5 Can Connectionist Networks Think?

The History and Status of the Debate — Map 5 of 7

An Issue Map™ Publication

Start Here

1 Alan Turing, 1950
Yes, machines can (or will be able to) think. A computational system can possess all important elements of human thinking or understanding.

I believe that at the end of the century ... one will be able to speak of machines thinking without expecting to be contradicted.

Alan Turing

Key: Icons for Conceptions of Thought

The following icons are used throughout this map to symbolize different conceptions of thinking:

- Symbolic accounts of the mind
- Connectionism
- Neuroscience

2
Connectionism networks can think. Connectionist networks can possess all important elements of human thinking or understanding. Connectionist networks are characterized by:
- an ability to learn via training, rather than explicit programming
- a "brain-like" parallel processing
- neural realism, or at least neural inspiration
- third tolerance of noisy or incomplete data
- superior performance on perceptual and motor tasks

Note: This general characterization of connectionism is intended to highlight those aspects of the field that are relevant to this map. Few connectionists would actually claim that "connectionist networks can think," because connectionist ideas are usually regarded as *simulations* of neural networks in the brain, which is where the real thinking is understood to take place.

Thinking = **Activity of a connectionist network**

Postulates of Connectionism

Sample Node

Sample Output Functions

1. There is a set of elements called *units* or *nodes*.
2. Nodes pass *signals* (numerical values) to each other.
3. Nodes are connected into networks, which resemble neural networks in the brain (see sidebar, "Connectionism And The Brain," on this map).
4. When a set of signals reaches a node, they are multiplied by *weights* and then are added together to produce an *activation value*.
5. The activation value is then passed through an *output function*, which decides what final value a unit should output.
6. Vectors (sets of numeric values) are the basic representational medium of a connectionist network. Vectors enter the network as input, are processed through vectors (and matrices) of weights and connections, and new vectors are produced as a result.
7. Neural networks are trained to compute *vector-to-vector functions*. That is, they learn to convert specific input vectors to specific output vectors.
8. Learning takes place at the weights, which are adjusted using a variety of learning procedures, generally involving exposure to a *corpus* of sample inputs (see sidebar, "Learning in Connectionist Networks," on this map).

If activation is above some level, fire one value, otherwise fire another value.

A continuous version of the threshold function, which outputs one of a range of values rather than just one of two values.

Connectionist Network Computing Vector-to-Vector Function

Proponents include David Rumelhart, Paul Smolensky, James McClelland, Geoff Hinton, Jerry Feldman, Paul and Patricia Churchland, Terrence Horgan, John Tienson, and David Touretzky. Other notable proponents include Jeff Elman, Stephen Grossberg, and Terrence Sejnowski.

History of Connectionism

Several historical precursors to connectionism are cited in the literature. It is sometimes claimed that Aristotle, with his focus on learning and intuition, was a proto-connectionist. The works of the British Empiricists (John Locke, George Berkeley, and David Hume) are also variously ascribed as precursors to connectionism, given that associative links between ideas are similar to weighted links between nodes. (Note that the Empiricists are also proposed as precursors to classical AI; see sidebar, "History of the Symbolic Data Assumption," on Map 3).

Formal plausibility of neural networks was pioneered by biologists and cybernetricists in the 1940s and 1950s. Warren McCulloch and Warren Pitts (1943) developed a "logical calculus" of neural activity, which showed that neurons (like elements could compute logical functions (see "The Logical Calculus of Neural Activity," Map 3, Box 6).

Donald Hebb (1949) proposed that learning in neural networks takes place where 2 connected units are simultaneously active. Frank Rosenblatt (1959) may have been the first neural connectionist. His *perceptron* (one-layer feedforward network) could learn to recognize patterns reasonably well, and he developed early forms of several modern learning schemes. Rosenblatt also introduced the use of computer simulation into neural network theory.

From the late 1950s to the 1970s there was a lull in connectionist research, which resulted from the early success of AI combined with Marvin Minsky and Seymour Papert's critique of perceptrons, which was developed throughout the 1960s (but not published until 1969). Their critique showed that one-layer networks couldn't compute certain kinds of functions (see "One-Layer Perceptrons Can't Compute Certain Functions," Box 10).

Connectionism began to resurface in the 1970s. One of the classic papers of the new generation was Stephen Grossberg's "Adaptive Pattern Classification and Universal Recoding: Parallel Development and Coding of Neural Feature Detectors" (1976), which introduced the term connectionism, outlined some of the basic properties of connectionist models, and argued for their superiority over classical AI systems. The major publication of this new generation was David Rumelhart, James McClelland, and the PDP Research Group's two-volume work, *Parallel Distributed Processing* (1986), which sold out upon publication and which rifled many AI researchers to the connectionist camp. The writers described modern connectionist research in detail and argued for the superiority of connectionism over classical symbolic AI.

