

Part I

Historical Background

Introduction

The computational theory of mind (CTM) emerged first in its “digital” form, then in its “connectionist” form. The purpose of Part I is to survey some of the historical antecedents to this emergence. Not all we will discuss contributed directly (or even indirectly) to the CTM, but it is all relevant to cognitive science, broadly construed. From these larger pictures we will try to extract a sketch of what these fields contributed directly to the CTM. We will see that the influences contributing to the shaping of cognitive science, and especially the computational theory of mind include the following. The connectionist form of the CTM is a descendant of both perceptrons (chapter 4) and associationism (chapter 1). After introducing connectionism we will return to associationism (chapter 12). From associationism and James we have the idea of (i) the conscious mind as an introspectable manipulator (association) of representations (ideas), and (ii) two levels of explanation: the introspective subjective psychological (software) and the objective neurological (hardware). From behaviorism we moved away from just introspection and into laboratory experimentation as practiced by current cognitive psychology. Information processing psychology gave us cognition as information processing, and more specifically in Miller, Galanter, and Pribram, TOTE units for explaining behavior are structured and function like computer programs. From biology we got most importantly the neuron doctrine: the nervous system and especially the idea that the brain is composed of networks of discrete units (neurons, axons, dendrites) joined together at synapses. The neuro-logical tradition, and especially McCulloch and Pitts, argued that the brain is composed of on/off units and circuits of such units that can be associated with the propositions of logic – the brain is equivalent to a machine table of a Turing machine, and if supplemented with unlimited memory is equivalent to a universal Turing machine. Perceptrons demonstrated that a computer hardware organized on the gross anatomy of the brain could be trained to discriminate certain categories of things in a broadly human way.

1

Associationism

1.1 Introduction: What is Associationism?

Associationism is the view that the mind is organized, at least in part, by principles of association. Associationists don't say just what makes a principle "associationist." Rather, they are content to state specific principles and call them "associationist" (the word gained currency with Locke, see below). But the *basic idea* behind associationism seems to be this: *items that "go together" in experience will subsequently "go together" in thought.* Typically, associationists are *empiricists* – they hold that all knowledge comes from experience both in the sense of being causally dependent on experience and in the sense of being justified solely by reference to experience. However, this is about where agreement ends, and each particular empiricist holds a doctrine slightly different from the others.

1.2 Generic Empiricist Associationism

These English psychologists – what do they really want? One always discovers them . . . seeking the truly effective and directing agent . . . in just that place where the intellectual pride of man would least desire to find it (in the vis inertiae of habit, for example, or in forgetfulness, or in a blind and chance mechanistic hooking together of ideas, or in something purely passive, automatic, reflexive, molecular and thoroughly stupid) – what is it really that always drives these psychologists in just this direction? Is it a secret, malicious, vulgar, perhaps self-destructing instinct for belittling man?

(Nietzsche, 1887)

Although empiricists differ in the details of their conception of the structure and operation of the mind, they can all be understood in terms of a common

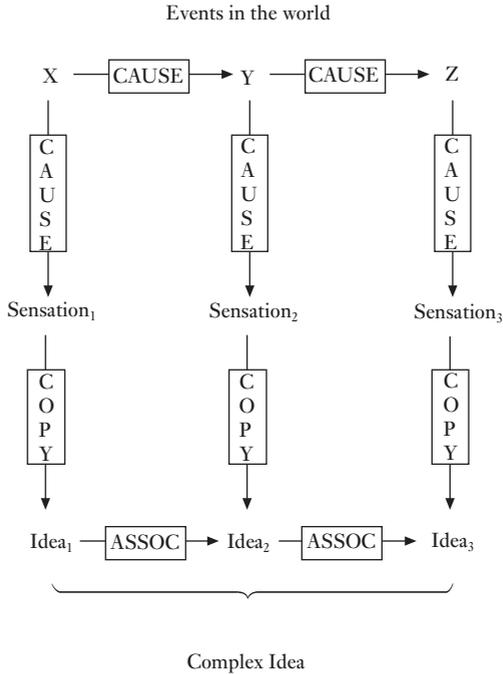


Figure 1.1 Generic empiricist associationism

overarching framework, one we call “generic empiricist associationism.” It fits no particular empiricist exactly, but gives instead a kind of composite photo of the movement (see figure 1.1).

Basic tenets

Figure 1.1 illustrates at least three basic tenets of generic associationism: (i) ideas, for instance (we return to this), are associated in the mind through experience; (ii) ideas can be decomposed into a basic stock of “simple” ideas, and ideas from the basic stock can be composed into more complex ideas; (iii) these simple ideas are derived from sensations. Sensations (sensory data) themselves are *not* governed by principles of association, but rather are caused by something outside the head (Hobbes, Locke, Hume: the world; Berkeley: God).

| Author | Date | Contiguity | Similarity | Contrast | Causality |
|------------------|--------|------------|------------|----------|-----------|
| Aristotle | 330 BC | X | X | X | |
| Thomas Hobbes | 1651 | X | | | |
| John Locke | 1700 | X | X | | |
| George Berkeley | 1733 | X | X | | |
| David Hume | 1739 | X | X | | (X)* |
| David Hartley | 1749 | X | | | |
| James Mill | 1829 | X | | | |
| John Stuart Mill | 1843 | X | X | | |
| Alexander Bain | 1855 | X | X | | |
| Herbert Spencer | 1855 | X | X | | |

*Subsequently reduced to contiguity.

