Content

- The Brain-Mind Problem: the philosophical approach
- Capsule History of Psychology
- Representation and Computation
- The Cognitive Revolution
- The Physical Symbol Systems Hypothesis
- Neural Networks: Biological and Artificial
- Cognitive Neuroscience: Multiple Levels - Multiple Methods
- Multiple Memory Systems
- Language acquisition, evolution and processing
- Embodied Cognition - Cognitive Robots
- Consciousness:
- Integrating Emotions and Cognitions.
The Brain-Mind Problem

- monism vs. dualism
- reductionism
- emergentism
- functionalism
- downward causation

<table>
<thead>
<tr>
<th>Monism:</th>
</tr>
</thead>
<tbody>
<tr>
<td>is the theory that there is only one fundamental kind, category of thing or principle.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dualism:</th>
</tr>
</thead>
<tbody>
<tr>
<td>is the theory that the mental and the physical or mind and body or mind and brain are, in some sense, radically different kinds of thing. (Interactionist dualism from Descartes to Popper and Eccles)</td>
</tr>
</tbody>
</table>
The Brain-Mind Problem: the philosophical approach

**Emergentism**

The properties of complex systems are **not reducible** to those of their constituent elements, though they could not exist without them. While many of the fundamental properties of matter, such as mass, are held to be merely quantitative and additive, emergent properties are said to be qualitative and novel or non-predictable.

---

**Reductionism**

Complex things can always be reduced to (explained by) simpler or more fundamental things. 

Ontological reductionism

\(<\rightarrow\) monism

Denial of reductionist ideas is holism
The Brain-Mind Problem

Functionalism
Functionalism is the doctrine that what makes something a thought, desire, pain (or any other type of mental state) depends not on its internal constitution, but solely on its function, or the role it plays, in the cognitive system of which it is a part. More precisely, functionalist theories take the identity of a mental state to be determined by its causal relations to sensory stimulations, other mental states, and behavior.

Downward Causation
all processes at the lower level of a hierarchy are restrained by and act in conformity to the laws of the higher level
Specifically: mental agents can influence the neural functioning

"Two way causation"
The nervous system can be considered as being open to various kinds of information, and that there would be no valid scientific reason to deny the existence of downward causation, or more precisely, a two-way causal relationship between brain and mind (Szentágothai)
Capsule History of Psychology

- Weber and Fechner
- Helmholtz
- Ebbinghaus
- James, William
- Freud
- Wundt
- Skinner
- Luria
- Vigotsky
- Piaget
- (Gestalt)
- Hebb
- Skinner

- a scientific study of human behavior and mental processes
- first: physicians who started experimenting with behavior through scientific methods
- memory experiments
- first experimental laboratory: Leipzig, Germany
- behaviorism
- psychoanalysis
- constructivism
Knowledge Representation

- formal logic
- rules
- concepts
- analogies
- images
- connections
The cognitive revolution: a historical perspective

George A. Miller

Cognitive science is a child of the 1950s, the product of a time when psychology, anthropology and linguistics were redefining themselves and computer science and neuroscience as disciplines were coming into existence. Psychology could not participate in the cognitive revolution until it had freed itself from behaviorism, thus restoring cognition to scientific respectability. By then, it was becoming clear in several disciplines that the solution to some of their problems depended crucially on solving problems traditionally allocated to other disciplines. Collaboration was called for: this is a personal account of how it came about.

Anybody can make history. Only a great man can write it.

Oscar Wilde’s aphorism is appropriate. At the time, the suggestion that we were making history would have been presumptuous. But anybody can make history; writing history is another matter. I know something of the scholarship required and nothing approaching it has gone into the story I will tell here. But I offer this personal account in the hope that it might interest and help the real historians of science.

At the time it was happening I did not realize that I was, in fact, a revolutionary, and two different stories became intertwined in my life. They unfolded concurrently but I will tell the psychological story first.

The cognitive revolution in psychology

The cognitive revolution in psychology was a counter-revolution. The first revolution occurred much earlier when a group of experimental psychologists, influenced by Pavlov and other physiologists, proposed to redefine psychology as the science of behavior. They argued that mental events are not publicly observable. The only objective evidence available is, and must be, behavioral. By changing the subject to the study of behavior, psychology could become an objective science based on scientific laws of behavior.

The behavioral revolution transformed experimental psychology in the US. Perception became discrimination, memory became learning, language became verbal behavior, intelligence became what intelligence tests test. By the time I went to graduate school at Harvard in the early 1940s the transformation was complete. I was educated to study behavior and I learned to translate my ideas into the new jargon of behaviorism. As I was most interested in speech and hearing, the translation sometimes became tricky. But one’s reputation as a scientist could depend on how well the trick was played.

In 1951, I published Language and Communication [1], a book that grew out of four years of teaching a course at Harvard entitled ‘The Psychology of Language’. In the preface, I wrote: ‘The bias is behavioristic – not fanatically behavioristic, but certainly tainted by a preference. There does not seem to be a more scientific kind of bias, or, if there is, it turns out to be behaviorism after all.’ As I read that book today it is eclectic, not behavioristic. A few years later R.F. Skinner published Verbal Behavior [2], a truly behavioral treatment of language and communication. By Skinner’s standards, my book had little or nothing to do with behavior.

In 1951, I apparently still hoped to gain scientific respectability by swearing allegiance to behaviorism. Five years later, inspired by such colleagues as Noam Chomsky and Jerry Bruner, I had stopped pretending to be a behaviorist. So I date the cognitive revolution in psychology to those years in the early 1950s.

Limitations of information theory

During those years I personally became frustrated in my attempts to apply Claude Shannon’s theory of information to psychology. After some initial success I was unable to extend it beyond Shannon’s own analysis of letter sequences in written texts. The Markov processes on which Shannon’s analysis of language was based had the virtue of being compatible with the stimulus–response analysis favored by behaviorists. But information measurement is based on probabilities and increasingly the probabilities seemed more interesting than their logarithmic values, and neither the probabilities nor their logarithms shed much light on the psychological processes that were responsible for them.

I was therefore ready for Chomsky’s alternative to Markov processes. Once I understood that Shannon’s Markov processes could not converge on natural language, I began to accept syntactic theory as a better account of the cognitive processes responsible for the structural aspects of human language. The grammatical rules that govern phrases and sentences are not behavior. They are...
mentalistic hypotheses about the cognitive processes responsible for the verbal behaviors we observe.

The end of behaviorism

Behaviorism was an exciting adventure for experimental psychology but by the mid-1950s it had become apparent that it could not succeed. As Chomsky remarked, defining psychology as the science of behavior was like defining physics as the science of meter reading. If scientific psychology were to succeed, mentalistic concepts would have to integrate and explain the behavioral data. We were still reluctant to use such terms as 'mentalism' to describe what was needed, so we talked about cognition instead.

Whatever we called it, the cognitive counter-revolution in psychology brought the mind back into experimental psychology. I think it is important to remember that the mind had never disappeared from social or clinical psychology. It was only experimentalists in the US who really believed that behaviorism would work. In my own case, when I became dissatisfied at Harvard between B.F. Skinner's strict behaviorism and S.S. Stevens' psychophysics, I turned to Jerry Bruner's social psychology, and in 1960 that led to the creation of Harvard of the Center for Cognitive Studies. Bruner's group at Bow Street had been calling themselves the 'Cognition Project' for some time, so we simply changed it from a project to a center. Bruner obtained a grant from the Carnegie Corporation of New York and Dean Bundy gave us space to house the enterprise. We assembled a group of bright young graduates and a few senior scholars who shared our interests. Peter Wason, Nelson Goodman and Noam Chomsky had the most influence on my thinking at that time.

Behaviorism flourished primarily in the US and this cognitive revolution in psychology re-opened communication with some distinguished psychologists abroad. In Cambridge, UK, Sir Frederic Bartlett's work on memory and thinking had remained unaffected by behaviorism. In Geneva, Jean Piaget's insights into the minds of children had inspired a small army of followers. And in Moscow, A.R. Luria was one of the first to see the brain and mind as a whole. None of these three spent time at the Center but we knew their work well. Whenever we doubted ourselves we thought of such people and took courage from their accomplishments. I'm happy to say the Harvard Center for Cognitive Studies was a success. The bright young graduates grew up to become important psychologists unafraid of words like mind and expectation and perception and memory. So that was how I experienced the cognitive revolution in psychology.

The cognitive revolution and cognitive science

While experimental psychologists were rethinking the definition of psychology, other important developments were occurring elsewhere. Norbert Wiener's cybernetics was gaining popularity, Marvin Minsky and John McCarthy were inventing artificial intelligence, and Alan Newell and Herb Simon were using computers to simulate cognitive processes. Finally, Chomsky was single-handedly redefining linguistics.

In the Historical Addendum to Newell and Simon's Human Problem Solving [3] they say: '1956 could be taken as the critical year for the development of information processing psychology' (p. 878). This is not difficult to justify. 1956 was the year that McCarthy, Minsky, Shannon and Nat Rochester held a conference on artificial intelligence at Dartmouth that was attended by nearly everyone working in the field at that time. In 1956 Shannon and McCarthy edited Automata Studies [4], and Minsky circulated a technical report that, after many revisions, and 5 years later, became his influential article, 'Steps toward artificial intelligence' [5].

