Matthew R. Johnson

Model systems of thought: A neuroscience perspective on cognitive framework

One of the core issues in psychology and neuroscience research is that the nervous systems and behavior of human beings are complex, yet laboratory experiments must be kept relatively simple in order to be well-controlled and thus provide definitive answers to research questions. In this chapter, I first discuss how "model organisms" are used to reduce the complexity of scientific investigations in low-level (cellular) neuroscience research; that is, even if researchers are ultimately interested in the human mind and brain, it is often easier to begin by asking reduced forms of their questions about animals with simpler nervous systems, and attempt to build towards an understanding of increasingly complex systems. A similar approach can be taken in psychological studies of human beings. It is impossible to simultaneously examine all the multitude of factors that drive human behavior, so instead we must study individual facets of human behavior in the laboratory, and hope to build towards a more unified understanding. Here, I argue for a "component process" approach to studying human thought, wherein we use extremely simple laboratory tasks in an effort to identify fundamental "building blocks" of cognition that may form the basis of more complex thoughts and behaviors.

Model systems

As a scientist, one of my pet peeves is when media pundits or public figures take scientific investigations to task without properly understanding the context of the research. For example, in my home country (the

United States), one might hear political candidates denouncing "excessive" government spending by saying things like, "Last year we spent millions of taxpayer dollars on research into microscopic roundworms! Why are we wasting money on such unnecessary experiments?"

Of course, it's true that some scientific studies are better and more important than others, but sometimes the types of experiments cited in the media as "ridiculous" or "unnecessary" are actually quite important. For instance, the microscopic roundworm *Caenorhabditis elegans* is one of the primary organisms used to study genetics and cellular biology. *C. elegans*, as it is called in the community, is extremely hardy, is an excellent candidate for genetic manipulations, and grows from egg to adulthood in only a few days. It has just 959 cells in its body, but 302 of those are neurons. Thus, it is infinitely easier (and more ethical!) to manipulate and study neural processes at the cellular level using *C. elegans* than using human beings.

Fortunately, many of the genes, proteins, structures, and other aspects of *C. elegans'* cellular neurobiology are either identical or at least homologous (similar) to corresponding aspects of the human nervous system. So it's not entirely accurate if a pundit denounces spending all that money "just to study roundworms." By studying roundworms, we are actually gaining knowledge about fundamental biological processes that apply to humans and many other species – not just roundworms. And we are gaining that knowledge while spending far less time and money than if we studied those processes in human beings directly.

We call species like *C. elegans* "model organisms" or "model systems" of human biology because, just like other types of models (for example, a model of a building or an airplane), they are a smaller or more simplified, but still generally accurate, representation of a larger or more complex thing. Just as an architect might build a model of a skyscraper to work out design problems before constructing the real one, biologists work out many fundamental principles in model organisms before attempting to apply those principles to solve more complex problems (for example, developing new medications) in human biology.

C. elegans are far from the only model organisms used to study neuroscience and behavior. Our basic understanding of how a signal travels through a nerve cell comes from experiments on squid neurons.

Studies of the sea slug *Aplysia californica* have contributed greatly to our knowledge of how memory operates on the cellular level. And the research that forms the foundation of our understanding of the mammalian visual system came from experiments on ordinary house cats. (All three of those bodies of research resulted in Nobel Prizes, by the way.)

Sometimes model organisms are chosen because they exhibit specific biological features that make them convenient to study. For example, certain squid neurons have a very large axon (the part of a neuron that conducts a signal towards other neurons) called, appropriately enough, the squid giant axon. The squid giant axon can be several hundred times thicker than the typical axon found in humans or other animals, making it significantly easier to pierce with a recording electrode in a laboratory experiment.

But aside from specific features like these, a good rule of thumb is that scientists try to use the simplest animal possible that exhibits the biological trait or behavior they want to study. A good scientist hates unnecessary complexity – he or she tries to study an experimental question in the purest form possible, while trying to avoid nuisance factors that might complicate interpretations of the data or make it more difficult to perform a well-controlled experiment. Anecdotally, Eric Kandel (who won the Nobel Prize for studies of memory in *Aplysia*) began his academic career intending to become a psychoanalyst of human beings, then joined a laboratory studying neural communication in systems of thousands of mammalian neurons, and eventually decided that his interest in the biological machinery of learning and memory would be best served by studying the even simpler nervous system of the modest sea slug (Kandel, 2001).

