2

Real Reduction in Real Neuroscience: Metascience, Not Philosophy of Science (and Certainly Not Metaphysics!)

John Bickle

PHILOSOPHICAL ACCOUNTS OF SCIENTIFIC REDUCTION

Consider what is, among philosophers, an apparently unconventional argument (at least in the sense that few seem to act upon it). Suppose we wish to understand scientific reductionism—its nature, aims, scope, and potential limits. Here's a strategy: Let us find a clear example of a "reductionistic" field of scientific inquiry, dubbed so not only by its practitioners but also by scientists working in other, related fields. Then, as unencumbered by epistemological and metaphysical assumptions as we can rend ourselves, let us investigate some paradigmatic examples of recent research from that field, with our choice of examples dictated by the field's most prominent researchers. (These choices will result from discussions with those researchers, from publication in the field's most respected journals, from decisions by prominent funding agencies, and the like.) And then let us analyze the shared practices across these examples that differentiate this field from other scientific fields investigating related phenomena, only admittedly less reductionistically. (A good analogy here might be what a historian of a science does qua historian, only we'll be working with recent and current case studies.) The resulting account should be an analysis of real reductionism in real scientific practice, as contrasted with artificial accounts of scientific reductionism that rest instead on philosophical assumptions about "what reduction has to be".

So characterized, that project strikes me as inherently reasonable. So why is it virtually non-existent in contemporary philosophy? It is virtually non-existent therein. Two accounts of "reduction" dominate the philosophical literature. One, intertheoretic reduction, has its roots in late-20th-century philosophy of
Most detailed variations are responses to Ernest Nagel’s groundbreaking work in chapter 11 of his (1961) book, *The Structure of Science*. These accounts rest on strong assumptions about the structure of scientific theories, the nature of scientific explanation, and layered hierarchical pictures of both extra-theoretic reality and the sciences themselves. Reduction of one theory to another is either syntactic derivability or some weaker notion that approximates derivability in various respects. Proponents of intertheoretic reduction often cite scientific examples from the history of physics (e.g., the 19th-century reduction of portions of classical equilibrium thermodynamics to statistical mechanics and the kinetic/corpuscular theory of gases) and genetics (e.g., the mid-20th-century reduction of Mendelian principles of inheritance to initial discoveries of molecular genetics). Many surveys of intertheoretic reductionism exist in the literature (my own is in chapters 1 and 2 of Bickle 1998). Few proponents of intertheoretic reductionism have ever worried much about whether reductionistic scientific practices remain constant, not only across distinct sciences but also across the time periods that separate current scientific practice from those of half a century (as in the case of genetics) to more than a century ago (as in the case of the gas laws and statistical mechanics).

The other currently popular account of scientific reduction among philosophers and cognitive scientists is even more philosophically loaded and removed from current scientific practice. This is “functional reduction”, first championed by philosophers pursuing consciousness studies (Chalmers 1996, Levine 1993) and most recently by Jaegwon Kim (2005). According to this view, scientific reduction is a two stage process. First scientists “functionalize” the concept targeted for reduction by characterizing it exhaustively in terms of its causes and effects. Then they pursue normal empirical investigations to discover which mechanisms in the actual world play this causal role (or at least approximate playing it). The scientific examples described to illustrate this account are telling. They are not even examples from the history of real science (like the ones that intertheoretic reductionists at least appeal to). Rather, they are from elementary school science education—like the boiling of water near sea level due to the dynamics of H2O molecules! One might reasonably assume that the actual practices of current reductionistic science differ substantially from those involved in the examples we use to teach children the rudiments of our scientific worldview! Reductionists should also be struck by the fact that the original proponents of functional reduction are anti-reductionists about some features of qualitative consciousness. Here then is another methodological lesson for reductionists (that should be rather obvious): don’t let your opponents define the key concept of your account!

I won’t try to explain why these two accounts of reductionism have held such sway in contemporary philosophy of mind and cognitive science. That would require a long story about the extent to which armchair metaphysics and normative epistemology have re-injected “analytic” philosophy over the
past three decades. My goals in this chapter are of a more positive note. I am going to sketch a metascientific analysis of real reductionism in real, current neurobiological practice and point out some of its most obvious differences from the two accounts just characterized. This account stems from a more general analysis, first articulated by neurobiologist Alcino Silva, of the conditions on sufficient evidence for establishing a cellular or molecular mechanism for a cognitive phenomenon, and further developed in collaboration between the two of us (Silva and Bickle, forthcoming). I’ll provide one detailed recent case study (out of easily one hundred I could have chosen) that illustrates our analyses. And I’ll do all of this without imposing any significant metaphysical or normative epistemological assumptions that are not already a part of the scientific practices I seek to articulate.