It was also in 1956 that Jerry Bruner, Jackie Goodenough and George Austin published A Study of Thinking [6], which took seriously the notion of cognitive strategies. In 1956 signal-detection theory was applied to perception by Tanner, Swets, Birdsall and others at Michigan. I published an article entitled 'The magical number seven, plus or minus two' [7] describing some limits on our human capacity to process information. In 1956 Ward Goodenough and Floyd Lounsbury published several articles on componential analysis that became models for cognitive anthropology, and J.B. Carroll edited a collection of papers by Benjamin Lee Whorf on the effects of language on thought.

In short, 1956 was a good year for those interested in theories of the mind, but it was only slightly better than the years just preceding and following. Many were riding the waves that began during World War II: those of servo theory, information theory, signal-detection theory, computer theory and computers themselves.

Moment of conception

Newell and Simon were right to put a finger on 1956, which was not only crucial in their own development but for all of us. Indeed, I can narrow it down even further. I date the moment of conception of cognitive science as 11 September, 1956, the second day of a symposium organized by the 'Special Interest Group in Information Theory' at the Massachusetts Institute of Technology [8]. At the time, of course, no one realized that something special had happened so no one thought that it needed a name; that came much later.

The chairman of the organizing committee was Peter Elias, who had only recently arrived at MIT from a Junior Fellowship at Harvard. The first day, 10 September, was devoted to coding theory, but it is the second day of the symposium that I take to be the moment of conception for cognitive science. The morning began with a paper by Newell and Simon on their 'logic machine'. The second paper was from IBM: Nat Rochester and collaborators had used the largest computer then available (an IBM 704 with a 2048-word core memory) to test Donald Hebb's neuro-psychological theory of cell assemblies. Victor Yngve then gave a talk on the statistical analysis of gaps and its relation to syntax.

Noam Chomsky's contribution used information theory as a foil for a public exposition of transformational generative grammar. Elias commented that other linguists had told him that language has all the precision of mathematics but Chomsky was the first linguist to back
up the claim. His 1956 paper contained the ideas that he expanded a year later in his monograph, *Syntactic Structures* ([9]), which initiated a cognitive revolution in theoretical linguistics.

To complete the second day, G.C. Szkiczi described some experiments on the speed of perceptual recognition, I talked about how we avoid the bottleneck created by our limited short-term memory, and Swets and Birdsall explained the significance of signal-detection theory for perceptual recognition. The symposium concluded on the following day.

I left the symposium with a conviction, more intuitive than rational, that experimental psychology, theoretical linguistics, and the computer simulation of cognitive processes were all pieces from a larger whole and that the future would see a progressive elaboration and coordination of their shared concerns.

The birth of cognitive science

By 1960 it was clear that something interdisciplinary was happening. At Harvard we called it cognitive studies, at Carnegie-Mellon they called it information-processing psychology, and at La Jolla they called it cognitive science. What you called it didn’t really matter until 1976, when the Alfred P. Sloan Foundation became interested.

The Sloan Foundation had just completed a highly successful program of support for a new field called ‘neuroscience’ and two vice-presidents of the Foundation, Steve White and AI Singer, were thinking that the next step would be to bridge the gap between brain and mind. They needed some way to refer to this next step and they selected ‘cognitive science.’ They created a Sloan Special Program in Cognitive Science in order to explore the possibilities.

I learned of the Foundation’s interest in 1977 from Kenneth A. Klivington, who was on the staff at the Foundation. My recollection is that Ken had talked to Marvin Minsky and others at MIT and was considering a recommendation that the Foundation invest in artificial intelligence. Shamelessly, I argued that in that case the Foundation’s money would be spent buying computers. I claimed that AI was merely part of a much larger movement. At that time the Sloan Foundation was sensitive to the charge that it had become part of the MIT endowment, so my lobbying for a broader constituency was well received.

**Interdisciplinary activities**

I argued that at least six disciplines were involved: psychology, linguistics, neuroscience, computer science, anthropology and philosophy. I saw psychology, linguistics and computer science as central, the other three as peripheral. These fields represented, and still represent, an institutionally convenient but intellectually awkward division. Each, by historical accident, had inherited a particular way of looking at cognition and each had progressed far enough to recognize that the solution to some of its problems depended crucially on the solution of problems traditionally allocated to other disciplines.

The Sloan Foundation accepted my argument and a committee of people from the several fields was assembled to summarize the state of cognitive science in 1978, and to write a report recommending appropriate action. The committee met once, in Kansas City. It quickly became apparent that everyone knew his own field and had heard of two or three interesting findings in other fields. After hours of discussion, experts in discipline X grew unwilling to make any judgments about discipline Y, and so forth. In the end, they did what they were competent to do: each summarized his or her own field and the editors – Samuel Jay Keyser, Edward Walker and myself – patched together a report (Keyser, S.J., Miller, G.A., and Walker, E., Cognitive Science in 1978. An unpublished report submitted to the Alfred P. Sloan Foundation, New York).

Our report had one figure, which is reproduced here (Fig. 1). The six fields are connected in a hexagon. Each line in the figure represents an area of interdisciplinary inquiry that was well defined in 1978 and that involved the tools of the two disciplines it linked together. Thus, cybernetics used concepts developed by computer science to model brain functions elucidated in neuroscience. Similarly, computer science and linguistics were already linked through computational linguistics. Linguistics and psychology are linked by psycholinguistics, anthropology and neuroscience were linked by studies of the evolution of the brain, and so on. Today, I believe, all fifteen possible links could be instantiated with respectable research, and the eleven links we saw as existing in 1978 have been greatly strengthened.

The report was submitted, reviewed by another committee of experts, and accepted by the Sloan Foundation. The program that was initiated provided grants to several universities with the condition that the funds be used to promote communication between disciplines. One of the smaller grants went to Michael Gazzaniga, then at the Cornell Medical School, and enabled him to initiate what has since become cognitive neuroscience. As a consequence of the Sloan program, many scholars became familiar with and tolerant of work in other disciplines. For several years, interdisciplinary seminars, colloquia and symposia flourished.
Cognitive sciences today
Unfortunately, the Alfred P. Sloan Foundation did not follow up this initiative, but the interactions stimulated in the early 1980s have left their mark. Some veterans of those days question whether the program was successful, and whether there really is something now that we can call 'cognitive science'. For myself, I prefer to speak of the cognitive sciences, in the plural. But the original dream of a unified science that would discover the representational and computational capacities of the human mind and their structural and functional realization in the human brain still has an appeal that I cannot resist.

References
6 Bruner, J.S. et al. (1956) A Study of Thinking, John Wiley
7 Miller, G.A. (1956) The magical number seven, plus or minus two. Psychol. Rev. 63, 81–97
8 Elias, P. et al. (1956) Information theory. IRE Trans. Information Theory, IT-2(3)
9 Chomsky, N. (1957) Syntactic Structures, Mouton
The Physical Symbol Systems Hypothesis

Formal logic: the symbols are words like "and", "or", "not", "for all x" and so on. The expressions are statements in formal logic which can be true or false. The processes are the rules of logical deduction.

Algebra: the symbols are "+", "x", "y", "1", "2", "3", etc. The expressions are equations. The processes are the rules of algebra, that allow one to manipulate a mathematical expression and retain its truth.

A digital computer: the symbols are zeros and ones of computer memory, the processes are the operations of the CPU that change memory.

Chess: the symbols are the pieces, the processes are the legal chess moves, the expressions are the positions of all the pieces on the board.
The Physical Symbol Systems Hypothesis

Newell and Simon:

"A physical symbol system has the necessary and sufficient means for intelligent action."

A physical symbol system "consists of a set of entities, called symbols, which are physical patterns that can occur as components of another type of entity called an expression (or symbol structure). Thus, a symbol structure is composed of a number of instances (or tokens) of symbols related in some physical way (such as one token being next to another). At any instant of time the system will contain a collection of these symbol structures. Besides these structures, the system also contains a collection of processes that operate on expressions to produce other expressions: processes of creation, modification, reproduction and destruction. A physical symbol system is a machine that produces through time an evolving collection of symbol structures. Such a system exists in a world of objects wider than just these symbolic expressions themselves."
Neural Networks: Biological and Artificial
- McCulloch-Pitts model
- Hebb’s learning rule
- Perceptron: learning machine
- Multi-layer Perceptron
- The Connectionist Movement: Parallel Distributed Processing
- Back to Neurobiology!
• Phrenology
• Localizationist view
• Aggregate field view
• Neuron doctrine
• Combining neuroscience and cognitive science: brain mapping
Cognitive Neuroscience: Multiple Levels - Multiple Methods

Experimental methods and disciplines

- **Anatomy**: discover the structure of the brain
- **Electrophysiology**
  - Single cell recording
  - Electroencephalography (EEG)
- Psychological experiments on people with brain damage, e.g. HM
- **Brain imaging**

Brain imaging

- EEG and MEG
- CT or CAT: Computerized axial tomography
- PET: Positron emission tomography
- fMRI: Functional magnetic resonance imaging
Multiple Memory Systems