Of course, the complexity of the experimental question determines, to some degree, what will be the optimal laboratory setup and model organism for addressing that question. For simple traits or behaviors, a roundworm or a fruit fly may suffice. Other areas of research – say, into drug addiction or aging-related dementia – are associated with the more complex brains of mammals, but still, a mouse or rat can often work nicely. For still other questions – say, how the brains of more intelligent mammals can recognize specific faces or objects – we might need to record signals from the brains of small monkeys.

Now, I myself am a cognitive neuroscientist – meaning I use safe, non-invasive neural recording techniques like functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) to study the brains and behavior of human beings. For all the sometimes frustrating complexities of our nervous systems and behavior patterns, these same complexities also make humans much more interesting subjects of study (at least, to me) than animals with simpler brains. There are logistical advantages to studying humans as well, foremost being the use of language. Although other animals can communicate with each other in various ways, no other species uses a full-featured language that allows it to express complex ideas with anywhere near the efficiency of human language. Thus, much like the squid giant axon is convenient for studies of neural transmission, the unique human facility for language actually makes us an excellent "model system" for studying certain neural or mental processes.

For example, say that I want to investigate how the visual system distinguishes between different individuals' faces, versus two different pictures of the same individual. To examine this question, I could make some kind of neural recording while subjects perform a face-identification task: On each trial, I show a subject two face pictures, and the subject presses a button to indicate whether he/she thinks the two pictures are of the same or different individuals. Later on, I'll analyze the neural recording data to look for differences between trials where the subject (correctly) judges that the faces are different and trials where they are judged to be the same. (It might be interesting to analyze error trials too, but let's forget about those for now.)

Now, you could certainly teach a monkey, and perhaps other species, to perform this task. But it could take weeks or months to train an animal to do it correctly, whereas a human subject could be given adequate instructions in under a minute. Assuming the neural recording techniques we can ethically use on human beings are sufficient for our purposes, it may be easiest to use human rather than animal subjects in this example.

Operationalization: Putting theoretical constructs into an experimental setting

I have just told you that my own research focuses on human beings, and that humans can be the most convenient research subjects for answering certain questions. So you might justifiably wonder why I spent one-third of this essay discussing advantages of working with model organisms. The main reason is that many of the points made about animal models can also apply to the way we study mental processes in humans. As stated above, we don't necessarily study *C. elegans* because we are particularly interested in roundworms; instead, we study them because we are interested in fundamental questions about how things like genes and proteins work, and *C. elegans* provides a simplified environment in which to study those biological concepts.

The corresponding point regarding psychology and cognitive neuroscience research is that the lofty big-picture questions that fascinate us may not always be apparent from a straightforward description of the experiments we perform. Whether a contemporary psychologist wants to study how we fall in love, what makes some people more intelligent than others, how we perceive the cues that let us perform social interactions, how language works, or any of a thousand other intriguing and complex questions, he/she may address that question with what seems like a relatively boring and simple experiment: Participants sit in a small, plain room with a single personal computer, viewing pictures or videos on the screen, performing some simple cognitive task, and responding by pressing buttons. Just like *C. elegans* gives us a simplified biological model of human genetics, tasks in psychology experiments give us simplified cognitive models of human thought. And just like the connection between sea slugs and human memory, the connection between a button-pressing task and falling in love may not always be apparent to the untrained eye.

This means that psychologists can discuss our research on two levels: The overarching questions that drive our work, or the nuts and bolts of the experiments we conduct and the data we collect. (And there can be various levels of theory connecting these two extremes.) In other words, simple cognitive tasks and laboratory experiments let us *operationalize* – define and measure in a straightforward way – more complex or general theoretical concepts. As professional scientists speaking amongst ourselves, we are accustomed to making the connections between these levels and inferring the big-picture context of an experiment automatically, but I think we sometimes forget about making such connections more explicit when addressing general audiences. So, with the remainder of this chapter, I'd like to introduce some of the big questions my research group studies, as well as some of the experiments we have designed to operationalize and answer those questions.