Renewed interest in connectionism was answered by the classical AI camp with a series of polemics, in particular an special issue of *Cognition* (1988) which contained 3 long critiques of connectionism by prominent cognitive scientists. The gist of the 3 *Cognition* arguments is the same: connectionism is a revival of associationist theories, which are cognitively weak; such theories can't account for the kind of systematicity and generativity that symbolic accounts explain so well; and, if connectionism can account for such phenomena, it offers a "mere implementation" of the classical view (see "The Past-Tense Debate," Map 4, Box 12, "Logic and Rule-Based Explanation," Box 24, and "The Connectionist Dilemma," Box 1).

The debate between connectionists and classicists, which has been called a "holy war" and a "battle to win souls," continues. Although the debate has often been heated, correct connections have increasingly towards "reconciliation" views, which try to give credit to the virtues of both classical and connectionist approaches (see sidebar, "Spectrum of Positions," on this map).

Further history is contained in Rumelhart and McClelland (1986b, pp. 41–44), and in Bechtel and Abrahamsen (1991, pp. 594–596).

Learning in Connectionist Networks

Connectionist networks learn by incrementally adjusting their weights in response to a corpus of sample inputs. The networks are repeatedly fed these inputs until they have been trained to perform as desired. Training takes place via a learning algorithm.

Learning is a central emphasis in connectionism, which has a vast technical literature on the various connectionist learning algorithms. The most widespread learning procedure is the backpropagation algorithm, which changes the weights of a feed-forward network so as to minimize difference between input. Other important learning algorithms include competitive learning and the Boltzmann machine learning algorithm.

Unmapped Territory
- Additional learning arguments

Are connectionist networks vulnerable to the arguments against physical symbol systems?

3 John Searle, 1990b; Roger Penrose, 1989
Connectionist networks are formal systems. Any function that can be computed by connectionist networks can be computed by a Turing machine, which means that connectionist networks are simulated on serial machines. Conversely, connectionist networks can be used to implement classical serial processing. Thus, arguments directed against the formal character of symbol manipulators apply equally well against connectionist networks.
Supported by
"The Chinese Room Argument," Map 4, Box 3;
"Mathematical Insight is Non-Algorithmic," Map 7, Box 23.

Note: The point that connectionist networks and symbol systems can simulate each other is widely accepted. Searle and Penrose use this point to show that their Chinese Room and Gödel arguments, respectively, apply to connectionist networks as well as to classical AI systems.

4 Jack Copeland, 1993
Simulations of connectionist networks are not duplications. It is invalid to argue from the fact that a connectionist network (simulated on a serial network) can't think to the conclusion that an actual connectionist network can't think. In fact, the invalidity can be demonstrated by modifying one of Searle's own arguments about simulation (see "Simulations Are Not Duplications," Map 2, Box 23).

3 Invalid Inferences

Searle argues that this is invalid reasoning:	But that implies that the contrapositive form is also invalid:	Thus, Searle's connectionism argument is also invalid:
1. An actual storm can make us wet. 2. Therefore, a simulation of a storm can make us wet.	1. A simulation of a storm can't make us wet. 2. Therefore an actual storm cannot make us wet.	1. A serial simulation of a connectionist network can't think. 2. Therefore, an actual connectionist network can't think.

5 John Searle, 1990a
The Chinese Gym argument. Suppose the basic Chinese Room were expanded to a large gym full of people speaking English-speaking men who have been given rules for the node to node passing of signals (from English speakers to men that tell them what symbols to pass to one another). Through this means, the men would simulate the connectionist network and would be able to produce Chinese speech, but none of the men understand Chinese. This example shows that instantiating a connectionist network is not enough to produce an understanding of Chinese.
Note: Also, see Map 4, In particular, "The Chinese Water Pipe Brain Simulator" (Map 4, Box 5) parallels this argument, with the difference that the Chinese water pipe argument uses one mini-instead of a gym full of men.

6 Jack Copeland, 1993
The systems reply to the Chinese Gym. Searle's Chinese Gym argument commits the same fallacy as the Chinese Room argument consists. It is invalid to infer that the gym as a whole doesn't understand Chinese from the fact that the individuals in the gym don't understand Chinese.
Note: Also, see "The Chinese Room, considered as a total system, think?" arguments on Map 4.