Figure 1.2 Principles of association (from Marx and Hillix, 1963: 106, figure 8; reproduced by permission of McGraw-Hill Companies)

1.3 Varieties of Associationism

Pure associationism holds that only associationist principles govern the operation of the mind, whereas *mixed* associationism holds that there are non-associationist principles at work as well. For associationists there are different dimensions of “going together.” Items can go together by having certain relations in space (e.g. spatial contiguity), they can have certain relations in time (e.g. temporal contiguity). They can go together by having more abstract relations such as cause and effect, similarity, and contrast. Here are some typical associationist principles:

- 1 *contiguity*: items that are contiguous in space or time are linked by association;
- 2 *similarity*: items that are similar are linked by association;
- 3 *contrast*: items that contrast are linked by association;
- 4 *causality*: items that are linked by cause and effect are linked by association.

Even though associationists provided no principled reason why the list of relations couldn’t be extended indefinitely, there was remarkable agreement over the years as to what in fact the operative associationist principles are (see figure 1.2).

Associationist processes

For associationists, there are also three major processes of association. One kind of process involves *which items follow one another in time*, such as recall-

ing something from memory or the temporal order of thoughts. Another kind involves *compounding*, such as taking simpler items and building more complex ones. A final kind of process involves *decomposition* or taking complex items and breaking them down into simpler ones:

- 1 *sequencing*: associationist principles can govern such processes as: recalling items in memory, and the temporal order of thoughts;
- 2 *compounding*: complex items can be formed from simple items by (a) mental mechanics, (b) mental chemistry;
- 3 *decomposition*: complex items can be broken down into their simpler constituents.

With compounding there is a major difference between those who, like Locke, use a kind of “mental mechanics,” from those, like J. S. Mill (see below), who argue for a kind of “mental chemistry” as well.

The domain of associationism

Finally, different associationists think that the “items” involved in associations are quite different: memories, ideas, images, thoughts, and things were all suggested, used, and defended candidates.¹ With these general observations in place, let’s see how these ideas are played out by two of the most famous associationists.

1.4 Locke and James

At first philosophical and psychological studies of associationism were often hard to distinguish because philosophy and psychology were hard to distinguish. Although some investigations were clearly philosophical in their focus (David Hume) and others were clearly psychological in their focus (William James) there were some who were both (e.g., David Hartley who, like Locke, was both a medical doctor and a philosopher). Muddying the issue is the fact that some philosophers wrote like psychologists and some psychologists wrote like philosophers. However, it seems that with Hartley, associationism changed from being a component of empiricist epistemology to being basically a psychological doctrine on its own. Associationist theorizing has almost the same history as theorizing on the mind up until Hobbes. At about the same time Descartes popularized the doctrine of innate ideas (and relegated associationism to a purely neurological doctrine). Philosophical associationism reached its

peak with the British Empiricists (ca. 1700–1850) and although the most extensive associationist theorizing probably occurs in James Mill’s (1829) *The Analysis of the Phenomena of the Human Mind*, the most influential discussion of associationism for philosophy was probably David Hume’s (1739), *Treatise of Human Nature*. However, contemporary cognitive science seems to owe more to Locke and James than to any of the other players in the associationist tradition.

John Locke, An Essay on Human Understanding

John Locke (1632–1704) was a contemporary (and friend) of Boyle as well as Newton. He studied metaphysics and logic at Oxford, and had an affair there which he said “robbed me of my reason.” At the end of the affair his reason seems to have returned; he never married, and went on to produce his famous work on the theory of knowledge (his *Essay*, 1700, took 20 years to write), and political theory (which influenced the Declaration of Independence).

Ideas

For Locke, unlike Descartes (see chapter 3), there are no innate ideas: “Let us then suppose the mind to be, as we say, white paper, void of all character, without any ideas. How comes it to be furnished? . . . I answer, in one word, from *experience*. In that, all our knowledge is founded, and from that it ultimately derives itself” (*Essay*, bk 2, ch. 1, para. 2). Mental contents (ideas) are derived either through external experience, *sensation*, or from internal experience, *reflection*, on the operations of the mind itself. Sensation and reflection yield *simple* ideas upon which mental operations, such as recognizing similarity and differences or abstracting, creates *complex* ideas of substance, relation, etc.

The world

Sensation gives us ideas of qualities of external things and there are two important classes of qualities:

primary qualities (such as solidity, extension, figure, motion, rest, number) are essential to, and dependent on, only their bearers, and are independent of any perceiving mind;

secondary qualities are not essential to their bearers, and are the powers of objects (by configurations of primary qualities) to cause experiences (such as color, sound, taste, smell) in perceiving minds.