The building blocks of thought

Simply put, my research focuses on one primary question: How do people think? Now, this might be the most general question in all of psychology, and in a sense, all psychologists study this question. However, many researchers focus on particular domains: How does emotion work? How does vision work? How does long-term memory work? How does a particular everyday task (such as reading or driving) work?

However, my research (much of which has been done in collaboration with my former PhD supervisor, Marcia K. Johnson, as well as others) concentrates primarily on thinking in the general sense – what processes combine to create the ongoing stream of conscious thought that we experience most of our waking hours, every day of our lives. I want to know what constitutes a thought, how we shift focus from one mental representation to another, how we synthesize information to create novel ideas, and what compels our attention to move on after thinking about something for a while, rather than dwelling on the same thought forever.

At first, it may not even be evident that these questions need to be asked. Thinking is something that "just works," right? Does it really need to be explained further? But like anything that "just works" – think of your car, or your smartphone – a lot of engineering may go into disguising the true complexity of the underlying machinery, in order to present a system whose operation appears straightforward and seamless on the outside. In other words, precisely because thinking feels so effortless, it may disguise the fact that, without your conscious knowledge, your brain is actually performing many complex operations every second just to keep your train of thought chugging along smoothly.

Of course, studying everyday thought in its usual form is not easy – people's normal train of thought is too rapid, too chaotic, and too difficult to track in a laboratory setting. Instead, we need to study thought in a "model system" that retains the core operations we are interested in, while eliminating unnecessary complications. To help accomplish this, we take what we call a "component process" approach – we try to break up complex mental processes into smaller components that we can study in isolation.

For example, imagine you're making plans for dinner tonight. You could probably break this thought process up into several simpler sub-processes – for example, retrieving information from long-term memory (where you will be tonight, how much money you have, some foods you like to eat), sorting through lists of options retrieved from memory (what food items you have at home, what restaurants will be nearby), and evaluating the pros and cons of each choice with regard to your goals/desires (what will taste best, what is quickest, what will cost the least, what is healthiest). Each of these sub-tasks could be broken down into even more basic processes as well.

When you plan dinner each day, your brain easily and automatically manages dozens of such tasks and sub-tasks, but in the laboratory, we generally want to isolate one or two simple processes at a time to figure out how they work. One process we have studied extensively so far is called *refreshing*, using what we call a "refresh task" (for review: M. K. Johnson et al., 2005).

At any given time, you probably have several different representations active in short-term memory that are related to whatever you're thinking about. Refreshing is the process of shifting your internal mental attention among those items. So if you're thinking about potential dinner options you have at home (chicken, pasta, salad…), after you have retrieved the list from long-term memory, you will likely refresh – or shift your mental spotlight onto – each option as you evaluate it.

In the laboratory, we try to control this process by explicitly telling people which items to think about, rather than letting their thoughts flow freely. In a typical refresh task (Figure 1), we might first present two items on the computer screen (usually either a pair of pictures or a pair of words) for a brief period of time (about 1500ms). The study participant will view these items and encode them into short-term memory. Then, after a short delay (about 500ms), we show the participant an arrow pointing to the location where one of those items was just presented. The arrow is the participant's cue to briefly (for about another 1500ms) turn his/her mental attention to the item that was presented in the cued location, and not to think about the other item. (For studies using picture stimuli, we typically tell participants to briefly visualize the item in their mind; for studies using word stimuli, we typically ask them to say it aloud. Generally, though, we obtain fairly similar results no matter what type of stimuli we use.) It's a very simple (and slightly boring) task, but it allows us to study this particular component of everyday thought in isolation, without the complications that tend to accompany more elaborate tasks.

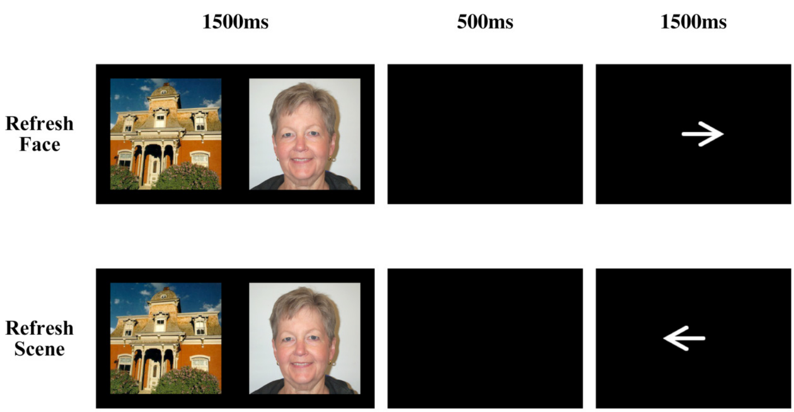

Figure 1: Structure of a typical refresh task used in functional MRI experiments.