7 Paul and Patricia Churchland, 1990
The Chinese Gym requires a prepostorous number of people. Searle's thought experiment could not reasonably be implemented. His gym would have to hold the entire populations of more than 10,000 Earths.

Are connectionist networks like human or neural networks?

13
The connectionist biological assumption. Connectionist networks are similar to real neural networks.
Note: Also, see the "Is the brain a computer?" arguments on Map 7, Box 7, and the "Relation between hardware and software similar to that between human brains and minds?" arguments on Map 7.

14 Christine Skarda and Walter Freeman, 1987
Connectionist networks are too simple. Real neural networks, observed in living organisms such as rabbits, exhibit complex dynamics and chaotic activity that connectionist networks lack. These complex dynamics seem too messy from an engineering standpoint but are critical to an understanding of neural dynamics.

is similar to

16 Jerry Fodor and Zenon Pylyshyn, 1988
"Brain-style" modeling can be misleading. Brains are flexible, processing focused on facts about the brain can be misleading. Although connectionism seems useful, it has led to the revival of such weak psychological theories as:
- associationism
- microfeature analysis
- statistically based learning

One should be deeply suspicious of the "brain modeling" that purports to address the mysteries of cognition (p. 64).

Fodor and Pylyshyn

17 Jerry Fodor and Zenon Pylyshyn, 1988
Facts about the brain may be irrelevant to facts about thinking. Brains and minds are not the same and there are other disanalogies. For example, nodes and rivers have little in common with the atoms they are constructed of. Similarly, thinking may have little in common with the neural structures it is implemented in. So, we should be cautiful about basing a theory of cognitive architecture on a theory about the brain.

is not similar to

Other connectionist arguments

9 Hubert Dreyfus, 1992
Connectionist computers lack a commonsense background. Connectionist networks are unable to make generalizations and classifications in the way human beings do, because they lack commonsense understanding of the world. For example, one of the array's early connectionist networks was trained to distinguish photos of tanks from photos of empty landscapes. The results were initially promising, but it was later discovered that the network was distinguishing inappropriately—by detecting whether or not there were clouds in the pictures with and without tanks had been taken on different days.
"Cons to the symbolic arguments applied to the symbol systems architecture, see the "Can a symbolic knowledge base represent human understanding?" arguments on Map 3.

Hubert Dreyfus

10 Marvin Minsky and Seymour Papert, 1969, revised 1988; David Rumelhart, James McClelland, and PDP Group, 1986
One-layer perceptrons can't compute certain functions. One-layer perceptrons can't compute certain functions, including:
- parity—whether an odd or even number of units in the input layer are active
- connectedness—whether all active units in the input layer are connected to all other active units (either directly, or via other units that are active).

11 David Rumelhart and James McClelland, 1986
Multilayer perceptrons can compute all relevant functions. The limitations described by Minsky and Papert apply only to multilayer networks. Minsky and Papert ignored such networks in part because there were no useful learning procedures for multilayer nets. Today, however, there are several effective algorithms for training multilayer networks.

12 Tim van Gelder and Robert Port, 1995
Connectionism fall into a computational mindset. Connectionists take some steps in the right direction, but they fall into a computationalist mindset. For example, when they substitute activation patterns for symbols, Connectionists should focus more on the dynamical systems approach, which views the mind as a complex system that evolves through time. Connectionists need to "take the leap out of the computational mindset and into time" (p. 3).

Are connectionist networks capable of the kind of systematicity and productivity that characterizes human thinking?

31 Tim van Gelder, 1990
The connectionist dilemma. The connectionist approach to cognitive science is impaled on the horns of a dilemma—it is either inadequate as a theory of the mind, or else it is just an implementation of the classical architecture.

Either	Or
Connectionism can't account for systematicity and related phenomena (productivity, compositionality, and inferential coherence),	Connectionism accounts for systematicity and related phenomena by using structured representations.
in which case	*in which case*
connectionism is inadequate as a theory of cognition.	connectionism is a mere implementation of the classical architecture.

In Either Case

"The mind cannot be, in its general structure, a connectionist network" (Fodor and Pylyshyn 1988, p. 33), and the classical symbolic paradigm is still the best explanation of mind.

44 Jerry Fodor and Zenon Pylyshyn, 1988
Connectionism is associationism. Processing units in a connectionist network are usually connected by associative links. But associationist theorists have long argued that such links can account for relationships between ideas, but not for phenomena. Symbolic AI, however, can explain systematicity with reference to structured representations and structure sensitive thought processes.

Unmapped Territory
- Additional associationism arguments

45 Brian McLaughlin, 1993a
The burden of proof is on connectionism. The burden of proof is on the connectionists to explain systematicity and related phenomena without implementing a classical architecture.