- Declarative
  - Semantic (facts)
  - Episodic (events)
- Nondeclarative
  - Procedural skills (e.g., motor, perceptual, cognitive)
  - Priming (perceptual, semantic)
  - Conditioning
  - Nonassociative (habituation, sensitization)

Exteroceptors

Cortex

Thalamus

Procedural (S-R) Caudate Nucleus

Emotional (S-Af) Amygdala

Declarative (S-S) Medial Temporal Lobe

Nucleus Accumbens

Matrix

Substantia Nigra

Globus Pallidus

Behavior
Minireview

Memory systems of the brain: A brief history and current perspective

Larry R. Squire*
Veterans Affairs Healthcare System, San Diego, CA 92161, USA
Departments of Psychiatry, Neurosciences, and Psychology, University of California, San Diego, La Jolla, CA 92039, USA

Received 23 April 2004; accepted 14 June 2004
Available online 4 August 2004

Abstract

The idea that memory is composed of distinct systems has a long history but became a topic of experimental inquiry only after the middle of the 20th century. Beginning about 1980, evidence from normal subjects, amnesic patients, and experimental animals converged on the view that a fundamental distinction could be drawn between a kind of memory that is accessible to conscious recollection and another kind that is not. Subsequent work shifted thinking beyond dichotomies to a view, grounded in biology, that memory is composed of multiple separate systems supported, for example, by the hippocampus and related structures, the amygdala, the neostriatum, and the cerebellum. This article traces the development of these ideas and provides a current perspective on how these brain systems operate to support behavior.

Published by Elsevier Inc.

Keywords: Declarative; Nondeclarative; Priming; Episodic; Semantic; Procedural; Habit; Conditioning

The idea that memory is not a single faculty of the mind is not itself new. One can find expressions of this idea in the writings of psychologists and philosophers more than a century ago. For example, it is often noted that Maine de Biran wrote in 1804 about mechanical memory, sensitive memory, and representative memory (Maine de Biran, 1804/1929), and that William James (1890) wrote separate chapters on memory and habit in his Principles of Psychology. Another perceptive writer was Bergson (1910). Focusing on habits, he wrote:

[It is] a memory profoundly different ... always bent upon action, seated in the present and looking only to the future... In truth it no longer represents our past to us, it acts it; and if it still deserves the name of memory, it is not because it conserves bygone images, but because it prolongs their useful effect into the present moment. (p. 93).

One can identify other antecedents as well. McDougall (1923) distinguished between explicit and implicit recognition, and Tolman (1948) wrote at length on the proposition that there is more than one kind of learning. Most often in the earlier literature, one finds this idea expressed as a dichotomy between two kinds of memory. Thus, Ryle (1949) distinguished between knowing how and knowing that, and Bruner (1969) contrasted memory without record and memory with record. In the 1970s a similar distinction was discussed in the artificial intelligence literature as procedural and declarative knowledge (Winograd, 1975).

Yet these writings did not lead to a single view of the matter. What was finally needed was not philosophical discourse or psychological intuition but experimental inquiry into how the brain actually stores information. The modern experimental era arguably began when Milner (1962) demonstrated that a hand-eye coordination skill (mirror drawing) could be learned over a period of days by the severely amnesic patient H.M. in the absence of any memory of having practiced the task before. While this finding showed that memory was not
unitary, discussions at the time tended to set aside motor skills as a special case that represented a less cognitive form of memory. The idea was that all the rest of memory is of one piece and that all the rest of memory is impaired in patients like H.M.

Subsequently, and into the 1970s, there were two strands of work that moved the focus beyond motor skills. One strand came from studies of experimental animals, and the other strand came from work with amnesic patients. In the case of animal studies, distinctions were drawn between recognition and associative memory (Gauffan, 1974), contextual retrieval and habit (Hirsh, 1974), and taxon and locale memory (O’Keefe & Nadel, 1978). These proposals had in common the idea that the hippocampus, and perhaps related structures, was involved in one particular type of memory. Yet, the proposals also differed from each other and their prescience was not widely appreciated, in part because at the time the findings in experimental animals did not conform well to the findings from human memory and amnesia. Indeed, one of the reasons that it took a long time to develop an animal model of human amnesia was that it was not understood what needed to be modeled. That is, the description of human amnesia was itself incomplete, and many tasks given to experimental animals were in fact tasks that animals could succeed at just as patients succeeded at motor skills.

However, after an animal model of human amnesia was established in the nonhuman primate (Mishkin, 1982; Squire & Zola-Morgan, 1983), a useful correspondence was established between the findings for experimental animals and humans, and since that time work with experimental animals has been invaluable for understanding the memory systems of the brain.

The second strand of relevant work in the 1960s and 1970s came from demonstrations of unexpectedly good learning and retention by amnesic patients on tasks other than motor skill tasks (Milner, Corkin, & Teuber, 1968; Warrington & Weiskrantz, 1968). However, there were two reasons why these findings, and others that followed in the subsequent few years, did not lead to proposals of multiple memory systems. First, even when the performance of amnesic patients was good, it sometimes fell short of normal levels. Second, in those cases when amnesic patients did perform normally, or near normally, a dominant view was that amnesia was therefore a retrieval deficit (Warrington & Weiskrantz, 1970, 1978; Weiskrantz, 1978).

Quite apart from issues of interpretation, it took some time to appreciate the crucial role of test instructions in determining whether amnesic patients performed as well as controls. For example, amnesic patients often performed well when they were given three-letter word stems as cues for previously presented words, a demonstration of what would later be termed priming. Only later was it appreciated that normal performance by amnesic patients depended on a nonmemory kind of instruction (Use this word stem to form the first word that comes to mind). With conventional memory instructions (Use this word stem as a cue to retrieve a previously presented word), controls performed better than amnesic patients (Graf, Squire, & Mandler, 1984).

The fact that amnesic patients do perform normally when these tests are structured appropriately (for an early example, see Jacoby & Witherspoon, 1982) showed that the phenomenon of priming is a distinct form of memory, separate from what is impaired in amnesia (Schacter & Buckner, 1998; Tulving & Schacter, 1990).

Evidence for the special status of priming also came from studies of normal subjects (Tulving, Schacter, & Stark, 1982). These authors wrote “. . . we are tempted to think that [these priming effects] reflect the operation of some other, as yet little understood, memory system.” (p. 341). Perhaps the most compelling evidence for the independence of priming and the kind of memory impaired in amnesia came much later from parallel studies of perceptual priming (Hamann & Squire, 1997; Stark & Squire, 2000) and conceptual priming (Levy, Stark, & Squire, in press). This work showed that severely amnesic patients can exhibit fully intact priming while performing at chance on conventional recognition memory tests for the same test items.

During the period that priming was discovered to be intact in amnesia, it also became appreciated that motor skills are not special after all but are a subset of a larger domain of skill-like abilities, all of which are preserved in amnesia. The first example was the task of mirror reading, a perceptual skill, which amnesic patients acquired at a normal rate despite poor memory for the words that they read (Cohen & Squire, 1980). Other demonstrations followed (e.g., the ability to resolve stereoscopic images, Benzing & Squire, 1989; cognitive skill learning, Squire & Frabjush, 1990; artificial grammar learning, Knowlton, Ramus, & Squire, 1992; and category learning, Knowlton & Squire, 1993).

Initially, these data were interpreted according to a distinction between declarative and procedure knowledge (Cohen & Squire, 1980). Other, similar dichotomies also came into use [e.g., explicit and implicit memory (Graf & Schacter, 1985); memory and habit (Mishkin, Malamut, & Bachevalier, 1984)]. However, during the 1980s, it became progressively difficult to fit the accumulating data to the two poles of a dichotomy. For example, emerging findings about priming led Tulving and his colleagues to write:

But even if we accept the broad division of memory into procedural and propositional forms and the division of propositional forms into episodic and semantic forms, there are phenomena that do not seem to fit readily into such a taxonomy. (Tulving et al., 1982, p. 336).
Work with experimental animals also influenced the movement away from dichotomies. First, in the early 1980s, the cerebellum was discovered to be essential for delay eyeblink conditioning (McCormick, Clark, Lavond, & Thompson, 1982), a form of learning that was entirely preserved both in animals with hippocampal lesions (Schmaltz & Theios, 1972) and in severely amnesic patients (Clark & Squire, 1998; Gabrieli, McGlinchey-Berroth, Gluck, Cermak, & Disterhoft, 1995). Second, the neostriatum was identified as important for the sort of gradual, feedback-guided learning that results in habit memory (Mishkin et al., 1984), and an elegant double dissociation was demonstrated in rats after fornix and caudate lesions in two tasks that appeared to measure declarative memory and habit memory, respectively (Packard, Hirsh, & White, 1989). A similar contrast between declarative memory and habit memory was later demonstrated for amnesic patients and patients with Parkinson’s disease (Knowlton, Mangels, & Squire, 1996). Finally, it was shown that still other types of learning, which involve the attachment of positive or negative valence to a stimulus, as in fear conditioning or conditioned place preference, have an essential dependence on the amygdala (Davis, 1992; Fanselow, 1994; LeDoux, 2004; McDonald & White, 1993).