‘MOLECULAR AND CELLULAR COGNITION’
AS A PARADIGMATIC CURRENT REDUCTIONISTIC SCIENTIFIC FIELD

The first step of any metascientific analysis is choosing the field from which to draw paradigmatic examples of scientific research that exhibit the concept one seeks to analyze. With regard to scientific reductionism, especially as it pertains to the relationship between cognition and the brain, the relatively new field of *molecular and cellular cognition* seems an apt choice. This field began in earnest in the early 1990s with the advent of using bioengineered genetic mutations in living, behaving rodents and the use of a variety of behavioral tasks that experimental psychology had developed for investigating specific cognitive phenomena. To date, learning and memory research has provided this field’s most impressive achievements, but work is now being pursued in more than 100 labs worldwide on virtually all the phenomena that comprise cognitive science. A professional society now exists, the *Molecular and Cellular Cognition Society* (or *MCCS*), whose web site (at www.molcellcog.org) states that a goal of the field is “to derive explanations of cognitive processes that integrate molecular, cellular, and behavioral mechanisms”. A related goal is to “promote the study of the molecular and cellular basis of cognitive function”. Molecular and cellular cognitivists set their field explicitly in contrast to cognitive neuroscience:

Unlike Cognitive Neuroscience, which historically has focused on the connection between human brain systems and behavior, the field of Molecular and Cellular Cognition studies how molecular (i.e. receptor, kinase activation), intra-cellular (i.e. dendritic processes), and inter-cellular processes (i.e. synaptic plasticity; network representations such as place fields) modulate animal models of cognitive function. (www.molcellcog.org)

Clearly, proponents of this field don’t hesitate to express their reductionistic attitude.
This attitude has even reached the latest neuroscience textbooks. In the introductory chapter of the most recent (4th) edition of their *Principles of Neural Science*, Eric Kandel, James Schwartz, and Thomas Jessell write:

This book . . . describes how neural science is attempting to link molecules to mind—how proteins responsible for the activities of individual nerve cells are related to the complexity of neural processes. Today it is possible to link the molecular dynamics of individual nerve cells to representations of perceptual and motor acts in the brain and to relate these internal mechanisms to observable behavior. (2001, 3–4)

These mind-to-molecular pathways “links” are reductions, at least in the sense that this concept is at work in actual current neuroscientific practice.

With a reductionistic scientific field in hand, our metascientific analysis next moves to finding the commonalities in scientific practices that unite investigations in this field and distinguish it from investigations of similar phenomena in less reductionistic fields. There are now hundreds of published experimental studies in molecular and cellular cognition to choose from. Space in this chapter limits me to a detailed presentation of only a single case. (I’ve presented others in recent publications, including Bickle 2003, chapters 2–4; 2005; 2006a; 2006b; forthcoming-b.) In the next section I’ll present a very recent example. In light of it, I’ll then present the Convergent Four principles of sufficient evidence for establishing a cellular or molecular mechanism for a cognitive phenomenon, and emphasize the two principles that constitute molecular and cellular cognition’s ruthlessly reductive core. I’ll then sketch the implicit account of real reductionism in really reductionistic neuroscience and contrast it with the two accounts popular in philosophy with which this chapter began.

**NEURONAL COMPETITION FOR PARTICIPATION IN A MEMORY TRACE IS DETERMINED BY RELATIVE CREB FUNCTION AT THE TIME OF TRAINING**

Electrophysiological studies in rodents have long suggested that only a small percentage of neurons in specific cortical regions encode a particular memory trace. For example, although roughly 80% of neurons in the lateral nucleus of the mouse amygdala receive sensory input during classical Pavlovian auditory fear conditioning, only about 20–30% display plasticity following the training phase. What factors determine which neurons are recruited to participate in a particular memory? Recent experiments by Sheena Josselyn, Alcino Silva, and their collaborators implicate as a key causal factor the functioning of gene expression transcription enhancer cyclic adenosine monophosphate (cAMP)/calcium responsive element binding protein, or CREB (especially the α and δ isoforms) in individual neurons at the time of training (Han et al. forthcoming).
Previous research has implicated CREB in the induction of late long-term potentiation (L-LTP), a form of activity-driven long-lasting (hours to days, even weeks) increased neurotransmission efficacy at individual chemical synapses. L-LTP requires new gene expression and protein synthesis. The multi-burst trains of activity in pre-synaptic axons necessary to induce L-LTP activate not only glutamatergic and N-methyl-D-aspartate (NMDA) receptors in the post-synaptic membrane, but also a class of dopaminergic post-synaptic receptors associated with a G-protein complex. This activity primes adenyl and adenylate cyclase molecules in the post-synaptic terminal to convert more adenosine triphosphate (ATP) molecules into cellular energy and cAMP. cAMP then serves as a second messenger, binding to regulatory subunits of protein kinase A (PKA) molecules and freeing up enough PKA catalytic subunits to translocate back to the neuron’s nucleus. There the PKA subunits phosphorylate CREB molecules, which in turn bind to cAMP responsive elements in the control region of both regulatory and effector genes, turning on new gene expression. The outcome is ultimately the synthesis of new proteins that are transported back to active synapses to restructure the cytoskeletons, keeping the synapses potentiated for hours to days (to weeks). Behaviorally, affecting these CREB-dependent mechanisms of L-LTP affects the consolidation of memories from labile, easily disrupted short-term to stable long-term form. Blocking any step in the cAMP-PKA-CREB process virtually eradicates memory consolidation, while enhancing steps can lead to faster and stronger consolidation. These basic effects have now been demonstrated experimentally for a large number of memory tasks, including hippocampus-dependent “declarative” or “explicit” memories.¹

