bogen 2001, 2002; Delehanty 2007, 2010; Hardcastle and Stewart 2002; Klein 2010a, b; Mole et al. 2007; Roskies 2007, 2010; Uttal 2001, 2011, 2013; van Orden and Paap 1997). A common strategy is to identify the points in the data production process where techniques are used or decisions are made that may jeopardize or compromise the reliability of that process. If we begin by considering the design stage of data production, experiments using fMRI may be regarded as failing with respect to *ecological validity* in so far as subjects perform cognitive tasks that are designed to be implemented while a subject lies still and constrained inside the scanner. However, the cognitive activities that such experiments are supposed to shed light on the neural mechanisms of take place in far less restricted environments. Experiments using fMRI will thus always be limited when it comes to satisfying the criterion of ecological validity. Additionally, given that it is not clear that correlational claims about cognitive functions under study within the confined conditions of the laboratory may be extended beyond that context, fMRI experiments may also be regarded as lacking *external validity*. This does mean, however, that investigators learn nothing about “real-world” cognitive activities when they use fMRI. Rather, it means that we need to think carefully about what kinds of claims are supported by the data.¹

¹It is also relevant to note, that some investigators have sought to increase the ecological validity of fMRI experiments with innovative methods that allow for 3-D (as opposed to 2-D) objects to be used within the scanner (See Snow et al. 2011).

A third issue pertains to *construct validity*. For example, performing a face recognition task in a scanner with 2-dimensional stimuli presented on a flat-screen monitor is clearly different from being presented with real 3-D faces. This prompts the question of whether learning about face recognition with 2-D faces is revealing with respect to all of the phenomena that we typically identify as instances of facial recognition, which includes recognition of 3-D faces. A second and related issue, also having to do with construct validity, is whether an experimental paradigm used to study face recognition is sufficient for individuating the cognitive capacity it is intended to measure. For example, face recognition is a complex cognitive function that requires both attentional and mnemonic processes. Thus, an experimental paradigm implemented in an fMRI scanner ought to be able to differentiate the function of attending to faces from face recognition. Although cognitive neuroscientists emphasize the importance of task analysis as a means to ensure the construct validity of their experimental paradigms, they often disagree about which experimental paradigms are the best for measuring different cognitive functions. Such disagreements have prompted skeptics like Uttal (e.g., 2001) to argue that an objective taxonomy of cognitive functions will never be possible. However, some philosophers regard this as far too skeptical a conclusion (See, for example, Landreth and Richardson 2004).

The important point is that if concerns about ecological, external, and construct validity do not shape the development of an experimental design and experimental paradigm and protocol, this will likely impact the kinds of interpretive claims that are warranted by the data. In contrast, since hypotheses typically make reference to the cognitive capacity of interest, if the experimental paradigm is insufficient for individuating that discrete cognitive capacity, then the data will be unreliable for discriminating among competing hypothetical claims pertaining to that cognitive capacity.

Philosophical scrutiny has also been directed at the reliability of data production processes that involve the use of fMRI technology. For example, philosophers have pointed to the fact that the occurrence of the BOLD signal does not directly correlate with task-related neural activity, thus making it a potentially unreliable indicator of such activity (e.g., Bechtel and Stufflebeam 2001; Bogen 2001, 2002; Klein 2010a, b; Roskies 2007). Third, investigators are not always able to distinguish task-related effects from mere artifacts when looking at the raw data. Guesswork is typically required to improve the signal-to-noise ratio in data collected from each subject (i.e., within-subject data) and to eliminate artifacts (e.g., head motion during scanning) before processing the data. Such guesswork leaves open the possibility of experimenter error. Fourth, the fMRI data for each experimental condition has to be determined and averaged across subjects. This requires that the data be mapped and fitted onto an atlas of the brain (e.g., Talarach atlas). Given differences in the shape and sizes of subjects' brains, averaging the data across subjects and fitting it into the atlas leaves open the possibility of data distortion. Another problem concerns the method of subtraction. In order to determine which area of the brain is involved in which cognitive task, investigators compare the BOLD signal observed on two task conditions that are thought to differ exclusively

with respect to one cognitive activity. For example, face recognition is thought to involve familiarity as well as recollection. One might thus imagine that a subject could be run in a face recognition paradigm and a familiarity paradigm and that activity observed in the familiarity paradigm could be subtracted from that in the face recognition paradigm to yield that area of the brain that is relevant for recognition. However, this method assumes that the two tasks actually discriminate between these two cognitive capacities, which may not be the case (For further discussion, see Bechtel and Stufflebeam (2001); Bogen (2001, 2002); Klein (2010a, b); Roskies (2007, 2010)).