Ideas of the world

Locke is sure we believe in an external world and an internal self: “Sensation convinces us that there are solid, extended substances; and reflection, that there are thinking ones; experience assures us of the existence of such beings” (*Essay*: ch. 23, para. 15). But the above picture of our mental contents raises the question, how do we come to know anything about the world (or ourselves)? Locke’s metaphor for our predicament, and the direction out of it, is striking: “the understanding is not so much unlike a closet wholly shut from light, with only some little opening left, to let in external visible resemblances, or *ideas* of things without; would the pictures coming into such a dark room but stay there, and lie so orderly as to be found upon occasion, it would very much resemble the understanding of a man in reference to all objects of sight and the *ideas* of them” (*Essay*: ch. 11, para. 17). Note two things (which we will be returning to later): (1) ideas of sensation are analogized to *pictures* and have “visible resemblances” to external things; (2) ideas of sensation are *linked* to what they are about. At this point Locke presses the above distinction in qualities into service: ideas caused by external objects *do resemble primary qualities* – such ideas “are resemblances of them, and their patterns do really exist in the bodies themselves” (*Essay*: ch. 16, para. 16). But ideas caused by external objects *do not resemble secondary qualities* – the world itself contained no sweetness or blueness, only extension in motion. But what kind of link is needed for ideas to be about what they are about? His answer seems to be that it is a *causal* link: “these several appearances being designed to be the marks whereby we are to know and distinguish things which we have to do with, our *ideas* do as well serve us to that purpose and are as real distinguishing characters, whether they be only constant effects [secondary qualities?] or else exact resemblances [primary qualities?] of something in the things themselves: the reality lying in that *steady correspondence* they have with the distinct constitution of real beings . . . it suffices that they are *constantly produced* by them” (*Essay*: ch. 30, para. 2; last two emphases added). So this is how we break out of the dark room of our senses to the world about us – causation and resemblance. We will return to these themes soon, and in depth, in chapters that follow.

We can distinguish in Locke two general concerns: compounding and succession. It is not clear exactly what the domain of association is for Locke (though we will assume it is “ideas”), nor how general and pervasive are its principles. Locke acknowledges three general operations of the mind:

(1) composition, (2) setting ideas next to each other without composition (relations of ideas), and (3) abstraction (general ideas). The first relates especially to association.

Composition and complex ideas

Locke calls composition the process where the mind: “puts together several of those simple ideas it has received from sensation and reflection, and combines them into complex ones” (*Essay*: ch. 11, sect. 6). And in reverse: “All our complex ideas are ultimately resolvable into simple ideas, of which they are compounded and originally made up” (*Essay*: ch. 22, sect. 4). Here we see a kind of “mental mechanics” at work, where complex ideas are built out of simpler ideas like a wall is built out of bricks and mortar. It may be that associative principles such as similarity and contiguity are operative in composition, but if so, they are merely two of many principles and by no means hold sway over the process in general: “The mind . . . arbitrarily unites into complex ideas such as it finds convenient; whilst others that have altogether as much union in nature are left loose, and never combined into one idea, because they have no need of one name. It is evident, then, that a mind, by its free choice, gives a connection to a certain number of ideas, which in nature have no more union with one another than others it leaves out” (*Essay*: bk III, ch. 5, sect. 6). According to Locke, then, complex ideas need not always result from ideas which arrive together, and complex ideas can be formed “arbitrarily” and by “free will” – hardly associationist principles.

Succession of ideas

In the 4th edition of his *Essay*, Locke added a new chapter entitled “Of the association of ideas,” thereby giving a name to a doctrine, which name turns out to have been more influential than the original doctrine. His interest in the association of ideas (he also used “connection” of ideas) seems restricted almost completely to the pathological, that is, to mental breakdowns and he never names or formulates explicit principles of association. Like Hobbes before him and Hume after him he distinguishes two types of association of ideas² – those that have “a natural correspondence and connexion one with another,” and “wholly owing to chance or custom; ideas that in themselves are not at all of kins, come to be so united in some men’s minds, that ’tis very hard to separate them, they always keep in company, and the one no sooner at any time comes into the understanding but its associate appears with it.” Locke says little here about the first category, but he goes on to make a number of points about the second: (i) the mind makes these combinations either volun-

tarily or by chance (hence people exposed to the same environment can be very different psychologically); (ii) strength of the first impression or “future indulgence” (positive reinforcement?) can so unite ideas “that they always afterwards kept company together in the man’s mind as if they were but one *idea*”; (iii) some antipathies “depend upon our original constitution, and are born with us.” Locke’s main concern here is a concern with rectifying *wrong* associations – pedagogical, not psychological, analysis. That will change dramatically in the hands of probably the most distinguished and influential associationist, William James.

William James, The Principles of Psychology

The most readable treatment of associationism from within psychology proper was chapter 14: “Association,” of William James’s *The Principles of Psychology*, a book he contracted in 1878 to publish in two years, but which finally took him twelve years to finish.³ William James (1842–1910) is probably America’s most distinguished psychologist to date. He was ten years younger than Wundt, but the year Wundt went from Zurich to Leipzig, 1875, James got \$300 for “physiological” apparatus from Harvard: “It is conventional to say that Wundt founded the world’s first psychological laboratory at Leipzig in 1879, although Wundt himself had facilities for experimental demonstration at Leipzig soon after he arrived there in 1875. In short, both James and Wundt had informal demonstrational laboratories (not research laboratories) in 1875 and thereafter” (Boring 1929: 509). At age 19 James entered Lawrence Scientific School of Harvard after schooling abroad and a year of studying art. He studied chemistry and comparative anatomy, and after two years entered Harvard Medical School. At age 23 he accompanied Louis Agassiz on a naturalist expedition to the Amazon. He then went to Germany for a year and a half of medical studies. After that, he had a multi-disciplined career, as befits a contributor to cognitive science. In 1872 he was made instructor of physiology at Harvard College. In 1876 he was made assistant professor of physiology. In 1880 he was made assistant professor of philosophy, and in 1885 professor of philosophy. In 1889 he was made professor of psychology (just in time for his book).