46 Daniel Dennett, 1991b
Cognition isn't always systematic. Not all organisms exhibit systematicity of thought processes. For example, there are creatures that can see some sense (from "This fox is going to eat me!" but cannot think "I am going to eat the fox.")

47 Robert Matthews, 1994
Three-concept monte. By the way they pose their challenge to connectionism, Fodor and Pylyshyn (and later, McLaughlin) make it hard for connectionists to reply to their challenge. Fodor and Pylyshyn start with talk of systematicity and implementation, they conceal the crucial issue of explanation. It turns out that the only explanations classicists are willing to accept are classical explanations.

Explanation

48 David Braddon-Mitchell and John Fitzpatrick, 1990
Systematicity can be explained by natural selection. Systematicity can be explained by natural selection rather than by a specific architecture such as the language of thought. The explanation of systematicity thus provides no evidence for the need for supplementary evidence that must be provided to support a specific architectural hypothesis and hence undermines the most decisive over time.

Systematicity — **Implementation**

49 Brian McLaughlin, 1993b
Natural selection does not explain systematicity. Natural selection can't explain systematicity, because natural selection doesn't describe its constitutive bases. Evolutionary theory can describe the historical origins of constitutive bases, but it can only do so after the constitutive bases have been described. So, evolutionary theory by itself isn't enough.

Can connectionist networks exhibit systematicity?

33 Tim van Gelder, 1990
Functionally compositional representations avoid the connectionist dilemma. Smolensky and others have developed "functionally compositional" representation schemes, which can account for systematicity without implementing a classical architecture. Such representations can be built up from parts, but they can be broken back down so that "they are non-constitutive—that is, they don't explicitly contain or "token" their parts."
Supported by
Tensor Product Representations Avoid the Dilemma," Box 38.

Note: Van Gelder points out that Gödel numbers also exhibit noncluastic constituent structure. See sidebar, "The Steps of Gödel's Proof," on Map 3.

func · tion · al con · pos · si · tion · al i · ty:
A representational scheme that can produce complex representations from parts and that can decompose a complex representation back down into those parts. The process can be repeated to create increasingly complex representations.

con · cat · en · at · ive com · pos · si · tion · al i · ty:
Functional compositionality, with the added feature that complex representations explicitly contain their parts. Parts are "literally present" or "tokened" in complex representations. Classical symbolic representations are concatenative in this sense.

34 Keith Butler, 1991
Unstructured representations can account for systematicity. Semantic structure can be represented by the causal history of an activation pattern. For example, "John loves Mary" is an unstructured activation pattern that can be traced back to earlier causal effects of "John," "loves," and "Mary." Unstructured representations can be further back to microfeatures of them. In a sense of causal history, unstructured connectionist representations can account for systematicity and related phenomena. So, Fodor and Pylyshyn are right to claim that connectionist representations are unstructured, but wrong to claim that such representations can't account for systematicity.

35 Paul Smolensky, 1988a
Tensor product representations avoid the dilemma. Tensor product representations (see sidebar, "Tensor Product Representations," on this map) have constituent structure but are distributed and context sensitive in a nonclassical way. Vectors representing "John," "loves," and "Mary" are bound together to produce complex representations.

Role: Subject

Filler: John

Tensor Product Representations

In a tensor product representation, vectors representing roles (e.g., "subject" and "verb") are combined with vectors representing role fillers (e.g., "John" and "Mary") by taking their *tensor product*. A tensor product is the vector that results from multiplying each element of one vector by each element of the other. In a connectionist network, this can be implemented by forming the role and filler vectors into separate input layers that connect at a set of multiplicative junctions.

Once a filler has been combined with a role it may be combined with other role or filler products by vector addition to form corresponding elements are simply added together.

	John (subject)	Mary (object)	Loves (verb)	John loves Mary
	1	0	1	1
	1	1	1	1
	1	1	1	1
	1	0	1	1
	1	0	1	1
	1	0	1	1
	1	0	1	1
	1	0	1	1

Microfeatures
upright container
finger-sized handle
burst odor

Cup — **Cup handle** — **Cup bottom** — **Cup** — **Coffee**

minus — **equals**

36 Jerry Fodor and Brian McLaughlin, 1990
The regress of contexts. Smolensky's coffee representation leads to an infinite regress of representations. The representation of a cup of coffee is represented as *cup-with-coffee*. But then *cup-with-coffee* presumably depends on some further representation, such as *cup-of-coffee-on-the-table*. And so on.