Building on this experimental background, Josselyn, Silva, and their collaborators first showed that only around 20% of neurons in mouse lateral amygdala (LA) displayed CREB activation after auditory fear conditioning (Han et al. forthcoming). CREB activity was measured using a standard immunocytochemical antibody technique for labeling LA the presence of phosphorylated CREB (pCREB) in individual LA neurons. Wild-type mice (with no bioengineered genetic mutations) were divided into a tone + shock group (who underwent auditory fear conditioning in a training chamber and were exposed to the conditioning tone in a novel chamber during the testing phase 24 hours later) and a number of control groups (e.g., tone alone, immediate shock, chamber alone, and home cage groups). Mice in the tone + shock group showed roughly 20% pCREB-positive neurons in LA following the testing phase. No control group showed more than 10%.

¹ See Bickle 2003, ch. 2 for a nontechnical description of the basic molecular biology of LTP and some techniques for engineering specific genetic mutations in mammals (with extensive references to the primary scientific literature). See the other references cited in the last paragraph of the previous section for nontechnical discussions of some specific experimental results using these genetically mutated mice in memory research.
This first result is consistent with CREB functioning being a key factor in neuronal recruitment to participate in specific memory traces, since the number of pCREB positive neurons following auditory fear conditioning matches up well with the number of potentiated neurons found in previous electrophysiological studies. But can CREB function be shown to be involved more directly in neuronal competition during memory training? Josselyn, Silva, and colleagues established this more direct experimental connection by microinjecting replication-deficient herpes simplex virus (HSV) vectors fused with a gene expressing green fluorescent protein (GFP) and either the wild-type CREB gene for the $\alpha$ and $\delta$ transcription enhancer isoforms ($CREB^{WT}$) or a gene for a dominant-negative repressor form of the CREB protein that competes with endogenous CREB for binding sites in gene control regions but inhibits gene expression ($CREB_{133A}$), in which the serine (S) residue that occurs at site 133 in CREB $\alpha$ and $\delta$ isoforms has been replaced by an alanine (A) residue (Han et al. forthcoming). GFP makes infected neurons easy to image and count using standard microscopy techniques; infected neurons literally synthesize a protein that distributes throughout their cytoplasms and glows green in microscopic images. The $CREB^{WT}$ insertion increases the amount of CREB $\alpha$ and $\delta$ molecules available and enhances CREB functioning in infected neurons over normal endogenous levels. The $CREB_{133A}$ insertion reduces CREB transcription enhancer function. The details of this experimental work will not be familiar to philosophers and cognitive scientists, even for those who comment regularly on the scope and limits of neurobiology; but this is the sort of molecular biological knowledge and manipulation that is common in current molecular and cellular cognition.

Han et al. (forthcoming) first manipulated CREB expression in a population of genetically mutated mice with greatly reduced levels of CREB transcription enhancers (CREB$^{\alpha\delta-}$ mice). The gene for CREB $\alpha$ and $\delta$ isoforms had been “knocked out” at the embryonic stem cell development phase in these mice. Previous behavioral studies have shown that these mice display significantly deficient consolidation of short-term into long-term memory on a large number of tasks. For example, in auditory fear conditioning, they only spend about 20% of testing time freezing after exposure to the conditioned tone in the testing phase 24 hours after standard one-trial tone-shock pairings, as compared to about 60% freezing time in wild-type littermate controls. (Freezing is a stereotypic rodent fear response in which the animal crouches, tucks its front paws inward beneath its chest, and ceases all movement except breathing.) Interestingly, CREB$^{\alpha\delta-}$ mutants are intact compared to wild-type littermate controls on short-term versions of this and other memory tasks, where the delays between training and test phases range from 30 minutes to 2 hours. This common result controls for motivational, perceptual, attentional, and motor confounds. The CREB expression manipulation generates a specific memory consolidation effect.