Mental life: thinking

Psychology is the science of mental life, both its phenomena and of their conditions.

(The Principles of Psychology)

For the purposes of cognitive science, James's conception of our mental life or "thinking" (James: "I use the word thinking for every form of consciousness indiscriminately") has the following central features:

- 1 It is conscious. "Consciousness from our natal day is the teeming multiplicity of objects and relations. The only thing which psychology has a right to postulate at the outset is the fact of thinking itself. The first fact for us then as psychologists is that thinking of some sort goes on."
- 2 It is introspectable. "Introspective observation is what we have to rely on first and foremost and always." However, James, unlike some later writers, did not think introspection was incorrigible: "introspection is difficult and fallible . . . the difficulty is simply that of all observation of whatever kind."
- 3 It is private. "My thought belongs with *my* other thoughts, and your thought with *your* other thoughts . . . the only states of consciousness that we naturally deal with are found in particular consciousness, minds, selves, concrete particular I's and you's."
- 4 It "flows like a stream." What one introspects is an unbroken flow of ideas that follow one another according to principles of association: "Consciousness, then, does not appear itself chopped in bits. . . . In talking of it hereafter, let us call it the stream of thought, consciousness, or of subjective life."
- 5 It is about something ("intentional"). Thoughts have "ideas" as their constituents and ideas are about something or other.
- 6 It is evolutionary. Higher cognitive functions evolved because of their adaptive value.

Association

Background

Although James speaks occasionally, as the British empiricists did, of the mind compounding idea parts into complex wholes, he was on the whole skeptical of the doctrine of complex ideas. James also, paradoxically, claims explicitly that "objects are associated, not ideas" and he goes on to say: "We shall avoid confusion if we consistently speak as if *association*, so far as the word stands for an *effect* were between things thought of . . . not ideas, which are associated in the mind. . . . And so far as association stands for a *cause*, it is between *processes in the brain*" (*Briefer Course*: 5). This is not completely clear: what exactly is an effect of what here? Maybe we should think of it in the way shown in figure 1.3.

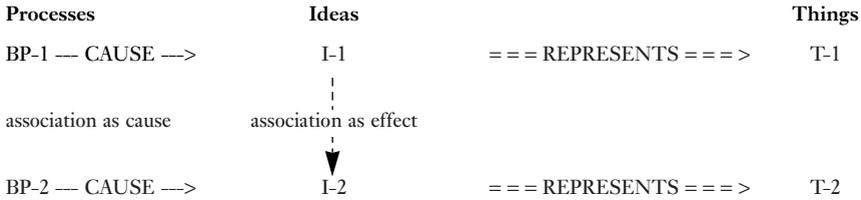


Figure 1.3 James on association, ideas, and things

Brain processes are the basic bearers of association. One brain process (BP-1) *causes* and becomes associated with another brain process (BP-2). But brain processes cause ideas (I) which are about or represent, objects, things (T) in the world that we think about, and by this means these things come to be associated – that is the *effect* of brain processes. Ideas, then, are the intermediary between brain processes and things – ideas both are caused by brain processes and represent these things.

Whatever exactly James meant by his remark, the real issue, he thinks, is accounting for the time course of thought: *how does the mind solve the problem of what to think next?* His general answer is that the sequencing of thoughts is in accordance with principles of association, and he suggests a variety of such principles including contiguity and similarity. But James never rests content with mere *descriptions* of patterns of association. He regularly presses for *explanations* at a deeper neural level. For instance, after formulating association by contiguity he says: “Whatever we name the law, since it expresses merely a phenomenon of mental habit, *the most natural way of accounting for it is to conceive it as a result of the laws of habit in the nervous system; in other words, it is to ascribe it to a physiological cause*” (*Principles of Psychology*: 561–2). “*The psychological law of association of objects thought of through their previous contiguity in thought or experience would thus be an effect, within the mind, of the physical fact that nerve currents propagate themselves easiest through those tracts of conduction which have been already most in use*” (*Principles of Psychology*: 563). And true to this explanatory strategy he postulates a pair of important, and prescient, neurological principles, the first for a pair of brain processes, the second for multiple brain processes:

- (P1) When two elementary brain-processes have been active together or in immediate succession, one of them, on reoccurring, tends to propagate its excitement into the other. (*Briefer Course*: 5)

Principle (P1) is, as we will see in chapter 3, similar to a principle later proposed by Hebb. The second principle is:

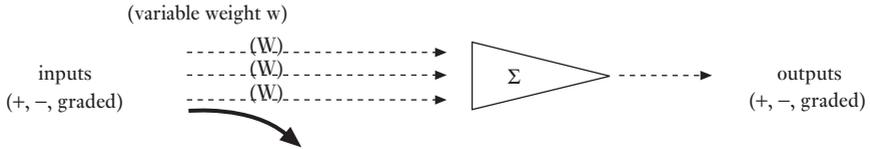


Figure 1.4 A James neuron

(P2) The amount of activity at any given point in the brain–cortex is the sum of tendencies of all other points of discharge into it, such tendencies being proportionate:

- 1 to the number of times the excitement of each other point may have accompanied that of the point in question;
- 2 to the intensity of such excitements; and
- 3 to the absence of any rival point functionally disconnected with the first point, into which the discharges might be diverted.