When we scan people's brains using fMRI while they perform a refresh task, we typically observe activity in several brain regions, most notably an area in the front-left portion of the brain known as the *dorsolateral prefrontal cortex* (DLPFC). More specifically, there is greater activity in DLPFC when people do a refresh task than when they perform other tasks like passively viewing pictures. This suggests that

DLPFC is a key area involved in shifting mental attention from one item to another. Furthermore, as studies frequently also observe DLPFC activity in much more complicated mental tasks involving extensive planning or manipulation of many items in short-term memory (tasks collectively called *executive functions*), we could theorize that some of that activity occurs because refreshing is a sub-component of those complex tasks.

In one series of studies (M. R. Johnson, Mitchell, Raye, D'Esposito, & M. K. Johnson, 2007; M. R. Johnson & M. K. Johnson, 2009), we scanned participants' brains with fMRI while they performed a refresh task in which the initial stimuli were a picture of a face and a visual scene, as in Figure 1. Note that on all trials, participants saw the exact same visual display: First a screen with one face and one scene picture, then an arrow. However, depending on which sides of the screen the face and scene were shown on, and which way the arrow pointed, the arrow cue would indicate to participants either to think about (refresh) the face or the scene. It's important that the actual visual items shown onscreen (face, scene, arrow) were identical regardless of which item participants refreshed, because the critical differences in brain activity were in visual brain areas. When participants turned their mental attention to the face (and ignored the scene), we observed more activity in brain areas associated with visual processing of faces. When they instead turned their attention to the scene (and ignored the face), we observed more activity in areas associated with visual processing of scenes. However, we observed DLPFC activity whenever participants refreshed an item, regardless of whether it was a face or a scene.

This suggests a more refined picture of how refreshing operates in the brain: Perhaps DLPFC is the area that generates the initial neural signal to shift our mental attention to a certain item, regardless of what the item is. This signal then *modulates* (affects) activity in visual areas corresponding to the refreshed item. In other words, when the instruction "visualize the scene you just saw" is given, the DLPFC instructs scene-processing brain areas to replay the neural activity they experienced when the scene was first viewed. This model may be somewhat oversimplified, but it is generally consistent with the results we observe in our experiments, as well as others' interpretations of how these processes and brain regions operate.

Here's another experiment we ran recently (M. R. Johnson et al., 2013). This one is purely behavioral – i.e., we measured no brain activity, and relied only on reaction times (RTs). In this study (Figure 2), we presented participants with two everyday words, followed by an arrow cue instructing them to refresh (say aloud) one of the words. So far, this is very similar to previous refresh tasks. However, in this study, we added something: Immediately after participants refreshed a word, we printed either that word (the *refreshed* item) or the other word that was initially presented, but not refreshed (the *unrefreshed* item) again on the screen. When that final word was presented, participants were instructed to read it aloud as quickly as possible, and we measured their RTs to say it.

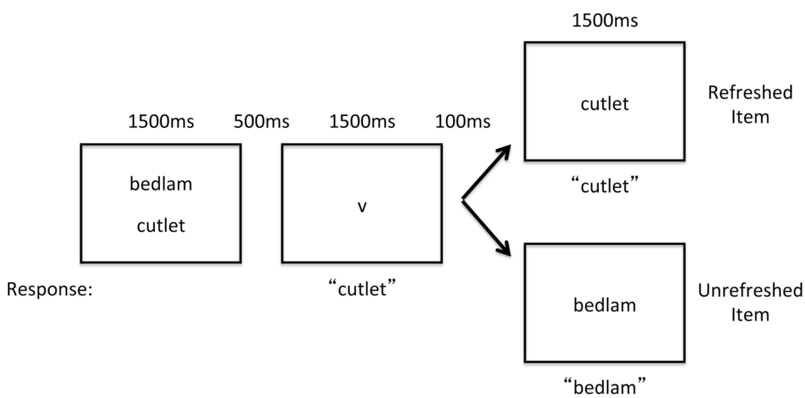

Figure 2: Structure of a refresh task used in a recent cognitive psychology experiment.