A final issue with fMRI concerns what can be concluded on the basis of fMRI images. As several philosophers have argued, fMRI images are themselves outcomes of data interpretation rather than products of the data production process (e.g., Bogen 2002; Klein 2010b; Roskies 2007). Thus, conclusions that are made on the basis of these images – i.e., using the images themselves to adjudicate among competing hypotheses concerning structure-function relationships in the brain – will fail if decisions made during the stages of data processing involve the introduction of errors that fail to preserve the integrity of the raw data. This is one reason why philosophers have argued that analytic scrutiny must be directed at the analytical techniques involved in the production of fMRI images (e.g., Bogen 2002; Klein 2010b; Roskies 2010).

Despite the apparent limitations of fMRI technology, it continues to be widely used in cognitive neuroscience. Many neuroscientists, however, are aware and openly acknowledge these limitations and are in search of more reliable approaches to locating regions of the brain that subservise cognitive functions, identifying patterns of connectivity between different brain regions, and understanding the processing of information through the brain (See, for example, Logothetis 2008; Culham 2013 <http://culhamlab.ssc.uwo.ca/fmri4newbies/>).

Cognitive neurobiological experiments have also been a target of philosophical analysis. First, when it comes to the process of data production, cognitive neurobiologists have traditionally been concerned almost exclusively with reliability of the data production process and less concerned with issues of external, ecological, and construct validity. Given that investigators aim to establish causal relationships between cellular and molecular activity and behavior, in order to rule out the possibility of confounding variables, animal subjects are often raised in impoverished environments and trained with types of stimuli having parameters they would be unlikely to encounter in the real world (See Sullivan (2007; 2009) for further discussion). The dissimilarity between the laboratory and the external world thus jeopardizes the ability to extend causal claims established in the laboratory to real-world contexts (See Sullivan (2009) for further discussion).

Another issue that arises with respect to experiments in cognitive neurobiology is that not all investigators are concerned with construct validity. Many investigators are less interested in the cognitive processes that occur when an animal is trained in an experimental learning paradigm than with obtaining data that indicates that an observable change in behavior has occurred. Such data is then used as a basis

for inferring that the cognitive function that the paradigm purportedly individuates has been detected. However, sometimes it is unclear what cognitive capacity a given experimental paradigm actually individuates, which compromises the ability to use data collected using that paradigm as a basis for making causal claims about the role of cellular and molecular activity in a discrete cognitive function (See Sullivan (2010) for further discussion).

Philosophers have also addressed the question of whether results from experiments using model organisms, which are commonplace in low-level neuroscience and the neurobiology of learning and memory, are extendable to the human case (e.g., Ankeny 2001; Burian 1993; Schaffner 2001; Steel 2008; Sullivan 2009). Model organisms include, for example, rodents, sea mollusks, and fruit flies. These organisms are referred to as “models” in so far as scientists use them to establish causal relationships that they aim to generalize to the human population. However, differences between the two populations (i.e., laboratory animals and human beings) and the two contexts (lab versus ordinary environment) complicate the extrapolation of findings from the one context to the other. This prompts the question: When is the extrapolation of causal claims from the one context to the other warranted? Proponents of extrapolation, such as Daniel Steel (2008), have sought to provide an account that puts the investigative strategy on firmer epistemological footing. Of course, strategies for improving extrapolation from model organisms to the human case will vary depending upon the kinds of causal claims at issue (e.g., Sullivan 2009).

Conclusion

Philosophical work on the epistemology of experimentation in neuroscience, as is evident in the above examples, has been directed primarily at the knowledge-generating capacities of specific investigative strategies, tools, and techniques. However, neuroscience is a rapidly expanding field with global aims, the achievement of which requires the development of new and complex technologies. Ideally, we would like to have a workable set of analytic tools that we could apply in different areas of neuroscience and direct at different investigative strategies with the aim of determining whether those investigative strategies are knowledge-generating. Identifying one general set of conceptual tools has been the aim of this article.

Cross-References

- ▶ [Brain Research on Morality and Cognition](#)
- ▶ [Human Brain Research and Ethics](#)
- ▶ [Neuroimaging Neuroethics: Introduction](#)