Han et al. (forthcoming) microinjected HSV vectors containing genes for GFP and either $CREB^{WT}$ or LacZ (as a control vector) into lateral amygdala
(LA) of CREB$^{\alpha\delta-}$ mutants or wild-type littersmates prior to the training phase of auditory fear conditioning. Although the viral vector only infected around 18% of LA neurons in all groups (as measured by counting the number of LA neurons in confocal microscopic images displaying GFP compared to the number that did not), CREB$^{WT}$ injections completely rescued long-term auditory fear conditioning in CREB$^{\alpha\delta-}$ mutants. CREB$^{\alpha\delta-}$ mutants receiving the LacZ control vector displayed the usual failure to consolidate long-term auditory fear conditioning memories, freezing only about 20% of the time following exposure to the conditioned tone 24 hours after training, compared to the $>70\%$ freezing time in wild-types microinjected with either CREB$^{WT}$ or LacZ. Yet CREB$^{\alpha\delta-}$ mutants receiving microinjection of CREB$^{WT}$ froze about 75% of the testing time after exposure to the conditioned tone, statistically identical to wild-type performances. Furthermore, the rescued consolidation of long-term auditory fear conditioning did not result simply from the CREB$^{WT}$ injections in LA facilitating the freezing response. This was demonstrated by a supplemental study in which all groups were subjected to a hippocampus-dependent contextual fear conditioning task. In this task mice are exposed to a novel chamber, allowed to explore it briefly, and then shocked. They are placed back in the training chamber 24 hours later and measured for their freezing response. Both CREB$^{\alpha\delta-}$ mutant groups, those receiving LA injections of CREB$^{WT}$ vector and those receiving the control LacZ vector, showed the usual reduction in freezing time during re-exposure compared to both wild-type littermate control groups. This control result indicates that the LA CREB$^{WT}$ vector injections had no effect on memory consolidation deficits on hippocampus-dependent tasks and that the rescued consolidation in the auditory fear conditioning task was not due simply to facilitating the freezing response. So increasing CREB function in less than 20% of LA neurons completely rescues the consolidation of long-term auditory fear conditioning in CREB-deficient mice.

Interestingly, increasing CREB function in LA neurons of wild-type mice also enhanced auditory fear conditioning. Han et al. (forthcoming) showed this by using low intensity shocks (0.4 mA as compared to 0.7 mA used in the earlier study) that elicit a less-than-maximal freezing (fear) response. Wild-type mice receiving the LA LacZ control vector spent about 40% of the time freezing upon tone exposure 24 hours after training. Wild-type mice receiving the LA CREB$^{WT}$ vector spent about 75% of the time freezing upon tone exposure in the testing phase. This difference reflects a statistically significant increase in tone-shock association due to increased CREB availability over normal endogenous levels.

But can one show that the specific neurons infected by the CREB$^{WT}$ vector were actually the neurons preferentially recruited into the tone-shock association memory trace? To visualize the neurons that were components of the memory trace, Han et al. (forthcoming) took advantage of the unique time-course of the transcription of an activity-dependent gene, *activity-regulated*
cytoskeleton-associated protein (Arc). Increased activity in a given neuron induces a rapid, transient increase in Arc transcription, so that Arc RNA localized in the cell nucleus 5–15 minutes after neuron activity can serve as a molecular signal of recent activity (Guzowski et al. 1999). Han et al. (forthcoming) used a cellular imaging strategy, fluorescent in situ hybridization, to detect the specific LA neurons that were active (Arc+) during the testing phase of the auditory fear conditioning task. Only those neurons that were active during the testing phase, and thus part of the fear conditioning memory trace, would be Arc+. Inactive neurons during the testing phase, presumably not part of the memory trace, would be Arc−. The Arc images of LA neurons could then be merged with the GFP images to count the percentage of LA neurons that were double labeled (GFP+ and Arc+). Those neurons would be the ones that were both infected by the CREB<sup>WT</sup> vector (as evidenced by their being GFP+) and hence subject to increased CREB function, and also recruited into the fear conditioning memory trace (as evidenced by their also being Arc+).

If increased CREB function at the time of training is a critical factor that influences the probability that a given LA neuron is recruited as part of a fear conditioning memory trace, then GFP+ neurons with elevated CREB function induced by the CREB<sup>WT</sup> vector microinjections should have a greater likelihood of being Arc+ following the testing phase than their GFP− neighboring neurons that were not infected by the CREB<sup>WT</sup> vector. What were the percentages in the various experimental groups? In wild-type mice who received the CREB<sup>WT</sup> vector prior to auditory fear conditioning training, slightly more than 20% of all LA neurons were Arc+ during the testing phase (another result that coheres nicely with other measures described above about the percentage of LA neurons incorporated into a given memory trace). However, GFP+ LA neurons (which were infected with the CREB<sup>WT</sup> vector and thus had higher rates of CREB function at the time of training) were roughly 3 times more likely to be Arc+ than were neighboring GFP− neurons (which had endogenous CREB function at time of training). In wild-type mice infected with the LacZ control vector coupled with GFP, once again slightly more than 20% of all LA neurons were Arc+ during the testing phase of the auditory fear conditioning task. However GFP+ neurons (and hence infected with the control LacZ vector that does not affect CREB function) and neighboring GFP-neurons were equally likely to be Arc+. This effect was even more pronounced in the CREB<sup>αδ</sup>− mice, who are deficient in consolidating fear conditioning memories into long-term form, but whose deficit was rescued by LA CREB<sup>WT</sup> vector microinjections. In CREB<sup>αδ</sup>− mutants receiving the LacZ control vector, the percentage of Arc+ neurons during the testing phase was significantly lower than in any other group (less than 10% of all LA neurons), and GFP+ neurons (infected with the inactive control vector) were no more likely to be Arc+ than their nearby GFP− neighbors. However, in CREB<sup>αδ</sup>− mutants receiving the CREB<sup>WT</sup> vector, once again nearly 20% of all LA neurons were Arc+ during the testing phase. And GFP+ neurons
(infected with the CREB\textsuperscript{WT} vector and thus with increased CREB function at time of training) were roughly 10 times more likely to be Arc\textsuperscript{+} than their nearby GFP\textsuperscript{−} neighbors. (For the quantified data, see Han et al. forthcoming, figure 2.) These data directly support the hypothesis that neurons with higher CREB function at the time of training are more likely to be recruited into a specific memory trace than are those with normal or low CREB function. The mechanism of this competitive recruitment process into specific memory traces has now been reduced down to particular molecular processes in individual neurons. Modifying these processes in either direction has predictable behavioral effects on memory consolidation on tasks dependent on neurons in the region of the brain whose molecular processes have been manipulated.\textsuperscript{2}