Given the wide variety of learning and memory tasks explored in these studies, and the number of different brain structures that were implicated, an account of memory based on a two-part dichotomy came to appear overly simplistic. Indeed, one wondered what the various kinds of memory that were preserved in amnesic patients had in common aside from the fact that they were not declarative. Accordingly, beginning in the mid 1980s, the perspective shifted to a framework that accommodated multiple (i.e., more than two) memory systems (see for example, Tulving, 1985). At that time, the term “nondeclarative” was introduced with the idea that declarative memory refers to one memory system and that “nondeclarative memory” is an umbrella term referring to several additional memory systems (Squire & Zola-Morgan, 1988). The seminal volume of this period, The Memory Systems of 1994 (Schacter & Tulving, 1994), presented a collection of writings that largely reflected this point of view.

The result of all this was that it was now possible to reach a clearer, more concrete, and ultimately a more accurate classification of memory by placing the work within a biological framework. Fig. 1 illustrates a taxonomy that incorporates these ideas (for the earliest version of this diagram, see Squire, 1987). Declarative memory is the kind of memory that is meant when the term “memory” is used in everyday language. It refers to the capacity for conscious recollection about facts and events and is the kind of memory that is impaired in amnesia and dependent on structures in the medial temporal lobe and midline diencephalon. Other characteristics of declarative memory allow the term to be extended to experimental animals and bring work with humans and animals into more comfortable contact.

Thus, declarative memory allows remembered material to be compared and contrasted. It supports the encoding of memories in terms of relationships among multiple items and events. The stored representations are flexible and can guide performance under a wide range of test conditions. Declarative memory is representational. It provides a way to model the external world, and as a model of the world it is either true or false. In contrast, nondeclarative memory is neither true nor false. It is dispositional and is expressed through performance rather than recollection. Nondeclarative forms of memory occur as modifications within specialized performance sys-

![Fig. 1. A taxonomy of mammalian long-term memory systems. The taxonomy lists the brain structures thought to be especially important for each form of declarative and nondeclarative memory. In addition to its central role in emotional learning, the amygdala is able to modulate the strength of both declarative and nondeclarative memory.](image-url)
tems. The memories are revealed through reactivation of the systems within which the learning originally occurred.

Declarative memory can be divided into semantic memory (facts about the world) and episodic memory (the capacity to re-experience an event in the context in which it originally occurred) (Tulving, 1983). Episodic memory requires the participation of brain systems in addition to those that support semantic memory, for example, the frontal lobes (Shimamura & Squire, 1987; Tulving, 1989). It is an interesting question whether nonhuman animals have a capacity for episodic memory (Tulving, 2002), but the idea is difficult to put to test (for relevant experiment and discussion, see Clayton & Dickinson, 1998; Tulving, in press).

The various memory systems can be distinguished in terms of the different kinds of information they process and the principles by which they operate. In the case of declarative memory, an important principle is the ability to detect and encode what is unique about a single event, which by definition occurs at a particular time and place. In the case of nondeclarative memory, an important principle is the ability to gradually extract the common elements from a series of separate events. Sherry and Schacter (1987) suggested that multiple memory systems evolved because they serve distinct and functionally incompatible purposes. For example, the gradual changes that occur in birdsong learning are fundamentally different from and have a different function than the rapid learning that occurs when a bird caches food for later recovery.

The memory systems of the brain operate in parallel to support behavior. For example, an aversive childhood event involving being knocked down by a large dog can lead to a stable declarative memory for the event itself as well as a long-lasting nondeclarative fear of dogs (a phobia) that is experienced as a personality trait rather than as a memory. The idea that memory systems operate independently and in parallel is nicely illustrated by a study of rats that was carried out in a four-arm, plus-shaped maze (Packard & McGaugh, 1996). First, the upper (north) arm was blocked, and rats were started from the lower (south) arm and trained to turn to the west to find food. Probe trials were introduced at various times by starting rats from the north arm (with the south arm now blocked). In probe trials given early in training, rats entered the rewarded (west) arm, that is, they returned to the place where food had been found. In probe trials introduced later in training, rats went to the nonrewarded arm, that is, they turned east, thereby repeating the left-turn response that they had previously made to find food. Place responding early in training was abolished by lidocaine injections into the hippocampus, and rats exhibited no preference for either arm. Correspondingly, later in training, the preference for a left-hand turn was abolished by lidocaine injections into the caudate nucleus. Interestingly, in this case rats did not behave randomly but now exhibited place responding (that is, they turned west). Thus, even though behavior was dominated later in training by the caudate nucleus, and by left-hand turns, information remained available about the place where food could be found. When the caudate nucleus was disabled, the parallel memory system supported by the hippocampus was revealed.

While one memory system may substitute for another in the sense just described, what is learned differs markedly depending on which memory system is engaged. In the rat, what is learned might be a spatial location or a turning response. In humans, the difference can also be quite striking. For example, consider a task introduced by Tulving, Hayman, & MacDonald (1991), in which simple three-word sentences are presented as novel “facts” to be learned, e.g., medicine cured hiccup. In a recent study, sentences were presented repeatedly across several sessions, and recall was subsequently tested by asking participants to complete the first two words of each sentence so as to form a sentence that had been studied (medicine cured hiccup) (Bayley & Squire, 2002). When the hippocampus and related structures were able to support performance, as in healthy volunteers, learning occurred rapidly and what was learned was accompanied by conscious knowledge about which answers were correct. Further, the learning was readily expressed even if a part of the sentence was replaced by a synonym (medicine relieved). By contrast, in severe amnesia, declarative memory was not available and learning was extremely slow. Importantly, what little was learned during 12 weeks of training was outside of awareness, confidence ratings were unrelated to success, and performance succeeded only when the first two words in the test sentences were the same words that had appeared during training.

These findings in humans and rats emphasize that what is important is not only the task that is to be learned but also what strategy is implemented during learning, which in turn reflects what memory system is engaged. Under some circumstances the strategy that is engaged is not optimal for solving a task. For example, hippocampal lesions in rats can facilitate the acquisition of a maze task that requires repeated visits to illuminated arms and that is dependent on the caudate nucleus (Packard et al., 1989). The hippocampal lesion presumably disrupts the tendency to use a nonoptimal declarative memory strategy, in the same sense that trying to memorize what one is doing can interfere with human skill learning. Indeed, when humans acquire a difficult habit learning task, structures important for habit learning and structures important for memorizing (i.e., declarative memory) can appear to compete for control of performance. Early in learning, fMRI revealed activity within the medial temporal lobe, as if participants were attempting to memorize the task structure (Poldrack et
monkeys learn the pattern discrimination task, one might
humans that is acquired as a skill, and in the way that
(2000). To achieve a two-choice discrimination task for
(2004). (Mishkin et al., 1984; Teng, Stefanacci, Squire, & Zola,
stimulus is the correct one (Squire, Zola-Morgan, &
1983). Yet, amnesic patients learn such tasks in a few tri-
ilarly by experimental animals. The best-known
example of this circumstance is visual pattern discrimina-
tion learning (e.g., + vs. →). Monkeys with large medial
temporal lobe lesions are intact at the learning and reten-
tion of pattern discriminations (Squire & Zola-Morgan,
2000; Teng, Stefanacci, Squire, & Zola, 2000). To achieve a two-choice discrimination task for
humans that is acquired as a skill, and in the way that
monkeys learn the pattern discrimination task gradually
during several hundred trials in a manner reminiscent of skill learning (Iversen, 1976) and that humans ap-
proach the task as a simple problem of memorization.
Whereas in humans the learning and retention of pattern
discriminations is dependent on the medial temporal
lobe, in monkeys the pattern discrimination task is depen-
dant on an inferior temporal lobe-neostriatal pathway
(Mishkin et al., 1984; Teng, Stefanacci, Squire, & Zola,
2000). The difference appears to lie in the fact that
monkeys learn the pattern discrimination task gradually
during several hundred trials in a manner reminiscent of skill learning (Iversen, 1976) and that humans ap-
proach the task as a simple problem of memorization.
Whereas in humans the learning and retention of pattern
discriminations is dependent on the medial temporal
lobe, in monkeys the pattern discrimination task is depen-
dant on an inferior temporal lobe-neostriatal pathway
(Mishkin et al., 1984; Teng, Stefanacci, Squire, & Zola,
2000). To achieve a two-choice discrimination task for
humans that is acquired as a skill, and in the way that
monkeys learn the pattern discrimination task, one might
ask humans to learn to discriminate between the paint-
ings of a master and the paintings of a talented forger.
The notion of multiple memory systems is now widely
accepted (Eichenbaum & Cohen, 2001; Schacter, Wag-
ner, & Buckner, 2000; Squire, Stark, & Clark, 2004).
Yet it is interesting that one can still find the exclusively
psychological perspective “... that there is only one
memory system, which preserves all experiences and is
used in all tasks’” (Whittlesea & Price, 2001). Similar
viewpoints have been advanced occasionally during the
past 20 years. Typically, the notion is that there is only
one memory system but that there are multiple processes
operating on this system or multiple ways of accessing
its contents. The difficulty with such views is that they
are unnecessarily abstract and make insufficient contact with
biology. For example, the findings from eyeblink
conditioning provide direct evidence for a kind of mem-
ory that can be acquired, stored, and retrieved in the ab-
sence of the forebrain. Other kinds of memory (e.g.,
perceptual learning, declarative memory) do require
the forebrain. The locus of memory storage is entirely
different in these cases, and the learning proceeds by dif-
f erent principles. Perhaps there is some level of abstrac-
tion at which synaptic changes within the cerebellum
and synaptic changes within the neocortex can be
viewed as different expressions of a single memory sys-
tem. However, such a perspective tends to ignore rather
than to embrace the enormous amount that has been
learned about neuroanatomy, the molecular and cellular
biology of synaptic change, and the organization of
brain systems.
In biology, the term “system” is defined in terms of
both structure and function. The study of memory has
benefited in recent years as discussion of memory systems has drawn increasingly on what is known about biological systems. Strictly functional constructs founded in psychological science alone are seldom suffi-
cient because psychology has matured to the point
where it is able to connect concepts about memory to
biology. And history shows that, as biological informa-
tion becomes available about structure and mechanism,
explanation becomes more concrete and less dependent
terminology.
During the past two centuries, the study of memory,
and the study of cognition in general, has been central to
three disciplines: first philosophy, then psychology, and
now biology. One can expect the contributions of biol-
ogy to the study of memory to become even more cen-
tral in the coming years as more is learned about the
molecular biology of synaptic change and the neurosci-
ence of brain systems.