Cup-with-coffee — requires — **Coffee**

Cup-of-coffee-on-the-table — requires — **Cup-with-coffee** — requires — **Coffee**

39 Jerry Fodor and Brian McLaughlin, 1990
Constituents of tensor product representations lack causal powers. The parts of a tensor product representation are not explicitly contained or "tokened" within it—the tensor product operation produces the complex representation. As such, tensor product constituents lack individual causal powers. Classical constituents, by contrast, have causal powers by virtue of being explicitly contained in complex representations.

Distributed parts of tensors lack causal powers

Explicitly tokened parts of tensors have causal powers

John | **Loves** | **Sally**

40 Terence Horgan and John Tienson, 1992
Noniclassical constituents can be causally effective. The constituents of a tensor product representation can have causal powers in the sense that naturally occurring (noniclassical) constituents do. For example, the wake left by a motorboat can cause a smaller boat to capsize, and the wake is not a classical constituent of the motorboat. Similarly, noniclassical constituents of tensor product representations may have numerous causal effects, including "setting a buoy to bobbing in a certain way, knocking down a skier, and contributing to the destruction of a sandcastle" (p. 212).

41 David Chalmers, 1990b
Implicitly structured representations engage in structure-sensitive processing. A tensor product representation of distributed representations of complex sentences into the active voice, and vice versa (thereby exhibiting systematicity). This ability shows that the implicit structure of its representations can be causally effective in processing.

John loves Sally

Sally is loved by John

42 Brian McLaughlin, 1993b
Chalmers's representations lack systactic structure. The representations in Chalmers's model lack syntactic structure, even if they represent sentences that possess syntactic structure. The representations lack syntactic structure, such representations cannot participate in structure-sensitive (or syntax-sensitive) processes.

Arguments about levels are irrelevant to the philosophy of connectionism and AI, and go back at least to the work of David Marr.

Unmapped Territory
- Additional levels arguments

Implemented Model

43 Julian Pollack, 1990
The RAAM Network. Passive-active transformations can be achieved, in part, via a recursive auto-associative memory (RAAM) network. A RAAM network compresses constituent vectors into complex representations, and it can also decompose complex representations back down into constituent parts. The processes of compression and decomposition can be repeated to deal with arbitrarily complex connectionist representations.

John | **Loves** | **Mary**

John Loves Mary

Constituent Structure of Mental Representations

Classical symbolic theories postulate a *language of thought* (LOT) underlying thought (Map 3, Box 6), according to which complex mental representations are built out of more simple representations. Complex representations can themselves be combined to form higher-order representations.

When a complex representation explicitly contains its parts, those parts are said to be tokened in the complex representation. The parts of a complex representation are usually referred to as constituents.

Classical theorists claim that mental processes of inference, transformation, composition, and so forth are structure sensitive in that they depend directly on the constituent structure of representations.

complex representation ("molecular representation")

John | **Loves** | **Sally**

constituents of the representation (at the lowest level, "atomic representations")

Other ways of referring to constituent structure include:
- compositional structure
- compositional semantics
- compositionality
- systactic structure
- combinatorial syntax and semantics
- language of thought

Can computers think in Images?

69
Computers can't understand images. Computers can't think because they can't see images in the way that people do. Computers can only deal with formal symbolic information.
Note: "Image" in these arguments usually refers to an "eyes-closed" image, which is imagined in one's mind without the real object mechanically present. Sometimes the term is also used to describe an "eyes-open" perception.

Paul Smolensky

Can images be realistically represented in computer arrays?

70
Images can't represent as computers by filled cells in an array. Images are represented in a machine's memory function as if they were arranged in a rectangular array. Images correspond to filled cells in the array, and the array-based series pixels are interpreted by higher-level systems.
Note: This argument is sometimes referred to as the cathode ray tube, or CRT, theory of mind, because cathode ray tubes are surface matrices.

Douglas Hofstadter

71 Michael Tye, 1991
Images are interpreted symbol-filled arrays. Images are two-and-one-half-dimensional arrays of symbols that encode two-dimensional information about physical surfaces along with viewer-centered information, such as depth and relative orientation of the surfaces. The symbols have an attached semantical interpretation that specifies their representational content.

72 *Anticipated by* Michael Tye, 1991
Connectionist objection to Tye's theory. Tye's theory is a kind of connectionist representation of memory in storage. Memory is not stored in a classical representational manner, but is reactivated in a connectionist dispositional fashion, based on the strength of connections between processing nodes.

73 Michael Tye, 1991
Tye's image theory can accommodate connectionist memory. The theory does not require that images be stored in representationally. The discrete parcels of information contained in images may be stored in a connectionist dispositional system.