This section has no doubt been rough sledding for many philosophers and cognitive scientists, so I’ll briefly summarize the points that will be emphasized in the metascientific analysis that follows.

- Numerous previous experiments had implicated CREB functioning in individual neurons as a mechanism of long-term memory consolidation, including in lateral amygdala (LA) neurons for auditory fear conditioning.
- Intervening to block CREB functioning using molecular-genetic techniques produces mice that cannot consolidate short-term fear associative memories into long-term form.
- Intervening to increase CREB functioning at time of training in less than 20\% of LA neurons (using HSV vector microinjection techniques) completely rescues long-term fear association memories in CREB-deficient mutant mice, and increases long-term fear memories in wild-type mice using a less-than-maximal aversive unconditioned stimulus.
- A fluorescent \textit{in situ} hybridization study reveals that individual LA neurons with increased CREB functioning at time of training are statistically much more likely to be recruited into the neuronal memory trace than neighboring neurons with normal endogenous or decreased CREB functioning.

THE CONVERGENT FOUR PRINCIPLES OF MOLECULAR AND CELLULAR COGNITION

Case studies like the one described in the previous section comprise the basis on which a purely metascientific account of real reductionism in actual scientific

\textsuperscript{2} In subsequent experiments, Josselyn, Silva, and their collaborators controlled for the possibility that neurons with increased CREB function simply have a lower threshold for inducing Arc (they don’t), and that inhibiting CREB function in roughly 20\% of LA neurons in wild-type mice (via HSV CREB\textsuperscript{S\textsubscript{A}} insertion) has no detrimental effects on memory consolidation in the auditory fear conditioning task (because enough LA neurons with relatively high CREB function remain available for recruitment into the memory trace). See Han et al. (forthcoming) for details on these control experiments.
practice can be generated and then assessed for philosophical significance. Based on a number of such cases, neurobiologist Alcino Silva was first to sketch (in unpublished writings) four principles that together amount to sufficient experimental evidence for establishing a cellular or molecular mechanism for a given “systems-level” cognitive phenomena, at least within the accepted practices of molecular and cellular cognition. Our recent collaborations have produced more detailed accounts of these Convergent Four (Silva and Bickle, forthcoming). These principles constitute metascientific fruits of an intimological investigation of molecular and cellular cognition—quite literally, the application of scientific practices to the study of scientific practice itself (Silva and Bickle, forthcoming). The account of real reductionism in actual reductionistic neuroscientific practice sketched in the next section derives directly from these principles.

**Principle 1: Observation.** Occurrences of the hypothesized mechanism are strongly correlated with occurrences of the behaviors used as experimental measures of the cognitive phenomenon. Experiments in many species and neural systems have documented the observation that learning is accompanied by changes in synaptic plasticity in the very brain regions required for that particular form of learning. Others have documented that maintenance of these synaptic changes are correlated with memory performance. Specific forms of synaptic plasticity, like late-phase long-term potentiation (L-LTP), have been correlated experimentally with memory performance in a variety of tasks. CREB function has been observed to be correlated with L-LTP. Meeting the Observation Principle does not by itself establish that the hypothesized mechanism is part of the causal nexus generating the behavioral measures. (Molecular and cellular cognitivists aren’t strict Humeans about causality!) But establishing these observations is often an early step in formulating the causal-mechanistic hypotheses that this field investigates experimentally. Before this Principle is met, investigators have no reason for pursuing the more detailed experiments required to establish sufficient evidence for a molecular mechanism for a cognitive phenomenon. In the case study discussed in the previous section, observation experiments had already long established that L-LTP followed from CREB function, that consolidation of long-term auditory fear conditioning (i.e., freezing during the testing phase) followed from L-LTP in lateral amygdala (LA) neurons, and more. (Indeed, even more than simple observation experiments already linked CREB function in LA neurons and long-term auditory fear conditioning prior to the study discussed above, as we’ll see in the discussion of Principle 4 below.) In addition to these previous results, Han et al. (forthcoming) began their investigations with an immunocytochemical antibody labeling study for pCREB that showed CREB functioning in roughly 20% of LA neurons during activation of long-term auditory fear conditioning memory (a result that matched earlier studies of plasticity in LA neurons
using electrophysiological techniques). This was a straightforward example of an observation experiment.

**Principle 2: Negative Alteration.** Intervening directly to decrease activity of the hypothesized mechanisms must reliably decrease the behaviors used as experimental measures of the cognitive phenomenon.