(Briefer Course: 5)

It might be useful here to diagram this second principle in terms of an imaginary “neuron,” which we dub the “James neuron” (see figure 1.4). In applying his associative principles James divides thought into two categories: *spontaneous* thought and *voluntary* thought (see figure 1.5).

Spontaneous trains of thought

Here we find three large categories of phenomena: total recall, partial recall, and focalized recall (association by similarity).

Total recall

This happens when there is unrestricted association between previous events and later recall. In James’s example a dinner party is followed by a brisk walk:

If a, b, c, d, e, for instance, be the elementary nerve-tracts excited by the last act of the dinner party, call this act A, and l, m, n, o, p be those of walking home through the frosty night, which we may call B, then the thought of A must awaken that of B, because a, b, c, d, e, will each and all discharge into l through the paths by which their original discharge took place. Similarly they will discharge into m, n, o, and p; and these latter tracts will also each reinforce the other’s action because, in the experience B, they have already vibrated in unison. The lines in Fig. 57 . . . symbolize the summation of discharge into each of the components of B, and the consequent strength of the combination of influences by which B in its totality is awakened. (Briefer Course: 6).

| Spontaneous thought | Voluntary thought |
|---|---|
| 1. Total (impartial) association [figures 57, 58] | 1. Recalling a thing forgotten [figure 61] |
| 2. Partial (mixed) association habit recency vividness emotional congruity [figure 59] | 2. Means–end reasoning [figure 61] |
| 3. Similarity association [figure 60] | 3. Generalization to all problem solving |

Figure 1.5 James’s taxonomy of the succession of thought

“Such processes as we have just described . . . would necessarily lead, if unobstructed, to the reinstatement in thought of the *entire* content of large trains of past experience” (*Briefer Course*: 6). (See figures 1.6(a) and 1.6(b).) Such massive and detailed association is not the norm.

Partial recall

Partial recall (see figure 1.6(c)) is the most common variety of association and in these cases only some of the past experiences have associational consequences; “In no revival of a past experience are all the items of our thought equally operative in determining what the next thought shall be. Always some ingredient is prepotent over the rest” (*Briefer Course*: 7). So the question arises as to which ingredient is prepotent and why. James’s answer is that “the prepotent items are those which appeal most to our INTEREST” (*ibid.*) “Expressed in brain–terms, the law of interest will be: *some one brain–process is always prepotent above its concomitants in arousing action elsewhere*” (*ibid.*). James surveys four principles of “interest” for determining “revival in thought”:

(1) *Habit* By this James means an association will favor elements that are most *frequent* in past experience: “Frequency is certainly one of the most potent determinants of revival. If I abruptly utter the word *swallow*, the reader, if by habit an ornithologist, will think of a bird, if a physiologist or medical specialist in throat–diseases, he will think of deglutition” (*Briefer Course*: 8).

(2) *Recency* James gives the example of a book, which habitually reminds him of the ideas it contains, but upon hearing of the suicide of the author, now remind him of death. He concludes: “Thoughts tend, then, to awaken their most recent as well as their most habitual [frequent] associates” (*Briefer Course*:

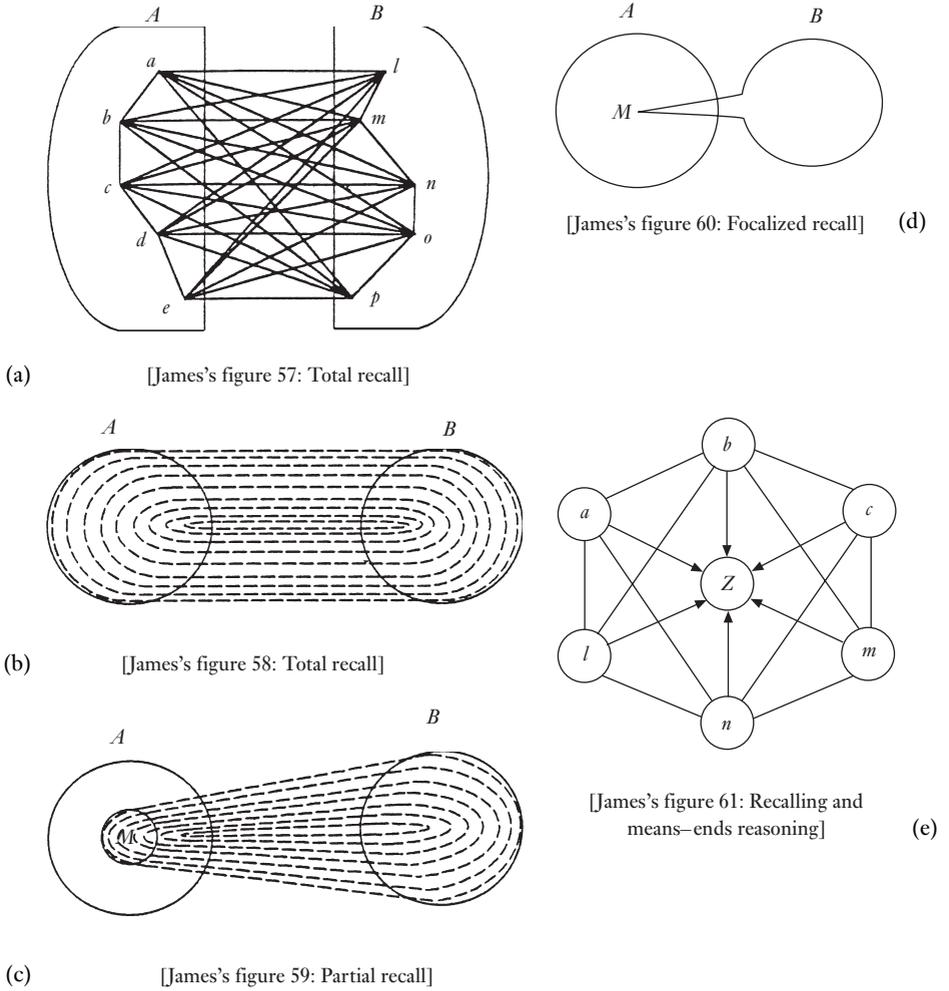


Figure 1.6 James's figures for the succession of thought (from *Briefer Course*)

- (a) James's figure 57: Total recall
- (b) James's figure 58: Total recall
- (c) James's figure 59: Partial recall
- (d) James's figure 60: Focalized recall
- (e) James's figure 61: Recalling and means-ends reasoning

8). And as usual, James tries to account for the phenomena at a lower level: “Excitement of peculiar tracts, or peculiar modes of general excitement in the brain, leave a sort of tenderness or exalted sensibility behind them which takes days to die away. As long as it lasts, those modes are liable to have their activities awakened by causes which at other times might leave them in repose. Hence *recency* in experience is a prime factor in determining revival in thought” (*Briefer Course*: 8–9).

(3) *Vividness* This is the strength or degree of an impression that the original experience carries and “*Vividness* in an original experience may also have the same effect as habit or recency in bringing about likelihood of revival” (*Briefer Course*: 9). For example: “If the word *tooth* now suddenly appears on the page before the reader’s eye, there are fifty chances out of a hundred that, if he gives it time to awaken any image, it will be an image of some operation of dentistry in which he has been the sufferer. Daily he has touched his teeth and masticated with them; this very morning he brushed them, chewed his breakfast and picked them; but rarer and remoter associations arise more promptly because they were so much more intense” (*ibid.*).

(4) *Emotional congruity* As for this, James writes: “A fourth factor in tracing the course of reproduction [in thought] is *congruity in emotional tone* between the reproduced idea and our mood. The same objects do not recall the same associates when we are cheerful as when we are melancholy. Nothing, in fact, is more striking than our utter inability to keep up trains of joyous imagery when we are in depressed spirits. . . . And those of sanguine temperament, when their spirits are high, find it impossible to give any permanence to evil forebodings or to gloomy thoughts” (*Briefer Course*: 9).

James sums up these four factors: “*Habit, recency, vividness, and emotional congruity* are, then, all reasons why one representation rather than another should be awakened by the interesting portion of a departing thought. We may say with truth that *in the majority of cases the coming representation will have been either habitual, recent, or vivid, and will be congruous*” (*Briefer Course*: 9).

Notice that although James *labels* these associational principles (APs), and gives us examples of them, he never explicitly *formulates* them. What might such a principle look like? James never says, but if such principles are supposed to control the time course of thought, they might look like these:

(AP1) If the subject is entertaining the thought A, and A is associated by habit (frequency) with thought B, then the subject will next think thought B – unless this association is overridden by some stronger principle of association.

(AP2) At any given time, the strongest principle of association is the operative one.

James also does not distinguish the fourth principle (emotional congruity) from the earlier three, yet it is possible that it is really quite different in that it does not seem to associate any particular thought (B) with any other particular thought (A). It says that a whole class of thoughts is more likely to be called up than the rest – the class of thoughts that are similar in emotional value.

Focalized recall, or association by similarity

This sort of association turns on a similarity between parts of things, shared qualities or relations (see figure 1.6(d)): “let us suppose that selective agency of interested attention . . . refined itself still further and accentuates a portion of the passing thought, so small as to be no longer the image of a concrete thing, but only of an abstract quality or property. Let us moreover suppose that the part thus accentuated persists in consciousness (or, in cerebral terms, has its brain process continue) after the other portions of the thought have faded. *This small surviving portion will then surround itself with its own associates* after the fashion we have already seen, and the relation between the new thought’s object and the object of the faded thought will be a *relation of similarity*. This pair of thoughts will form an instance of what is called ‘*Association by Similarity*.’ . . . Similarity, in compounds, is *partial identity*. When the *same* attribute appears in two phenomena, though it be their only common property, the two phenomena are similar in so far forth” (*Briefer Course*: 9–10). James’s example here is first thinking of the moon, then a gas flame (by similarity of color), then a football (by similarity of shape). Note that the moon and the football share no relevant associated property themselves.

Voluntary trains of thought

James wants to extend his associationist account of the time course of “spontaneous trains of thought” to “voluntary trains of thought”: “Hitherto we have assumed the process of suggestion of one object by another to be spontaneous. . . . This is reverie, or musing; but great segments of the flux of our ideas consist of something very different from this. They are guided by a distinct purpose or conscious interest; and the course of our ideas is then called *voluntary*” (*Briefer Course*: 11). As usual, James also redescribes this at the physiological level: “Physiologically considered, we must suppose that a purpose means the persistent activity of certain rather definite brain-processes. . . . This interest is subserved by the persistently active brain-tracts we have supposed” (*ibid.*). At the physiological level the crucial difference between spontaneous and vol-

untary sequences of thought is that the latter involves persistently active neural processes while the former does not.