74 Michael Tye, 1991
Connectionism cannot accommodate all relevant information. Evidence suggests that image generation is a constructive process that draws on separate packets of information stored in long-term memory. The connectionist dispositional theory lacks an account of how these packets can be stored discretely.

75 David Marr and Keith Nishihara, 1982
Visual perception utilizes two-and-one-half-dimensional arrays of symbols. Visual processing occurs though stages of construction:
1. Visual information is gathered by recognition of intensity differences.
2. A two-dimensional "primal sketch" is made, which orders the visual information into a representation of edges and other surface details.
3. A two-and-one-half-dimensional sketch a two-and-one-half-dimensional sketch is constructed to include depth and movement features the objects in the visual field.
4. Finally, the two-and-one-half-dimensional sketch is compared with other stored two-and-one-half-dimensional sketches.

The result is a three-dimensional structural description with a hierarchical structure.

Are images quasi-pictorial?

77 Geoffrey Hinton, 1979a
The quasi-pictorial view underlies some image effects. Some effects of mental manipulating images suggest that images are represented by interpreting them as pictures. The symbol-filled arrays. However, if we interpret images as structured representations, those effects can be explained. Such structured descriptions represent objects and parts and objects as nodes and relationships as labeled arcs.

78 Geoffrey Hinton, 1979b
Structural descriptions affect imaging tasks. Imaging shows that how a shape is perceived depends on the structural descriptions that are used to construct it.

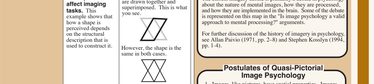
Imagine 2 triangles. Take one, flip it over, and superimpose it on the first. This is what you see:

Imagine 2 parallelograms that are drawn together and superimposed. This is what you see:

However, the shape is the same in both cases.

76 Stephen Kosslyn and James Pomerantz, 1977
Images are quasi-pictorial representations. Images are depictive representations, properties that correspond to those of their correlated physical-objects. Images have structure by virtue of their ties to higher perceptual processes.
Supported by
"Imagery Psychology," Box 88.

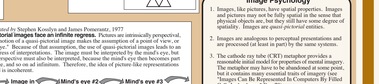
79 *Anticipated by* Stephen Kosslyn and James Pomerantz, 1977
Quasi-pictorial images face an infinite regress. Pictures are intrinsically perspectival, and so the notion of a quasi-pictorial image makes the assumption of a point of view, or "mind's eye," from which the image is interpreted. But this mind's eye faces an infinite regress of interpretations. The image must be interpreted by the mind's eye, but then that perspective must also be interpreted, because the mind's eye then becomes part of the image, and so on and so ad infinitum. Therefore, the idea of picture like representations in the mind is incoherent.



80 Stephen Kosslyn and James Pomerantz, 1977
The mind's eye should be seen as a classification scheme. The mind's eye need not be interpreted as a location in space. The mind's eye should be seen as a classification (or "visual buffer") that interprets sensory information in a series of stages. In the process of interpretation, sensory information is classified in terms of conceptual categories that are correlated with objects and their properties.

81 Stephen Kosslyn, Steven Pinker, George E. Smith, and Steven P. Schwartz, 1979
Machine implementations using images have displayed real computational power. Computer imagery does not lead to an infinite regress about a mind's eye. In fact, computer imagery has been used to solve problems in a variety of tasks, even to the point of being able to solve problems that were not straightforwardly implemented for. Such success could not result from an incoherent theory.

Image in mind's eye #1



Is image psychology a valid approach to mental processing?

88
Image psychology. Members of this school of psychology claim that images play an essential role in thinking.
Note: Image psychologists are not directly concerned with the issue of whether computers can process imagery, and in fact most of the work in this area that images can be represented as filled arrays is an error. However, the imagery debate in psychology is relevant to the representational issue, so it is represented here.

89 Zenon Pylyshyn, 1973
The definition of image is too vague. The notion of "image" is too vague to be of any scientific value, with the commonsense notion of a picture. But the notion of a picture is misleading for the study of mind, because it suggests that "image" processing involves something actually present in the brain when they perceive imagery.

90 Stephen Kosslyn, Steven Pinker, George E. Smith, and Steven P. Schwartz, 1979
Computer simulations show how the image theory can be made precise. Image theory can be precisely described in computational data structures. This is not just a vague picture in the head metaphor, because operations on images (such as rotation and scanning) can be simulated on a computer.

91 Allan Paivio, 1971
Computational theories of imagery give inadequate account of basics. Kosslyn's computational theory fails to explain either the developmental origins of images or their structure and function. The theory simply assumes some units are primitive without explaining them. An operational theory would succeed because it produces factual information rather than mere formal models.