Acknowledgments
Supported by the Medical Research Service of the
Department of Veterans Affairs, the National Institute
of Mental Health, and the Metropolitan Life Founda-
tion. I thank Robert Clark and Nicola Broadbent for
comments and discussion.

References
Bayley, P. J., & Squire, L. R. (2002). Medial temporal lobe amnesia:
gradual acquisition of factual information by nondeclarative
Benzing, W. C., & Squire, L. R. (1989). Preserved learning and
memory in amnesia: Intact adaptation-level effects and learning of
stereoscopic depth. Behavioral Neuroscience, 103, 538–547.
Bergson, H. L. (1910). Matter and Memory (authorized translation by
Waugh (Eds.), The pathology of memory (pp. 253–259). New York:
Academic Press.
of pattern analyzing skill in amnesia: Dissociation of knowing how
Aggleton (Ed.), The amygdala: Neurobiological aspects of emotion,


Language acquisition, evolution and processing

- How does the mind turn sounds into words (phonology)?
- How does the mind turn words into sentences (syntax)?
- How does the mind understand words and sentences (semantics)?
- How does the mind understand discourse (semantics, pragmatics)?
- How does the mind generate discourse?
- How does the mind translate between languages?
- How does the mind acquire the capacities just described?
- To what extent is knowledge of language innate?
Hypotheses about how the mind uses language

• Symbolic
  – Linguistic knowledge consists largely of rules that govern phonological and syntactic processing.
  – The computational procedures involved in understanding and generating language are largely rule-based.
  – Language learning is learning of rules.
  – Many of these rules are innate.
  – The leading proponent of this general view has been Noam Chomsky.
  – Rule-based models of language comprehension and generation have been developed in the SOAR system and within other frameworks

• Connectionist
  – Linguistic knowledge consists largely of statistical constraints that are less general than rules and are encoded in neural networks.
  – The computational procedures involved in understanding and generating language are largely parallel constraint satisfaction.
Syntax

- Grammatical rules that specify how words and other morphemes are arranged so as to yield acceptable sentences.
- Implicit knowledge provides linguistic intuitions of ungrammatical strings and equivalency of meaning of different grammatical constructions.
Semantics

Semantics is the study of meaning, that is, how people mentally represent the meaning of words and sentences.

A morpheme is the minimal unit of speech used to code a specific meaning. (e.g., pill and kill are morphemes; so are -s and -ed.

Morphemes taken together comprise a mental lexicon, the dictionary of long-term memory.
Pragmatics

The manner in which speakers communicate their intentions depends on the social context.

Direct speech acts (e.g., command: “Open the window!”) may be socially awkward. Indirect speech acts assume the guise of a different speech act to achieve the same result (e.g., inform: “It is really hot in here.”)
Pragmatics

The cooperative principle guides conversations so that speakers utter appropriate statements. We speak audibly, use language understood by the listener, and follow the rules.

Listeners draw appropriate inferences called conversational implicatures. For example, if I say “I am out of gas” you might say “There’s a gas station around the corner.”
Contrasts to Animal Communication

- Only language uses symbols to represent objects. Words are detached from their referents unlike the calls of a bird or chimpanzee. Displacement in space and time is thus possible with language.

- Productivity is ability to create novel sentences that can be understood by other speakers of the language. Although chimps can learn ASL and sign novel expressions, there is a vast difference in productivity.
Universal Grammar

- Genetically determined knowledge of language allows rapid acquisition, in theory. At issue is whether a general learning process is sufficient.
- A Language Acquisition Device (LAD) requires parameter setting from exposure to examples but restricts range (e.g., SOV, SVO, VSO, VOS in a pinch but not OVS or OSV).
Neural Systems

- Broca’s aphasia is an inability to speak fluently without effort and correct grammar. Speech is halting and agrammatic.
- Wernicke’s aphasia is a comprehension dysfunction. Speech is fluent and effortless but also semantically vacuous.
Figure 10.7. Broca's area and Wernicke's area in the left hemisphere.
Thought and Language

- Identity hypothesis equates thought with subvocal speech.
- Modularity regards language as independent of general cognitive resources.
- Alternatively, language might be dependent on such general resources.
- Linguistic relativity asserts thought is dependent on the form of language.
Figure 10.10. Possible relations between thought and language.
Thought and Language

- Only the identity hypothesis can be rejected unequivocally (curare paralyzes speech musculature but thought remains from neocortical activity).

- Supporting modularity, linguistic savants and Williams Syndrome show speech can be preserved despite mental retardation. However, general intellectual impairment from accidents or disease typically causes aphasia.
Thought and Language

- Color names differ across cultures but perception of the color spectrum as revealed by color typicality effects is universal.

- However, polysemy in English creates ambiguity in comprehension. Euphemisms can cloud thought. Language A can be translated to Language B but often there is a cognitive cost involved. Thus, a weak version of linguistic relativity finds
Thought and Language

- Object permanence (sensori-motor thought) precedes first words at around 12 months.

- Gestural abilities (e.g., pointing, pretending, imitating) predict the transition from babbling to first words.
TUTORIAL ON EMBODIMENT

Welcome to the tutorial on embodiment

Tutorial Content:

1. Introduction
   1.1. Outline
   1.2. Target Audience

2. Traditional Artificial Intelligence, its Successes and Problems
   2.1. The Study of Intelligence
   2.2. Successes of Traditional AI
   2.3. Problems of Traditional AI

3. Embodiment
   3.1. Theoretical Scheme
   3.2. Locomotion Case Studies
      3.2.1. Brain-less and power-less creatures
         3.2.1.1. Euthanized trout
         3.2.1.2. Passive dynamic walkers
         3.2.1.3. Exercise
         3.2.2. Brain-less and powered creatures
            3.2.2.1. Behavioral diversity - crazy bird
            3.2.2.2. Self-stabilization
            3.2.2.3. Exercise
            3.2.3. Simple control exploiting body dynamics and interaction with environment
               3.2.3.1. Leg coordination in insect walking
               3.2.3.2. Exploiting self-adaptation: cockroaches climbing over obstacles
               3.2.3.3. Exercise
   3.3. Grasping Case Studies
      3.3.1. Cheap grasping with a robotic hand
3.3.2. Cheap grasping with an universal gripper
3.3.3. Exercise
3.4. Additional Online Lectures

4. Information Theoretic Implications of Embodiment
4.1. The Role of Sensory Morphology in Perception
4.1.1. Eye morphology
4.1.2. Slippage detection through skin morphology
4.2. Information Self-structuring through Sensory-motor Coordination
4.2.1. Visual perception case study
4.2.2. Quantifying information structure
4.3. Exercise

5. Embodied Cognition
5.1. Cognitive Science Paradigms
5.1.1. Cognitivism
5.1.2. Connectionism
5.1.3. Embodied dynamicism and enactivism
5.1.4. Summary
5.2. Learning Categories
5.2.1. Grounding toddler learning in sensory-motor dynamics
5.2.2. Categorization in artificial agents
5.2.3. Where mathematics comes from?
5.3. Cognition from Bottom-up
5.3.1. Body schema and forward models
5.3.2. Why is that cognition?
5.3.3. Cognitive development and the iCub humanoid robot

Editors:
Matej Hoffmann, Dorit Assaf, Rolf Pfeifer
Artificial Intelligence Laboratory
Department of Informatics
University of Zurich
Switzerland

Contact:
Matej Hoffmann (hoffmann@ifi.uzh.ch)
Dorit Assaf (assaf@ifi.uzh.ch)
Classical and modern view of intelligence.
Illustrations by Shun Iwasawa
2.1 Creature Consciousness An animal, person or other cognitive system may be regarded as conscious in a number of different senses.

Sentience. It may be conscious in the generic sense of simply being a sentient creature, one capable of sensing and responding to its world (Armstrong 1981). Being conscious in this sense may admit of degrees, and just what sort of sensory capacities are sufficient may not be sharply defined. Are fish conscious in the relevant respect? And what of shrimp or bees?