Experiments that establish negative alteration are often the centerpieces of current molecular and cellular cognition investigations. For example, the engineered genetic mutation ("knock-out") that produces the CREB$^{\alpha \delta -}$ mice used in the studies reported in the previous section is a negative alteration. The mutation decreases CREB function and experimenters then track reliable decreases in behaviors that measure specific types of memory consolidation (with the appropriate controls to rule out sensory, attentional, motivational, and motor confounds).

More specific genetic interventions, with the use of either selective promoter regions on genetic insertions that limit where genes of interest are expressed or inhibited, or the use of pharmacological tools that limit the temporal dimensions of the genetic manipulation, are often used to provide evidence of negative alteration. For example, Abel et al. (1997) coupled a transgene that overexpresses regulatory subunits of PKA molecules to a promoter region that binds $\alpha$ calmodulin kinase II. So while the transgene was present in every cell of the mouse’s body, it was only expressed in high amounts in forebrain neurons (including hippocampus). This enabled the experimenters to demonstrate a negative alteration on hippocampus-dependent memory tasks with these mice, but no significant alteration on amygdala-dependent tasks (where the transgene was expressed in lesser amounts). A second example is the CREB$^{\text{IR}}$ mouse, developed by Silva, Mashushige, and collaborators (Kida et al. 2002), in which an inducible CREB repressor fusion protein competes with endogenous CREB for CRE binding sites, but inhibits gene expression there. The CREB repressor protein has the usual alanine-for-serine residue change at position 133, but has been fused with a ligand binding domain from a human estrogen receptor that itself has been mutated to be activated by the drug tamoxifen (TAM). Hence the CREB repressor fusion protein is only activated, and hence only inhibits CREB function, when these mice have been injected with TAM; as soon as the injected TAM has been metabolized, CREB function returns to normal endogenous levels. This creates a 6–12 hour CREB negative alteration, inducible and reversible in the same mice, and enabled experimenters to demonstrate a transient loss of memory consolidation (and reconsolidation after reactivation) in mutated animals dosed with TAM just before training (Kida et al. 2002).

**Principle 3: Positive Alteration.** Intervening directly to increase activity of the hypothesized mechanisms must reliably increase the
behaviors used as experimental measures of the cognitive phenomenon. Although positive alterations of learning and memory have been carried out successfully in insect studies for more than a decade, cases are still few and far between in mammal studies. That is what makes the recent studies described above especially intriguing. Both the complete rescue of long-term auditory fear conditioning in CREB\textsuperscript{δ−} mutants and the enhanced effect in wild-type mice using low intensity training shocks following HSV CREB\textsuperscript{WT} microinjections into lateral amygdala (LA) are examples of positive alteration. In both cases, the inserted genetic material increased CREB function in roughly 20\% of LA neurons and reliably increased the measured freezing response to tone presentation during the testing phase of auditory fear conditioning. Techniques that generate evidence of positive alteration in mammals are genuine methodological breakthroughs in current molecular and cellular cognition.

**Principle 4: Integration.** The hypothesis that the proposed mechanisms are key components of the causal nexus that produces the behaviors used as experimental measures of the cognitive phenomenon must be connected up with as much experimental data as is available about both the hypothesized mechanism and the cognitive phenomenon.

Principle 4 is the most abstract of these conditions on sufficient evidence, and certainly the one requiring the most extensive explication.\(^3\) On the one hand, it serves to rule out silly objections to claimed mechanisms based on these conditions such as “removing oxygen from the animal’s environment significantly alters its behavior in this memory task. Is oxygen consumption thereby a mechanism of memory?” or “...Does memory thereby reduce to oxygen consumption?” (These are counterexamples that philosophers sometimes raise to the Convergent Four, attempting to be cute.) Clearly, the empirical background against which serious experimental studies are performed has already ruled out such silly mechanisms or reductions.\(^4\) Yet Principle 4 is intended to accomplish far more than just this. Data meeting it often provide the empirical reasons that motivate molecular and cellular cognitivists to attempt the specific negative and positive alteration experiments that they do, down to the particular gene expression and protein synthesis they manipulate (including the particular molecular-biological techniques they employ) and the behavioral measures they use to track the effects of their manipulations. A lot of information is usually

\(^3\) Silva and Bickle (forthcoming) is a first attempt to begin this explication.

\(^4\) Not to mention the fact that positive alterations into these silly “mechanisms” don’t produce significant effects on the behavioral measures used; or if they do, then the proposed “silly” mechanisms actually are key components of the causal nexus. This fact demonstrates the independence of Principle 3 from Principle 4 and the necessity of including Principle 3 in these conditions that are jointly sufficient for establishing a molecular or cellular mechanism for a cognitive phenomenon.
known about the molecular biology and the behaviors that molecular and cellular cognitivists combine in their negative and positive alteration studies, and this goes far beyond the observational correlations that fall under Principle 1 (Observation).