92 Stephen Kosslyn and James Pomerantz, 1977
Pylyshyn uses the wrong notion of an image. Pylyshyn criticizes the "picture-in-the-head" theory of images, according to which they are like metal photographs, but images are rather more like still have some spatial properties.

93 Zenon Pylyshyn, 1981
The picture-in-the-head metaphor severly influences the image theory. Even if no one takes the picture-in-the-head metaphor seriously, the metaphor is implicitly drawn on when image theorists appeal to the spatial properties of images.

94 Allan Paivio, 1971
Images are not functional representations of thought. Thought cannot be explained simply in terms of verbal stimulus-and-response patterns. The phenomena of meaning, reference, and language involve a coordination of images with words that is not fully explained by the behaviorist program.

95 Zenon Pylyshyn, 1973
Images are propositional. To be explanatory, images must play a role in cognition that is the mere experience of images gives us no reason to believe that images play such a central role. Images in themselves (prior to interpretation) are epiphenomenal that allow a causal substrate underlying mental events to stop a wave.
Note: Pylyshyn argues elsewhere that, in general, conscious processes lack explanatory power.

96 Stephen Kosslyn and James Pomerantz, 1977
Images are propositional even if they're not primitive explanatory concepts. It may be true that images are not primitive, low level entities. However, explanation does not always require reference to primitive, low level entities. For example, to understand how a car works, you would learn much about architectural design just by studying bricks, mortar, steel, and so on. Likewise, the proper level of analysis of perception requires the inclusion of images.

97 Zenon Pylyshyn, 1981
Images are cognitively impenetrable. Cognitively impenetrable, in that they can be altered in various ways by what a subject knows. For example, in a biological study of mitosis, students are seen as a composition of petals, sepals, leaves, and stems, whereas an artistic study of the same cells sees them as a composition of colors, shadings, and edges. Cognitive impenetrability cannot be explanatory primitives; they break down into simpler parts depending on context.

cog · ni · tive · ly pen · e · tra · ble: A mental phenomenon is cognitively impenetrable if it cannot be accessed or altered by other thought processes. For example, a perceptual illusion is cognitively impenetrable, because it can't be changed to a perception of gestalt.

According to Zenon Pylyshyn, cognitively impenetrable phenomena are cognitively encapsulated, because they remain basic throughout changes of context.

98 Mark Rollins, 1989
The cognitive impenetrability condition doesn't isolate basic elements. The impenetrability condition presupposes that mental architecture must be cognitively impenetrable is flawed:
- Many functions of the body, such as digestion and the beating of the heart, are not cognitively impenetrable.
- The cognitive impenetrability condition assumes a fixed level for elements in the functional architecture. But functional decompositions can change levels in the network.

99 Stephen Kosslyn, Steven Pinker, George E. Smith, and Steven P. Schwartz, 1979
Cognitive impenetrability doesn't isolate image theory. The cognitive impenetrability condition does not refute the image theory because the image theory does not assume that all image processes are cognitively impenetrable. The image theory acknowledges that some processes are cognitively penetrable whereas others are not. The theory only recognizes that determining which are which is an important research issue.

100 Stephen Kosslyn, 1973
Scanning visual images. Subjects were requested to memorize a series of pictures and later to imagine them one at a time. They were asked to focus on one end of the imagined picture and to identify when a specified feature appeared. Both the amount of time required for the subjects to correctly respond increased with the distance they have taken them to scan an actual picture, which suggests that mental representations have spatial properties analogous to those of external images.

Experiment

103 Stephen Kosslyn, 1994
Mental rotation. Subjects were shown an object and were interpreted with a related version of the same object. Further the object was rotated from its initial position, the longer it took for subjects to see if it was the same or not as the original. The longer the rotation angle, the longer it took for subjects to see if it was the same or not as the original. Therefore, the idea of picture like representations in the mind is incoherent.

104 Kosslyn Cerella & N. Shagoel, J. Metzler; and L. A. Cooper with earlier versions of this experiment.

101 Charles L. Richman, Dale J. Mitchell, and J. Steven Reznick, 1979, *as articulated by* Zenon Pylyshyn, 1981
Evidence from scanning and rotation experiments is problematic. Kosslyn's experimental results may be avoid for the following reasons:
- Subjects' tacit knowledge of real objects makes them think that they are supposed to work with images in the same way that they work with objects. As a result, they are literally scanning or rotating images, but are mentally simulating a real scanning or rotation of a real object. That is, the *task demands* of the experiment confound the results.
- The experiment is able to affect the subjects' with "nonverbal cues, tacit messages, ... loaded answers to questions, and so on" (p. 544). This gives rise to a variety of undesirable experimenter effects.