Wakefulness. One might further require that the organism actually be exercising such a capacity rather than merely having the ability or disposition to do so. Thus one might count it as conscious only if it were awake and normally alert. In that sense organisms would not count as conscious when asleep or in any of the deeper levels of coma. Again boundaries may be blurry, and intermediate cases may be involved. For example, is one conscious in the relevant sense when dreaming, hypnotized or in a fugue state?

Self-consciousness. A third and yet more demanding sense might define conscious creatures as those that are not only aware but also aware that they are aware, thus treating creature consciousness as a form of self-consciousness (Carruthers 2000). The self-awareness requirement might get interpreted in a variety of ways, and which creatures would qualify as conscious in the relevant sense will vary accordingly. If it is taken to involve explicit conceptual self-awareness, many non-human animals and even young children might fail to qualify, but if only more rudimentary implicit forms of self-awareness are required then a wide range of nonlinguistic creatures might
count as self-conscious.

What it is like. Thomas Nagel’s (1974) famous ‘what it is like’ criterion aims to capture another and perhaps more subjective notion of being a conscious organism. According to Nagel, a being is conscious just if there is something that it is like to be that creature, i.e., some subjective way the world seems or appears from the creature’s mental or experiential point of view. In Nagel’s example, bats are conscious because there is something that it is like for a bat to experience its world through its echo-locatory senses, even though we humans from our human point of view can not emphatically understand what such a mode of consciousness is like from the bat’s own point of view.

Subject of conscious states. A fifth alternative would be to define the notion of a conscious organism in terms of conscious states. That is, one might first define what makes a mental state a conscious mental state, and then define being a conscious creature in terms of having such states. One’s concept of a conscious organism would then depend upon the particular account one gives of conscious states (section 2.2).

Transitive Consciousness. In addition to describing creatures as conscious in these various senses, there are also related senses in which creatures are described as being conscious of various things. The distinction is sometimes marked as that between transitive and intransitive notions of consciousness, with the former involving some object at which consciousness is directed (Rosenthal 1986).
Integrating Emotions and Cognitions
Cognition and emotion

**Dr. Luiz Pessoa**, Indiana University, Bloomington, IN

The relationship between **cognition and emotion** has fascinated important thinkers within the Western intellectual tradition. Historically, emotion and cognition have been viewed as largely separate. In the past two decades, however, a growing body of work has pointed to the interdependence between the two.

Introduction

Cognition refers to processes such as memory, **attention**, language, problem solving, and planning. Many cognitive processes are thought to involve sophisticated functions that may be unique to primates. They often involve so-called controlled processes, such as when the pursuit of a goal (e.g., maintaining information in mind) needs to be protected from interference (e.g., a distracting stimulus). A prototypical example of a neural correlate of a cognitive process is the sustained firing of cells in dorsolateral prefrontal cortex as a monkey maintains information in mind for brief periods of time (Fuster and Alexander, 1971; Kubota and Niki, 1971). With the advent of functional MRI (fMRI), it appears that cognitive processes engage cortical regions of the brain (Gazzaniga et al., 2008).

Whereas there is relative agreement about what constitutes cognition, the same cannot be said about emotion. Some investigators use definitions that incorporate the concepts of drive and motivation: emotions are states elicited by rewards and punishers (Rolls, 2005). Others favor the view that emotions are involved in the conscious (or unconscious) evaluation of events (Arnold, 1960) (i.e., appraisals). Some approaches focus on basic emotions (Ekman, 1992) (e.g., fear, anger), others on an extended set of emotions, including moral ones (Haidt, 2003; Moll et al., 2005) (e.g., pride, envy). Strong evidence also links emotions to the body (Damasio, 1994). Brain structures linked to emotion are often subcortical, such as the amygdala, ventral striatum, and hypothalamus. These structures are often considered evolutionarily conserved, or **primitive**. They are also believed to operate fast and in an **automatic fashion**, such that certain trigger features (e.g., the white of the eyes in a fearful expression (Whalen et al., 2004)) are relatively unfiltered and always evoke responses that may be...
important for survival. Accordingly, an individual may not be necessarily conscious of a stimulus that may have triggered brain responses in an affective brain region, such as the amygdala. For discussion, see (Ohman, 2002; Pessoa, 2005).

Because of the inherent difficulty in providing clear definitions for both cognition and emotion, they will not be further defined here. We now turn to illustrating some of the interactions between emotion and cognition. Given the enormous scope of this topic, by necessity, this review will be relatively narrow in scope and will emphasize the brain systems involved in the interactions between emotion and i) perception and attention; ii) learning and memory; and iii) behavioral inhibition and working memory. Other valuable sources include (Damasio, 1994; LeDoux, 1996; Damasio, 1999; Dolan, 2003; Rolls, 2005; Phelps, 2006). A key conclusion from this review and from other current discussions of the relationship between cognition and emotion is that it is probably counterproductive to try to separate them. Instead, current thinking emphasizes their interdependence in ways that challenge a simple division of labor into separate cognitive and emotional domains. In particular, in the context of the brain, the general dichotomization alluded to above in terms of cortical-cognitive and subcortical-emotional brain areas is now viewed as largely simplified and breaks down rather quickly when more in-depth analyses are carried out; e.g., (Pessoa, 2008).

Before proceeding, however, a brief historical note is in order. The emotion/cognition debate came into sharp focus with the report of the mere-exposure effect (Kunst-Wilson & Zajonc, 1980), which led to a strong belief that affect was primary to and independent of cognition. It can be said that the mere-exposure effect and other behavioral findings shifted ongoing debates to focus on affect as being related to unconscious processing and subcortical activity, with cognition being related to conscious processing and cortical involvement. Interestingly, behavioral findings were interpreted in the context of the “low route” suggested by LeDoux (1996), which was purported to carry affective information subcortically. These early behavioral studies provided a strong impetus to the wave of neuroscience research in the late 1990s (and beyond) that investigated related phenomena. For some of the early theoretical arguments, see Fazio et al. (1986), Leventhal & Scherer (1987), Bornstein (1989), Lazarus (1994), Zajonc (1994), and Bargh (1997); also see Storbeck, Robinson, & McCourt (2006) and Storberk (2008).

**Perception and attention**

Viewing emotion-laden visual stimuli is linked to heightened and more extensive visual system activation (Pessoa et al., 2002a; Vuilleumier, 2005). For instance, viewing faces with emotional expressions evokes increased responses relative to viewing neutral faces throughout ventral occipitotemporal visual cortex (Figure 1).

Visual responses are also stronger when subjects view emotional scenes (e.g., a war scene) compared to neutral scenes (e.g., a lake scene). Increased visual activation is observed in both late visual areas, such as the fusiform gyrus and superior temporal sulcus, and early visual cortex in occipital cortex. Recent studies suggest that, in humans, even retinotopically organized visual cortex, including visual areas V1 and V2 along the calcarine fissure, are modulated by the affective significance of a stimulus (Padmala and Pessoa, 2008).
Enhanced visual activation when viewing emotional stimuli is consistent with the observed improvements in behavioral performance in several visual tasks. For instance, angry and happy faces are detected faster in visual search tasks (Eastwood et al., 2001), and possibly other emotional stimuli, too, such as a snake or spider (Ohman et al., 2001) compared to neutral stimuli. Stronger evidence comes from studies of the attentional blink paradigm, in which subjects are asked to report the occurrence of two targets (T1 and T2) among a rapid stream of visual stimuli. When T2 follows T1 by a brief delay, participants are more likely to miss it, as if they had blinked (hence the name). The attentional blink is believed to reflect a capacity-limited processing stage, possibly linked to a process of consolidation of the detected item for conscious reports. Interestingly, the attentional blink has been shown to be modulated by emotional stimuli, as subjects are significantly better at detecting T2 when it is an emotion-laden word (e.g., rape) than when it is a neutral word (Anderson, 2005).

Converging evidence for a link between perception, attention, and emotion comes from additional studies. For example, patients with unilateral inattention due to spatial hemineglect (often as a result of right hemisphere parietal lesions) are better at detecting happy or angry faces compared to neutral ones (Vuilleumier and Schwartz, 2001). These findings are consistent with the notion that emotional faces may direct the allocation of attention. For instance, in one study, emotional faces were flashed at spatial locations that subsequently displayed low-contrast visual stimuli (Phelps et al., 2006). They found that detection of the target was strongest when the fear face served as the spatial cue, suggesting that emotional stimuli can provide additional attentional guidance above and beyond a generic spatial cue. How is the increase in perceptual processing and attentional capture that is observed during the perception of affective stimuli mediated in the brain? Growing evidence links the amygdala, a subcortical region, with these effects.

For instance, patients with amygdala lesions do not exhibit improved detection of T2 emotional targets during the attentional blink (i.e., a decrease in the magnitude of the blink effect) (Anderson and Phelps, 2001), and show less evidence of increased responses in visual cortex during the viewing of fearful faces (Vuilleumier et al., 2004); see Figure 2.