Another nice feature of the study described in the previous section for our metascientific purposes in this section is the illustration it provides of Principle 4 at work. Many studies, from the behavioral down to the molecular biological, had already implicated CREB function in lateral amygdala (LA) neurons as a molecular mechanism of long-term consolidation of auditory fear conditioning. But no previous study had integrated these findings to directly implicate CREB function as the molecular mechanism for the recruitment of individual LA neurons into specific memory traces. Josselyn, Silva, and their colleagues took advantage of another recent discovery from the molecular biology of neuronal activity, the activity-dependent and temporally limited availability of Arc RNA in neuron nuclei, and an in situ hybridization technique for measuring this signal of recent neuronal activity. This molecular-biological insight enabled them to merge images of CREBWT-infected LA neurons (GFP+), which had enhanced CREB function, with images of Arc+ neurons at the time of the testing phase of auditory fear conditioning. They were thus able to demonstrate a significantly higher probability of GFP+ LA neurons also being Arc+ than their nearby GFP− neighbors. Integrating this new molecular-biological knowledge and imaging techniques with already-available molecular and behavioral knowledge about CREB function and long-term fear conditioning consolidation provided the novel direct evidence that CREB functioning in individual LA neurons is indeed a causal mechanism of neuronal recruitment into circuits subserving specific fear memory traces.

Another interesting feature of the Integration Principle is a way that prior experimental work gets incorporated into ongoing research. In current molecular and cellular cognition research, most of the time experimental work already exists that suggests a key mechanism for the cognitive phenomenon at issue; and typically this earlier work itself already meets most of the Convergent Four principles. Often the causal interventions used in this previous work have taken place at higher “levels” of biological organization than the new experiments being pursued. This is what philosophers and cognitive scientists typically refer to as “relating different levels” of theory and explanation. In the study discussed in the previous section, earlier experiments had already established that neurons in the lateral amygdala (LA) were anatomically connected to the motor pathways that generate the behavioral measures of auditory fear conditioning, and to the sensory inputs from auditory cortex. Other work (by Joseph LeDoux, James McGaugh, and others) had established that L-LTP takes place in LA neurons during auditory fear conditioning and that CREB functioning occurs during the neuronal plasticity that resulted from the training phase of the task. These connections had already been established as more than mere observed
correlations: alteration experiments had been performed successfully⁵ and theoretical integration had been proposed. In fact, these connections had already been established down to the number of LA neurons that received auditory input and the number that underwent plasticity in response to the tone-shock pairing. The “higher-level” results connecting CREB to LTP, LTP to LA neuron plasticity, and LA neuron activity to long term auditory fear conditioning thus became part of the integrative theoretical background for this study described in the section above. The Han et al. (forthcoming) study in turn established a positive alteration of CREB functioning to increase auditory fear conditioning behaviors, and CREB functioning as the mechanism for the recruitment of specific LA neurons into the circuits for particular memory traces. In this way Principle 4 captures how new results build in prior ones—who the prior ones themselves met at least some of the Convergent Four principles on their own.

The study described in the previous section is just one of at least one hundred others that could be cited as providing experimental illustrations of the Convergent Four Principles. Alcino Silva and I offer the Convergent Four as our first metascientific hypothesis based upon an intimological investigation of the scientific practices of molecular and cellular cognition (Silva and Bickle forthcoming). I’ll close this essay in the final section by sketching a second metascientific hypothesis: the nature of reductionism in the actual practices of this reductionistic branch of contemporary neuroscience, drawn from the core of the Convergent Four Principles.

**THE RUTHLESSLY REDUCTIVE CORE OF THE CONVERGENT FOUR**

Notice that Principles 1 and 4 require “higher level” scientific investigations.⁶ To establish the required observations between hypothesized mechanism and behaviors, and to integrate knowledge of molecules and behavior to establish the theoretical plausibility of the proposed mechanisms for the cognitive phenomenon in question, we need precise knowledge of what the system does under controlled experimental conditions. This means having both precise data about the system’s behaviors (as grist for our lower level mechanistic explanations)

⁵ With the possible exception of successful positive alteration experiments linking CREB function in LA neurons and auditory fear conditioning, which the Han et al. study also provided.

⁶ I enclose “higher level” in scare quotes to indicate that very little hangs on its explication (here or in the previous discussion of Principle 4). I don’t assume anything fancy by this term and nothing in my argument relies on any detailed account of “levels”. Here I simply refer to the common assumption in neuroscientific practice that locates appeals to neural systems at a higher level than appeals to the cellular physiology of its component neurons, and the latter at a higher level than the molecular-biological processes that take place around and inside of their membranes.
and good behavioral measures for the cognitive phenomenon at issue. These are jobs for cognitive scientists and experimental psychologists, not electrophysiologists or molecular geneticists. We also need to know where to start inserting our cellular and molecular interventions. The “decomposition and localization” investigations of cognitive neuroscientists are crucial for this knowledge.\(^7\) We also need to know what types of neuronal activity to intervene into. Action potential frequency? Action potential dynamics? Field potentials? Something else entirely? The work of neurocomputational modelers and simulators is important here. Each of these activities has distinct molecular mechanisms, and so requires different molecular-biological techniques to intervene into. Molecular and cellular cognition needs a lot of higher level cognitive science and neuroscience to accomplish its potential reductions—and it now regularly draws upon such scientists in order to get these details right. Molecular and cellular cognition is a reductionistic brand of current neuroscience, perhaps even “ruthlessly” so. But that in no way precludes its use of higher level cognitive science and scientists.