Voluntary thought is traditionally a stumbling block for associationist theories since it would seem that here, if anywhere, rational, logical procedures can occasionally prevail over associative links. James approaches the question in two stages. First he tries to account for “recalling a thing forgotten” in associationist terms. Then he tries to extend this account to problem solving. James poses the issue of voluntary thought in terms of problems and their means of solution: “But in the theoretic as well as in the practical life there are interests of a more acute sort, taking the form of definite images of some achievement which we desire to effect. The train of ideas arising under the influence of such an interest constitutes usually the thought of the *means* by which the end shall be attained. If the end by its simple presence does not instantaneously suggest the means, the search for the latter becomes a *problem*; and the discovery of the means forms a new sort of end . . . an end, namely, which we intensely desire . . . but of the nature of which . . . we have no distinct imagination whatever” (*Briefer Course*: 11). Thus problem solving is pictured as predominantly means–end reasoning. James immediately extends this: “The same thing occurs whenever we seek to recall something forgotten” (*ibid.*). “The desire strains and presses in a direction which it feels to be right, but towards a point which it is unable to see. In short, the *absence of an item* is a determinant of our representations quite as positive as its presence can ever be” (*ibid.*). As usual, James tries to redescribe this at the physiological level: “If we try to explain in terms of brain-action how a thought which only potentially exists can yet be effective, we seem driven to believe that the brain tract thereof must actually be excited, but only in a minimal and subconscious way” (*ibid.*). James thinks that both kinds of problem have a common structure: “Now the only difference between the effort to recall things forgotten and the search after the means to a given end is that the latter have not, whilst the former have, already formed a part of our experience” (*ibid.*).

Recalling a thing forgotten

In the case of recalling a thing forgotten: “The forgotten thing is felt by us as a gap in the midst of certain other things. . . . We recollect the general subject to which it pertains.” James schematizes the process of recalling a thing forgotten as shown in figure 1.6(e).

James explains: “Call the forgotten thing *Z*, the first facts with which we felt it was related *a*, *b*, and *c*, and the details finally operative in calling it up *l*, *m*, and *n*. Each circle will stand for the brain processes principally concerned in the thought of the fact lettered within it. The activity of *Z* will first be a mere tension; but as the activities in *a*, *b*, and *c* little by little irradiate into *l*,

m , and n , and as *all* these processes are somehow connected with Z , their combined irradiations upon Z , represented by the centripetal arrows, succeed in rousing Z also to full activity” (*Briefer Course*: 12).

Problem solving: means–end reasoning

James conceives of problem solving as related to recall. Going back to figure 1.6(e) (James’s figure 61), he says: “The end here stands in the place of a , b , c , in the diagram. It is the starting point of the irradiations of suggestion; and here, as in that case, what the voluntary attention does is only to dismiss some of the suggestions as irrelevant, and hold fast to others which are felt to be more pertinent – let these be symbolized by l , m , n . These latter at last accumulate sufficiently to discharge all together into Z , the excitement of which process is, in the mental sphere, equivalent to the solution of the problem. The only difference between this case and the last [recalling something forgotten] is that in this one there need be no original sub–excitement in Z , co–operating from the very first” (*Briefer Course*: 12). And James concludes, generalizing (hastily): “*From the guessing of newspaper enigmas to the plotting of the policy of an empire there is no other process than this. We trust to the laws of cerebral nature to present us spontaneously with the appropriate idea*” (ibid.; emphasis added).

James never addresses doubts one might have as to whether all reasoning is means–end reasoning as he described it (nor whether he has described all means–ends reasoning correctly). Consider the problem of balancing one’s checkbook: one adds up columns of numbers, subtracts others, compares results, etc. Are these processes like (James’s version of) means–ends reasoning? When adding a column of numbers and carrying a “1,” do we voluntarily attend to associates and wait for the requisite associate to pop into consciousness? It would seem not. Furthermore, the solution in this case is better described as the end, not the means – the means being the principles of arithmetic. Here again we see Hobbes’s early distinction between associative processes and “calculation” at work. Perhaps James’s theory is appropriate only for the former, and it was a mistake for him to generalize to all problem solving. Note that James is skeptical about the possibility of a complete account of such reasoning: “It is foreign to my purpose here to enter into any detailed analysis of the different classes of mental pursuit. In scientific research we get perhaps as rich an example as can be found. . . . No rules can be given by which the investigator may proceed straight to his result. . . . But the final stroke of discovery is only prepared, not effected by them [associations]. The brain–tracts must, of their own accord, shoot the right way at last, or we shall grope in darkness . . . we are at the mercy of the spontaneous workings of Similarity in our brain” (*Briefer Course*: 12–13).