Thus, it appears that the amygdala may underlie a form of emotional modulation of information that in many ways parallels attentional effects that are observed with non-emotional information (Pessoa et al., 2002a; Vuilleumier, 2005). There are several ways in which emotional modulation may be accomplished. First, it is possible that direct projections from the amygdala to visual processing regions enhance visual processing. The...
The amygdala sends projections across all levels of the visual system, including anterior regions in temporal cortex and posterior regions in occipital cortex (including V1 and V2) (Amaral et al., 1992). Thus, the amygdala is well situated to modulate sensory processing according to the affective significance of a visual object. A second possibility is that the amygdala interacts with other brain regions that are important for the control of attention, such as frontal and parietal regions (Barbas, 1995), which, by their turn, modulate visual processing. In the latter scenario, the amygdala (possibly indirectly) would recruit attentional circuits so as to enhance the sensory processing of emotion-laden stimuli.

A final issue that should be addressed when considering interactions between emotion and perception/attention is whether the perception of emotion-laden stimuli is automatic, namely independent of attention and awareness. This question has received considerable attention because specific answers to this question (no or yes) suggest potentially different relationships between emotion and cognition (more or less independence between the two, respectively). Interestingly, evidence both for and against automaticity has been presented. For instance, emotional faces evoke responses in the amygdala even when attention is diverted to other stimuli (Vuilleumier et al., 2001; Anderson et al., 2003). Perhaps even more strikingly, amygdala responses are sometimes reported for emotional faces of which subjects are not conscious (Morris et al., 1998; Whalen et al., 1998; Etkin et al., 2004; Whalen et al., 2004). Furthermore, cases of affective blindsight have been reported. These and other related findings suggest that at least some types of emotional perception occur outside of cognitive processing – and may rely on direct subcortical pathways conveying visual information to the amygdala (LeDoux, 1996). At the same time, recent findings have suggested that the perception of emotion-laden items requires attention, as revealed by attentional manipulations that consume most processing resources, leaving relatively few resources for the processing of unattended emotional items (Pessoa et al., 2002b; Bishop et al., 2004; Pessoa et al., 2005; Bishop et al., 2007; Hsu and Pessoa, 2007; Lim et al., 2008). Furthermore, it also appears that amygdala responses evoked by unaware stimuli depend somewhat on the manner by which awareness is operationally defined (Merikle et al., 2001), such that no unaware responses are observed when awareness is defined, for instance, via signal detection theory methods (Pessoa et al., 2006). Overall, the automaticity debate remains unresolved and controversial (Pessoa, 2005; Wiens, 2006; Bishop, 2007).

Memory and learning

Research on classical fear conditioning suggests that the amygdala is involved in the acquisition, storage, and expression of a conditioned fear response – such as when an animal learns that a neutral stimulus (e.g., tone) predicts an aversive event (e.g., mild shock). Whereas fear conditioning is believed to involve a more primitive form of affective learning, instructed fear illustrates a situation in which cognition and emotion interact more explicitly (Figure 4C). In this paradigm, subjects are verbally informed of the possibility of an aversive event given the presentation of one type of neutral stimulus (e.g., tone), while the presentation of another neutral stimulus (e.g., light) indicates that the aversive event will not occur. Interestingly, instructed fear generates robust physiological results to the threat stimulus that resemble the responses to a conditioned stimulus (e.g., tone) in fear conditioning, even though the aversive event is never administered to the subjects (only a verbal threat occurs) (Hugdahl and Ohman, 1977; Phelps et al., 2001) (Figure 3D). Research with humans indicates that the left amygdala appears to be necessary for instructed fear (Funayama et al., 2001). Another example of cognitive-affective learning involves observational fear, in which an acquired fear response is learned via social observation (Figure 4B). In this case, both humans and nonhuman primates are capable of learning the affective
properties of stimuli through observing the emotional reactions of a conspecific (Ohman and Mineka, 2001). As in the case of instructed fear, observational fear results in the expression of conditioned fear that is similar to the one observed during fear conditioning (Olsson and Phelps, 2004) (Figure 3C).

Emotional content can change the formation and recollection of a memory event, consistent with findings in both human and animal studies. Compared to neutral items, humans remember better emotionally arousing information, including emotionally charged stories, film clips, pictures, and words. For instance, in one study participants viewed two videos, one composed of neutral film clips and another composed of emotional film clips (Cahill et al., 1996). Although the two types of clips were taken from the same source and were equated in terms of levels of understandability, subjects were better at remembering emotional relative to neutral clips when tested approximately 3 weeks following the initial viewing of the films. In another study (Bradley et al., 1992), subjects viewed a large array of emotional and neutral pictures from the International Affective Picture System, a stimulus set that has been normed in terms of the dimensions of valence (positive/negative) and arousal (calm/excited). Participants initially rated the pictures along the dimensions of valence and arousal. An incidental free-recall test was administered both immediately and at one year following the rating sessions. Pictures rated as highly arousing were remembered better than all other pictures, including those rated as moderately arousing. Interestingly, the pattern of results was very similar when the subjects were tested a year later, namely, highly arousing pictures were better remembered.

In humans, the amygdala is known to be a critical structure for the enhancement of memory by emotion, consistent with both lesion (Adolphs et al., 1997) and neuroimaging work (for a review, see Phelps, 2004). Recent studies have begun to delineate some of the specific functions of this structure. For instance, it appears that the right amygdala is more strongly involved in emotional memory formation, whereas the left amygdala is engaged by the retrieval of those memories (Sergerie et al., 2006), suggesting a potential hemispheric dissociation of amygdala involvement at different stages of emotional memory. In addition, amygdala responses are also linked to a novelty effect on memory tasks – i.e., the tendency to classify items as new as opposed to old (Sergerie et al., 2007).

In humans, there is some support for the notion that the enhancement of memory due to emotion is due mainly to the arousal dimension of emotional items and not valence (positive/negative) per se (Phelps, 2006), a notion that is more firmly established in nonhuman animal studies (McGaugh, 2004). In these studies, the effects of emotion on memory have

Figure 3: Fear learning in the human amygdala. (a) The outlined box contains the area of the medial temporal lobe that includes the bilateral amygdala. (b–d) Amygdala activation to the CS is seen bilaterally after fear conditioning (b) and observational fear learning (c), and unilaterally (d) in the left amygdala after instructed fear. Reprinted by permission from Macmillan Publishers Ltd: Nature Neuroscience (Olsson and Phelps, 2007),
been revealed by a vast array of experimental manipulations, including inhibitory avoidance training, contextual fear conditioning, cued fear conditioning, water-maze spatial and cued training, among others. Typically, the effects of emotion on memory are investigated via drug administration, including agonists and antagonists of specific brain receptors. For instance, in one experiment, rats were trained to swim to an escape platform after being placed in a water tank (Packard et al., 1994). To mimic the effects of arousal, a group of animals received an injection of d-amphetamine immediately after training; a control group received a saline injection. Behavioral testing revealed that d-amphetamine administration in the amygdala enhanced memory both on a spatial task and on a non-spatial cued task. A growing body of animal studies strongly supports a model in which emotion influences memory by modulating memory storage (McGaugh, 2004). In particular, the amygdala and the closely associated basal forebrain system involving the stria terminalis appear to play a major role in this modulatory process. These structures are thought to play a central role on memory consolidation by modulating activation in a network of brain regions, including the hippocampus, which is centrally involved in memory formation, but also additional brain structures, such as the nucleus accumbens, caudate nucleus, entorhinal cortex, in addition to other cortical regions (McGaugh, 2002) (Figure 5).

Behavioral inhibition and working memory

An important dimension of cognition involves behavioral inhibition. Response inhibition, namely the processes required to cancel an intended action, is believed to involve control regions in prefrontal cortex (e.g., dorsolateral prefrontal cortex, anterior cingulate cortex, and inferior frontal cortex) (Rubia et al., 2003; Aron et al., 2004). Response inhibition is often investigated by using so-called go/no-go tasks in which subjects are asked to execute a motor response when shown the go stimulus (e.g., press a key as fast as possible when you see a letter stimulus), but to withhold the response when shown the no-go stimulus (e.g., do not respond when you see the letter Y). Typically, the go and no-go stimuli are shown as part of a rapid stream of stimuli (e.g., a sequence of letters). A recent study investigated the interaction between the processing of emotional words and response inhibition (Goldstein et al., 2007). Response inhibition following negative words (e.g., worthless) engaged the dorsolateral prefrontal cortex. Interestingly, this region was not recruited by negative valence or inhibitory task demands per se; instead, the dorsolateral cortex was sensitive to the explicit interaction between behavioral inhibition and the processing of negatively valenced words.

Working memory, another important cognitive operation, involves the maintenance and updating of
information in mind when the information is no longer available to sensory systems (e.g., when keeping a phone number in mind for a few seconds before dialing the number). Evidence for cognitive-emotional interaction comes from working memory studies, too. For instance, when participants were asked to keep in mind neutral or emotional pictures, maintenance-related activity in dorsolateral prefrontal cortex was modulated by the valence of the picture, with pleasant pictures enhancing activity and unpleasant pictures decreasing activity relative to neutral ones (Perlstein et al., 2002). Interestingly, emotional pictures did not affect dorsolateral responses during a second experimental condition during which participants were not required to keep information in mind, indicating that the modulation of sustained activity by emotional valence was particular to the experimental context requiring active maintenance. In another study, participants watched short videos intended to induce emotional states (e.g., clips from uplifting or sad movies), after which they performed challenging working memory tasks (Gray et al., 2002). Bilateral lateral prefrontal cortex activity reflected equally the emotional and working memory task components (Figure 6). In other words, prefrontal activity did not stem from the working memory task alone or by the mood ensuing from the viewing of the video, but resulted from an interaction between cognition and emotion.