References

- Ankeny, R. (2001). Model organisms as models: Understanding the 'lingua franca' of the human genome project. *Philosophy of Science*, 68, S251–S261.
- Bechtel, W., & Stufflebeam, R. S. (2001). Epistemic issues in procuring evidence about the brain: The importance of research instruments and techniques. In W. Bechtel, P. Mandik, J. Mundale, & R. S. Stufflebeam (Eds.), *Philosophy and the neurosciences: A reader* (pp. 55–81). Oxford: Blackwell.
- Bogen, J. (2001). Functional imaging evidence: Some epistemic hotspots. In P. K. Machamer, P. McLaughlin, & R. Grush (Eds.), *Theory and method in the neurosciences*. Pittsburgh: University of Pittsburgh Press.
- Bogen, J. (2002). Epistemological custard pies from functional brain imaging. *Philosophy of Science*, 69, S59–S71.
- Bogen, J., & Woodward, J. (1988). Saving the phenomena. *The Philosophical Review*, 97, 303–352.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge: Harvard University Press.
- Burian, R. M. (1993). How the choice of experimental organism matters: Epistemological reflections on an aspect of biological practice. *Journal of the History of Biology*, 26, 351–367.
- Campbell, D. D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand-McNally.
- Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. Cambridge: Cambridge University Press.
- Cook, T. D., & Campbell, D. D. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand-McNally.
- Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Culham, J. (2013). Functional imaging for newbies. (<http://culhamlab.ssc.uwo.ca/fmri4newbies/>)
- Delehanty, M. (2007). Perceiving causation via videomicroscopy. *Philosophy of Science*, 74(5), 996–1006.
- Delehanty, M. (2010). Why images? *Medicine Studies*, 2(3), 161–173.
- Franklin, A. (1986). *The neglect of experiment*. New York: Cambridge University Press.
- Franklin, A. (1999). *Can that be right? Essays on experiment, evidence, and science*. Boston: Kluwer.
- Guala, F. (2003). Experimental localism and external validity. *Philosophy of Science Supplement*, 70, 1195–1205.
- Guala, F. (2005). *The methodology of experimental economics*. Cambridge: Cambridge University Press.
- Hardcastle, V. G., & Stewart, C. M. (2002). What do brain data really show? *Philosophy of Science*, 69, S72–S82.
- Klein, C. (2010a). Philosophical issues in neuroimaging. *Philosophy Compass*, 5(2), 186–198.
- Klein, C. (2010b). Images are not the evidence in neuroimaging. *British Journal for the Philosophy of Science*, 61(2), 265–278.
- Landreth, A., & Richardson, R. C. (2004). Localization and the new phrenology: A review essay on William Uttal's *The New Phrenology*. *Philosophical Psychology*, 17, 108–123.
- Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, 453, 869–878.
- Mayo, D. (1991). Novel evidence and severe tests. *Philosophy of Science*, 58, 523–552.
- Mayo, D. (1996). *Error and the growth of experimental knowledge*. Chicago: University of Chicago Press.
- Mayo, D. (2000). Experimental practice and an error statistical account of evidence. *Philosophy of Science*, 67(3), S193–S207.
- Mole, C., Plate, J., Waller, R., Dobbs, M., & Nardone, M. (2007). Faces and brains: The limitations of brain scanning in cognitive science. *Philosophical Psychology*, 20(2), 197–207.

- Roskies, A. (2007). Are neuroimages like photographs of the brain? *Philosophy of Science*, 74(5), 860–872.
- Roskies, A. (2010). Neuroimaging and inferential distance: The perils of pictures. In M. Bunzl, & S. Hansen (Eds.), *Foundations of functional neuroimaging*. Cambridge: MIT Press.
- Schaffner, K. (2001). Extrapolation from animal models: Social life, sex, and super models. In P. K. Machamer, P. McLaughlin, & R. Grush (Eds.), *Theory and method in the neurosciences*. Pittsburgh: University of Pittsburgh Press.
- Schmuckler, M. (2001). What is ecological validity? A dimensional analysis. *Infancy*, 2(4), 419–436.
- Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.
- Snow, J. C., Pettypiece, C. E., McAdam, T. D., McLean, A. D., Stroman, P. W., Goodale, M. A., & Culham, J. C. (2011). Bringing the real world into the fMRI scanner: Repetition effects for pictures versus real objects. *Science Reports* 1, Article 130.
- Steel, D. P. (2008). *Across the boundaries: Extrapolation in biology and social science*. Oxford: Oxford University Press.
- Sullivan, J. A. (2007). *Reliability and validity of experiment in the neurobiology of learning and memory*. Dissertation, University of Pittsburgh.
- Sullivan, J. (2009). The multiplicity of experimental protocols: A challenge to reductionist and non-reductionist models of the unity of neuroscience. *Synthese*, 167, 511–539.
- Sullivan, J. (2010). Reconsidering “spatial memory” and the Morris water maze. *Synthese*, 177, 261–283.
- Sweatt, J. D. (2009). *Mechanisms of memory*. San Diego: Elsevier.
- Uttal, W. R. (2001). *The new phrenology*. Cambridge: MIT Press.
- Uttal, W. R. (2011). *Mind and brain: A critical appraisal of cognitive neuroscience*. Cambridge, MA: MIT Press.
- Uttal, W. R. (2013). *Reliability in cognitive neuroscience: A meta-meta-analysis*. Cambridge: MIT Press.
- Van Orden, G., & Paap, G. C. (1997). Functional neuroimages fail to discover pieces of mind in parts of the brain. *Philosophy of Science*, 64(S1), S85–S94.
- Woodward, J. (1989). Data and phenomena. *Synthese*, 79, 393–472.
- Woodward, J. (2000). Data, phenomena and reliability. *Philosophy of Science*, 67(3), S163–S179.