Yet in the end (at least at the present time), it is the experiments that illustrate Principles 2 and 3 that cinch the empirical case for a proposed lower level mechanism for a cognitive phenomena. It is certainly these experiments that typically constitute the unique contributions of molecular and cellular cognition studies. Even the case study described two sections ago, which in the previous section I argued made a significant contribution to Principle 4 for establishing the connection between CREB function and auditory fear memory consolidation, made an equally important contribution to Principle 3. That is, it established a positive alteration in CREB function to enhance the behavioral measures of auditory fear memory consolidation.

What then is the nature of the reductionism implicit in Principles 2 and 3? Unlike classic intertheoretic reduction, real reductionism in molecular and cellular cognition does not require an explicit, complete set of laws or explanatory generalizations that characterize the behaviors of reduced and reducing kinds in all contexts or circumstances. Reduction is not a logical relationship between such laws or generalizations. Unlike more recently developed and championed “functional” reduction, real reductionism does not require the reduced concepts to be characterized exhaustively in terms of their causes and effects; instead, it requires cognitive concepts to be operationalized methodologically, in terms ultimately of measures in specific behavioral protocols and paradigms, for the purposes of controlled experiments. In other words, instead of logical derivation of laws or explanatory generalizations, or functionalization of concepts, real reductionism in genuinely reductionistic neuroscientific practice is a matter of:

*Intervening causally, directly into processes at increasing lower levels of biological organization (cellular, intra-cellular molecular, molecular genetic)*

\(^7\) Bechtel and Richardson (1992) remains the most useful discussion of this strategy.
and then

*Tracking* the effects of these interventions in living, behaving organisms using a variety of measures widely accepted as indicative for the cognitive phenomenon being investigated.

When these interventions generate evidence for negative and positive alterations (Principles 2 and 3 above) in light of background evidence connecting the hypothesized cellular or molecular mechanism to the behaviors serving as measures for the cognitive phenomenon, a reduction is said to have been accomplished. The cognitive phenomenon reduces to the cellular or molecular mechanisms intervened into, within the anatomical circuits leading ultimately to the motor peripheries generating the measured behaviors. Molecular and cellular cognitivists then talk explicitly about having discovered “a molecular biology of cognition” (Bailey, Bartsch, and Kandel 1996), or of “providing an experimental framework for a molecular understanding” of a specific cognitive phenomena (Abel et al. 1997).⁸ Note that these are assertions made within—following Rudolph Carnap, we might say “internal to”—the practices of molecular and cellular cognition.⁹ They are neither metaphysical identity claims nor normative epistemological assertions.

Meeting Principles 2 and 3 establishes experimentally that hypothesized mechanisms are actually doing the job. When Principles 2 and 3 are established experimentally (and not simply predicted to be established), our best causal-mechanistic story for the cognitive phenomenon in question now resides at the lowest level of effective interventions (in conjunction with the anatomical circuitry that gets the neuronal activity out to the motor periphery). Once again, this assertion is a part of a metascience of molecular and cellular cognition—a description of the assertive practices that take place in the research of that field. It is not offered here as a metaphysical or normative epistemological claim. It is implicit in the ways that molecular and cellular cognitivists discuss their results and develop their experimental strategies (Bickle 2003, 2005, 2006a, forthcoming-a, forthcoming-b; Silva and Bickle forthcoming). From the perspective of molecular and cellular cognition, when all of the Convergent Four Principles are met for a hypothesized mechanism and a cognitive phenomenon, “higher level” explanations of that phenomenon lose their status as causal-mechanistic—although such explanations still provide empirical support for two of the four principles that our currently best causal-mechanistic explanation rests upon. If you ask higher-level explanations to provide more than this,

⁸ See Bickle (forthcoming-a) for more discussion of the differences between intertheoretic, functional, and metascientific reductionisms. Bickle (2006a) provides diagrams that illustrate contrasts between intertheoretic and metascientific reductionisms and a first try at distinguishing metascientific reductionism from the recent “new mechanical” philosophy of science.

⁹ See Bickle (2003, chapter 1), for more on the connection between Carnap’s internal/external existence question distinction and the brand of metascience advocated here.
however, after the Convergent Four principles have been met for a cellular or molecular mechanism, then you are asking for something beyond the role that the “ruthlessly reductive” practices of molecular and cellular cognition ascribe to them. That isn’t necessarily a mistake. We don’t yet know the explanatory scope of molecular and cellular cognition (although it already extends way beyond the range that most philosophers and cognitive scientists realize—see my publications cited below from Bickle 2003 onward). But your account thereby also isn’t “neurobiologically plausible”, at least in light of the practices and results of molecular and cellular cognition circa today. And that is a field of neuroscience whose practitioners increasingly populate publications in the best scientific journals, procure the largest share of external grants, and get awarded the most prestigious prizes.

REFERENCES