**Impact of cognition on emotion**

Although this short review focuses on the impact of emotional content on cognitive functions, here we briefly discuss another important line of studies that has investigated cognitive-emotional interactions, namely, cognitive emotion regulation (Ochsner and Gross, 2005; Ochsner and Gross, 2008). A particularly informative regulation strategy is “cognitive reappraisal”, which involves rethinking the meaning of affectively charged stimuli or events in terms that alter their emotional impact. Reappraisal appears to depend upon interactions between prefrontal and cingulate regions that are frequently implicated in cognitive control and systems like the amygdala and insula that have been implicated in emotional responding. Interestingly, having the goal to think about stimuli in ways that maintain or increase emotion may boost amygdala activity whereas having the goal to decrease emotion may diminish it. Furthermore, changes in emotional experience and autonomic responding may correlate with the concomitant rise or fall of prefrontal and/or amygdala activity. Although much of the work on the cognitive...
regulation of emotion has relied on a relatively strict separation between cognition and emotion, which are in this context viewed as engaged in tug-of-war for the control of behavior, this framework is likely overly simplistic. As proposed by Ochsner and Gross (2008), a more fruitful approach will entail developing an integrated framework for specifying what combinations of interacting subsystems are involved in emotional responding, as individuals exert varying degrees and types of regulatory control over their emotions.

Anatomical basis for cognitive-emotional interactions

In attempting to understand the relationship between emotion and cognition, it is important to consider anatomical information. Advances in our understanding of brain connectivity suggest that a given brain region is only a few synapses away from every other brain region (Sporns et al., 2004; Sporns and Zwi, 2004). Indeed, it appears that the brain is configured according to a small-world topology in which the path length between nodes is small – typically, cortical areas are connected directly or via just one or two intermediate areas (Hilgetag et al., 2000; Sporns et al., 2000) – and nodes are highly clustered (Sporns, 2006). Thus, a careful consideration of brain connectivity is informative in understanding potential cognitive-emotional interactions.

In the past decade, several quantitative analyses of brain connectivity have been undertaken (Young et al., 1994; Stephan et al., 2000). Not surprisingly, prefrontal areas are among those most distant from the sensory periphery, suggesting that they receive highly-processed and integrated sensory information. Such potential insulation of the prefrontal cortex from the periphery is thought to be a key anatomical feature of this region and presumably confers the primate brain with a greater degree of flexibility (Mesulam, 2002). Highly processed information would also be able to support more abstract processing that is required for cognition. Interestingly, the amygdala, a region often linked to emotional processing, appears to be equally removed from the sensory periphery – although in some species, direct sensory thalamic projections may be present (LeDoux, 1996). In addition, the amygdala makes very widespread projections. Overall, it appears that the amygdala is very well situated to integrate and distribute information (Figure 7).

It is also instructive to consider the connectivity of the hypothalamus (Risold et al., 1997), as it has been long recognized for its importance in emotional behaviours (Swanson, 2000, 2003). In particular, via its descending connections that innervate brainstem motor systems, this structure is thought to play a key role in the implementation of goal-directed behaviors. Hypothalamic signals also can be conveyed to the cortex, mostly by way of the thalamus. Critically, prefrontal cortical territories project directly to the hypothalamus. Thus, the
hypothalamus appears to be organized in such a way that it can generate both relatively reflexive behaviors and behaviors that are voluntarily triggered by inputs from the cerebral cortex (Swanson, 2000). Overall, this structure appears to be connected with all levels of the nervous system, including the neocortex (Swanson, 2000), enabling important hypothalamic regulatory signals to have widespread effects on the brain.

It is also important to consider the role of the *ascending systems*. For instance, the basal nucleus of Maynert is a major part of the so-called magnocellular basal forebrain system (Heimer and Van Hoesen, 2006). The projections from this system reach all parts of the cortical mantle (Heimer and Van Hoesen, 2006), and are involved in cortical plasticity in sensory cortex in the context of classical conditioning (Weinberger, 1995), in addition to arousal and attention mechanisms (see citations in (Sarter and Bruno, 2000; Heimer and Van Hoesen, 2006)). In particular, basal forebrain corticopetal cholinergic projections appear to be crucial for diverse attentional functions, including sustained, selective, and divided attention (Sarter and Bruno, 1999; Sarter et al., 1999; Sarter and Bruno, 2000). Of importance in the present context, the basal forebrain receives both cortical and amygdala inputs (for citations, see (Sarter and Bruno, 2000)). Notably, recent anatomical evidence suggests the existence of specific topographically organized prefrontal-basal forebrain-prefrontal loops (Zaborszky et al., 1999; Zaborszky, 2002; Zaborszky et al., 2005), so that specific prefrontal cortical targets of the basal forebrain connect back to sites from which the corticopetal fibers originate. Such loops provide a direct substrate for

Figure 7: Brain connectivity graph. Quantitative analysis of brain connectivity reveals several clusters of highly interconnected regions (represented by different colors). In this analysis, the amygdala (Amyg, centre of figure) was connected to all but 8 cortical areas. These connections involved multiple region clusters, suggesting that the amygdala is not only highly connected, but that its connectivity topology might be consistent with that of a hub that links multiple functional clusters. In this manner, the amygdala may be important for the integration of cognitive and emotional information. Figure labels represent different cortical areas with the exception of Hipp (hippocampus) and Amyg, which represent subcortical areas. Figure reproduced from Young et al. (1994) with permission from Freund Publishing House Ltd. Analysis of connectivity: Neural systems in the cerebral cortex, Reviews in the Neurosciences; copyright (1994).
cognitive-emotional integration, for example by allowing amygdala signals to be broadcast widely, including to frontoparietal regions known to be important for the control of attention. More generally, the overall anatomical arrangement of the basal forebrain may involve multiple functional-anatomical macrosystems (Alheid and Heimer, 1988; Zahm, 2006) with wide-ranging effects on brain computations and important clinical implications (Alheid and Heimer, 1988; Sarter and Bruno, 1999). In summary, the picture that emerges from anatomical connectivity data suggests a remarkable potential for integration of information.

Conclusion: from interactions to integration

Historically, emotion and cognition have been viewed as separate entities. One factor that may have contributed to this separation in the past century is methodological. For instance, data arising from single-unit or lesion studies usually allow the researcher to only derive conclusions concerning the specific areas being targeted. Research in the past two decades suggests, however, that such a view is likely deficient and that, in order to understand how complex behaviors are carried out in the brain, an understanding of the interactions between the two may be indispensable. Indeed, some studies have suggested that it may be important to go beyond understanding interactions, some of which are suggested to be mutually antagonistic, to understanding how cognition and emotion are effectively integrated in the brain. As stated recently, at some point of processing functional specialization is lost, and emotion and cognition conjointly and equally contribute to the control of thought and behavior (Gray et al., 2002). While these statements were offered as a summary of specific findings concerning working memory performance following mood induction (see above), they may aptly characterize a vast array of real-world situations. In other words, whereas many behaviors may be reasonably well characterized in terms of cognitive-emotional interactions such that emotion and cognition are partly separable, in many situations, true integration of emotion and cognition may also take place (Figure 8). The latter further blurs the distinction between cognition and emotion. See Duncan and Barrett (2007) for a similar view.

References

Figure 8: Potential relationship between anatomical sites, neural computations and behaviors. Brain areas (for example, A2), which are connected to form networks (ellipses), are involved in multiple neural computations (for example, NC2, NC3 and NC4) and specific computations (for example, NC4) are carried out by several areas (for example, A2 and A3). Therefore, the structure–function mapping is both one-to-many and many-to-one; in other words, many-to-many. Multiple neural computations underlie behavior. Each behavior has both affective and cognitive components, indicated by the affective and cognitive axes. Note that the axes are not orthogonal, indicating that the dimensions are not independent from each other. Brain areas with a high degree of connectivity (hubs) may be especially important for regulating the flow and interaction of information between regions. Reprinted by permission from Macmillan Publishers Ltd: Nature Neuroscience (Pessoa, 2008), copyright (2008). See (Mesulam, 1990) for a related scheme.


McGaugh JL (2004) The amygdala modulates the consolidation of memories of emotionally arousing

http://www.scholarpedia.org/article/Cognition_and_emotion


Storbeck J: The discussion in this paragraph draws on the accompanying review to this article by J. Storbeck.


**Recommended reading**


External links

- Luiz Pessoa's website (http://www.emotioncognition.org/)

See also

Cognition, Emotion

Sponsored by: Eugene M. Izhikevich, Editor-in-Chief of Scholarpedia, the peer-reviewed open-access encyclopedia
Sponsored by: Anil Seth, University of Sussex, UK
Reviewed by: Dr. Jorge Armony, Department of Psychiatry, McGill University, CANADA
Reviewed by: Anonymous
Accepted on: 2009-01-10 21:12:59 GMT

Categories: Cognitive Neuroscience | Consciousness