James is also skeptical about a complete brain science: “The *elementary* process of revival can be nothing but the law of habit. Truly the day is distant when physiologists shall actually trace from cell-group to cell-group the irradiations which we have hypothetically invoked. Probably it will never arrive” (*Briefer Course*: 13). James sums up as follows, again returning to the neural level: “To sum up, then, we see that *the difference between the three kinds of association reduced itself to a simple difference in the amount of that portion of the nerve-tract supporting the going thought which is operative in calling up the thought which comes*” (ibid.). “The order of *presentation of the mind’s materials* is due to cerebral physiology alone” (ibid.). Thus, the overall thrust of James’s analysis of the sequencing of thoughts is to *describe* them in terms of associationist principles, then try to *explain* their occurrence in terms of neurological principles.

1.5 The End of Classical Associationism

At least three factors led to the downfall of classical associationism in psychology in the late nineteenth and early twentieth centuries. Each factor was away from British empiricists’ focus on the introspectable association of ideas. *First*, there was the bankruptcy of the *introspective methodology* itself. By the end of the nineteenth century it came to be characterized by endless squabbles with no method for resolving them. After the establishment of Wundt’s laboratory in Leipzig (1879) psychology was becoming a science and there was little or no perceived place in it for the impressionistic literary style that heaped unverifiable anecdote upon unverifiable anecdote. It was time to close the “bloodshot inner eye” of introspection. *Second*, there was a set of related factors that led eventually to *behaviorism* and *stimulus-response* psychology. Perhaps the initial development was Ebbinghaus’s (1885) work on the association of stimulus and response in learning nonsense syllables. This can be considered one of the first laboratory applications of associationist principles, and at the same time one of the first steps in the creation of a science of experimental psychology related to learning, memory, and thought. This was followed by Thorndike’s (1911) work on animal learning, which paved the way for Watson (1913) and the behaviorist movement of the 1930s and 1940s, and for the acceptance of the importance of Pavlov’s (1927) work on the conditioned reflex (we turn to this in the next chapter). *Third*, there was Hartley’s and James’s systematic appeals to the *neural* level of explanation for psychological phenomena. This explanatory strategy, coupled with the development of neuroscience (and “physiological psychology”), with its techniques and

theories for studying the nervous system (Golgi, Cajal, Sherrington), called into question the desirability of (purely) psychological principles at all (we turn to this in chapter 3). With the demise of introspectionist methodology came the demise of the objects of introspection – ideas. The new elements of mind were stimuli and response, and their neural substrata – not introspectable at all. And as ideas were replaced by stimuli and responses, introspection was replaced with laboratory experimentation. There was also the increased prominence of reinforcements, reward, and conditioning – procedures rarely discussed by the British empiricists.

Notes

- 1 As we will see, it is William James's official position that it is things (out in the world), not ideas, that are associated.
- 2 "Ideas" for Locke, unlike Hume later, cover all mental contents: ideas of sensation and of reflection.
- 3 A shorter version occurs in chapter 16 of William James's *Psychology (Briefer Course)*. Figure numbers are those of the *Briefer Course*.

Study questions

What is associationism?

What is a general statement of what makes a principle "associationist"?

What is "pure" vs. "mixed" associationism?

Are sensations governed by principles of association – why/why not?

What two principles of association did most associationists subscribe to?

What are the two major types of mental processes that associationist principles are supposed to account for?

What is association by contiguity?

What is association by similarity?

What is association by contrast?

What is association by cause and effect?

Locke and James

What were Locke's main contributions to associationism?

What is psychology the study of, according to James?

- What six features characterize our mental life, according to James?
- What associational process is the main focus of James's work?
- What level of explanation is basic for James?
- What is James's view on the likelihood of a complete physiological account of thought?
- What basic principle governs association involving two brain processes active together?
- What basic principle governs association involving multiple brain processes active together?
- What does "association as an effect is a relation between things" mean?
- What does "association as a cause is a relation between brain processes" mean?
- What is the role of interest in spontaneous thought?
- What are the labels James gives to the four principles of partial recall?
- Give an example (from James) of each.
- Try to state these explicitly as associational principles.
- What is association by similarity? Give James's example.
- What does James mean by saying that similarity is partial identity?
- What distinguishes voluntary from spontaneous thought?
- Into what two subcategories does James divide voluntary thought?
- What is the main difference between recalling something forgotten and means–ends reasoning?
- What kind of reasoning seems to pose a problem for James, and why?
- What three factors contributed to the end of classical associationism?

Suggested reading

General: The single most complete survey of associationism is Warner (1921), which is obviously a bit dated and which, curiously, does not discuss James. Boring (1929), chapter 10, covers British empiricism, and chapter 12 covers the Mills and Bain. Marx and Hillix (1963), chapter 6, covers both traditional associationism and early behaviorism, which it treats as associations between stimuli and responses. For a more contemporary perspective, see the introduction to Anderson and Bower (1974).

For more on *Locke* see Cummins (1989), chapter 4, and McCulloch (1995), chapter 2, for more on Locke on representation. For more on *James* see Flanagan (1991), chapter 2, contains an excellent discussion of James's philosophy of mind and psychology from a cognitive science perspective, and some of our general remarks follow his. For some *other empiricists* we did not cover: on Hume, Wilson (1992) is a particularly relevant study of Hume (see references therein). On Bain, Young (1970), chapter 3, contains a discussion of Bain from a contemporary point of view. A recent *selection of associationist* writings can be found in Beakley and Ludlow (1992), part IV. Hunt (1993), chapter 3, contains a readable brief survey of empiricist and rationalist psychological doctrine, and chapter 6 contains a general discussion of James.