Before running the study, we expected participants to be faster at saying the refreshed item than the unrefreshed item, because we thought their mental representations of the refreshed item would be more active and thus more accessible. In fact, we found the exact opposite – they were significantly faster to say the unrefreshed item. (We have also replicated this result using pictures in a modified version of the task.) Although this was initially surprising, we now suspect that we have discovered the mental-attention version of a well-known visual attention effect called *inhibition of return* (IOR). Briefly, IOR is the finding that when participants are viewing (and shifting attention to) different items on a visual display, they are slower to return attention to a recently-visited location than to shift it to an unvisited location (Posner & Cohen, 1984; Posner, Rafal, Choate, & Vaughan, 1985). Some researchers interpret this as an aspect of our visual system's design that facilitates *foraging* (Klein, 2000). In other words, when viewing a scene, it usually makes sense to move your eyes around to different parts and explore the whole thing, rather than examining the same location over and over.

We will need more experiments to validate the following conjecture, but it is exciting to consider whether our mental attention system – whatever moves the "spotlight" of consciousness from one thought to another – is designed for foraging as well. Put another way, perhaps these studies will help reveal a mental mechanism responsible for our having a "stream" of consciousness rather than a "lake" – explaining why our thoughts tend to flow smoothly toward new possibilities rather than get stuck forever in a single place. And perhaps something about this mechanism is disrupted in conditions like autism, where patients may focus obsessively on one object or activity, or attention deficit hyperactivity disorder, where patients may shift attention away from the current train of thought too easily, and thus have difficulty staying on task.

These possibilities, and many more, remain to be explored. Perhaps future experiments will reject these hypotheses and suggest new ones. Perhaps these conjectures about the nature of thought are too broad to be justified by a few experiments on refreshing words or pictures. Perhaps these laboratory tasks represent an oversimplified or inaccurate model of natural thought processes. At least, however, the "model system" approach allows us to begin with well-controlled experiments and clear-cut results, and use these as the fundamental building blocks of more complex – and complete – theories of human thought.

References

Johnson, M. K., Raye, C. L., Mitchell, K. J., Greene, E. J., Cunningham, W. A., & Sanislow, C. A. (2005). Using fMRI to investigate a component process of reflection: prefrontal correlates of refreshing a just-activated representation. *Cognitive, Affective & Behavioral Neuroscience*, 5(3), 339–61.

Johnson, M. R., Higgins, J. A., Norman, K. A., Sederberg, P. B., Smith, T. A., & Johnson, M. K. (2013). Foraging for thought: an inhibition of return-like effect resulting from directing attention within working memory. *Psychological Science, 24,* 1004–12.

Johnson, M. R., & Johnson, M. K. (2009). Top-down enhancement and suppression of activity in category-selective extrastriate cortex from an act of reflective attention. *Journal of Cognitive Neuroscience*, 21(12), 2320–7. doi:10.1162/jocn.2008.21183.

Johnson, M. R., Mitchell, K. J., Raye, C. L., D'Esposito, M., & Johnson, M. K. (2007). A brief thought can modulate activity in extrastriate visual areas: Top-down effects of refreshing just-seen visual stimuli. *NeuroImage*, 37(1), 290–9. doi:10.1016/j.neuroimage.2007.05.017.

Kandel, E. R. (2001). Eric R. Kandel – Nobel Prize Autobiography. Retrieved September 4, 2012, from <http://www.nobelprize.org/nobel_prizes/medicine/laureates/2000/kandel-autobio.html>.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, 4(4), 138–147. doi:10.1016/S1364-6613(00)01452-2.

Posner, M., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. Bouwhuis (Eds.), *Attention and performance Vol. X* (pp. 531–556). Hillsdale, NJ: Erlbaum.

Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology*, 2(3), 211–228. doi:10.1080/02643298508252866.