102 Stephen Kosslyn, Steven Pinker, George E. Smith, and Steven P. Schwartz, 1979
Illegal properties with the imagery experiments have been discoffirmed. Task demands and experimenter effects do not refute scanning and rotation experiments. In fact, the "illegal properties" of these experiments have been discoffirmed. The idea of a real object is discoffirmed by experiments that make no mention of physical-motion and "scanning" or "rotation" terms are replaced by "shift attention."
- The claim that subjects give answers they think the experimenters want to hear is disconfirmed by the fact that subjects often give responses that the experimenter didn't anticipate. These responses could not have been suggested by the experiments.

Postulates of Image Psychology

1. Verbal processing cannot explain all cognitive processing; imagery is not epiphenomenal.
2. Imagery and verbal processing are alternative coding systems, or "modes of symbolic representation."
3. Behaviorisments emphasis on linguistic phenomena because it claims that only verbal reports are empirically accessible. However, experiments can be designed that make mental imagery empirically accessible as well.
4. The more concrete or "*living* like" a stimulus is, the more likely it will be associated with an image, rather than with a verbal process.
5. Imagery is a parallel processing system that stores and manipulates spatial information. Verbal processing, by contrast, operates in series.
6. Chains of symbolic transformation can involve images, words, or both. These include perceptual processes, learning, memory, and language.
7. Basic concepts like "image," "mediation," "word," "processing," and so forth should be defined operationally in order to give these concepts empirical significance.

These postulates are adapted from Allan Paivio (1971). Other proponents (whose opinions may differ on specific points) include Stephen Kosslyn, Steven Pinker, George E. Smith, and Roger Shepard, Arthur Shaun, and others. Stephen Kosslyn's work grew out of image psychology.

op · er · a · tion · al def · i · ni · tion: A definition of a concept in terms of operations, or procedures, that a subject can be operationally defined in terms of the number of times a subject hits a button in a controlled environment. Also, see the "Is the test, behaviorally or operationally construed, a legitimate intelligence test?" arguments on Map 7.

Can computers recognize Gestalts?

82
Gestalt recognition is impossible for computers. Gestalt recognition involves the immediate comprehension of a pattern as a unified whole. Computers can only sequentially process the components of a pattern; they can't recognize Gestalt wholes.
Note: Map 3 contains a variety of Gestalt-inspired objections to symbol systems.

83 Edwin Boring, 1946
Gestalt theories are consistent with computer inference. From a computational point of view, perception involves unconscious inferences similar to those made by electronic computers. From a Gestalt point of view, perception results from the operation of dynamical fields of force in the brain. Both theories provide legitimate scientific explanations of the phenomenon of perception, just from different perspectives. There is no fundamental contradiction between the 2 approaches.

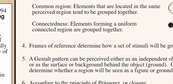
84 Douglas Hofstadter, 1978, *as articulated by* Joseph Rychlak, 1991
Gestalt properties of images can be implemented recursively. Recursive loops can explain the Gestalt figure–ground relationship. Recursive loops allow the background of an image to be constructed as a complement (or opposition) to the figure.

86 Rudolph Arnheim, 1969
Computers can process images from the top down. Humans recognize images from the top down, taking them in as a whole and only later focusing on details. Computers only recognize images from the bottom up, processing local features individually and then aggregating them together.

Postulates of Gestalt Psychology

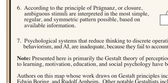
1. The whole is different from the sum of its parts.
2. The whole is studied in terms of its form or organization. A Gestalt—which in German means "pattern" or "shape"—is best understood as kind of perceptual configuration.
3. *Laws of grouping* describe the general patterns of organization obeyed by perceptual wholes.

No grouping.



Proximity: Elements that are closer to one another are grouped together.

Similarity: Elements that are more similar to one another are grouped together.



Closure: Elements that form closed units are grouped together.

Good continuation: Elements forming continuous lines or curves are grouped.

Common region: Elements that are located in the same perceived region are grouped together.

Connectedness: Elements forming a uniform connected region are grouped together.

4. Frames of reference determine how a set of stimuli will be grouped into a perceptual whole.
5. A Gestalt pattern can be perceived either as an independent object (figure), or as a part of a larger pattern (ground). Whether it is perceived as one or the other depends on how it is seen as a figure or ground.
6. According to the principle of Prägnanz, or closure, ambiguous stimuli are interpreted in the most simple, regular,