

Genealogy of the “Grandmother Cell”

CHARLES G. GROSS
Department of Psychology
Princeton University
Princeton, New Jersey

A “grandmother cell” is a hypothetical neuron that responds only to a highly complex, specific, and meaningful stimulus, such as the image of one’s grandmother. The term originated in a parable Jerry Lettvin told in 1967. A similar concept had been systematically developed a few years earlier by Jerzy Konorski who called such cells “gnostic” units. This essay discusses the origin, influence, and current status of these terms and of the alternative view that complex stimuli are represented by the pattern of firing across ensembles of neurons. *NEUROSCIENTIST* 8(5):512–518, 2002. DOI: 10.1177/107385802237175

KEY WORDS Grandmother cell, gnostic cell, Konorski, visual coding, labeled line

The term “grandmother cell” refers to a neuron that would respond only to a specific, complex, and meaningful stimulus, that is, to a single percept or even a single concept. As originally conceived, a grandmother cell was multimodal, but the term came to be used mostly for representing a visual percept. As we shall see, the term arose because the first such neuron was postulated to represent a grandmother. There might be many grandmother cells responding to a specific stimulus, such as one grandmother, but their response properties would be the same. Because of this redundancy, the loss of a grandmother cell or two might not result in loss of the percept. “Coding by grandmother cells” is at the other extreme from “ensemble,” “coarse,” or “population” coding in which a grandmother or other stimulus is coded by the pattern of activity

over a group of neurons. In ensemble coding, there are no “grandmother cells” that detect the unique collection of features that characterize a grandmother. Rather, each member of the ensemble responds somewhat differently, for example, to a granny’s wrinkles, white hair, or to several different old women; the coding of a specific grandmother is done by a unique pattern of activation across the ensemble.

Starting in the early 1970s, the term *grandmother cell* moved from laboratory jargon and jokes into neuroscience journals and serious discussions of the bases of pattern perception (e.g., Barlow 1972; Blake more 1973; Anstis 1975; Frisby 1980; Marr 1982; Churchland 1986). The term is now nearly ubiquitous in introductory neuroscience and vision textbooks, where it often plays the role of straw man or foil for a discussion of ensemble or coarse coding theories of sensory representation (e.g., Cowey 1994; Gazzaniga and others 1998; Rozenzweig and others 1999). This essay considers the origins of the term *grandmother cell* and similar expressions and, more briefly, the roots of ideas about ensemble coding.

Jerry Lettvin and the Birth of Mother and Grandmother Cells

Jerry Lettvin originated the term “grandmother cell” around 1969 (Barlow 1995) in his M.I.T. course titled “Biological Foundations for Perception and Knowledge.” When discussing the problem of how neurons can represent individual objects, he told a (tall) tale of how the neurosurgeon A. Akakhievitch had located a group of brain cells that “responded uniquely only to a mother . . . whether animate or stuffed, seen from before or behind, upside down or on a diagonal or offered by caricature, photograph or abstraction.” At this point, Lettvin introduced the mother-obsessed character from Philip Roth’s (1969) novel *Portnoy’s Complaint* and Akakhievitch ablated all of the mother cells in Portnoy’s brain. As a result, Portnoy completely lost the concept of his mother (see Box 1). Akakhievitch then went on to the study of grandmother cells.

From this origin, the term grandmother cell seems to have spread so quickly that Horace Barlow in his 1972 article “Single Units and Sensation: A Neuron Doctrine for Perceptual Psychology” didn’t even explicitly define the term in criticizing the idea and, in 1973, Colin Blakemore could write of the “great debate [that] has become known as the question of the ‘grandmother cell.’ Do you really have a certain nerve cell for recognizing the concatenation of features representing your grandmother?”

Jerzy Konorski’s Gnostic Units

Although unknown to Lettvin, the grandmother cell idea had actually been set out in detail as a serious scientific proposal a few years earlier by the Polish neurophysiologist and neuropsychologist Jerzy Konorski in his *Integrative Activity of the Brain* (1967), a wide-ranging set of speculations on the neurophysiology of perception and learning (see Fig. 1 and Box 2). His ideas on the organi-

I thank G Berman, D Cha, DF Cooke, MFA Graziano, T Moore, and F Tong for their detailed comments. Supported by N.I.H. grant EY11347.

Address correspondence to: Charles G. Gross, Department of Psychology, Princeton University, Princeton, NJ 08544 (e-mail: cggross@princeton.edu).

Box 1. Lettvin's Story about Mother and Grandmother Cells (ca. 1969)

In the distant Ural mountains lives my second cousin, Akakhi Akakhievitch, a great if unknown neurosurgeon. Convinced that ideas are contained in specific cells, he had decided to find those concerned with a most primitive and ubiquitous substance—mother. . . . And he located some 18,000 neurons that responded uniquely only to a mother, however displayed, whether animate or stuffed, seen from before or behind, upside down or on a diagonal, or offered by caricature, photograph, or abstraction.

He had put the mass of data together and was preparing his paper, anticipating a Nobel prize, when into his office staggered Portnoy, world-renowned for his Complaint. On hearing Portnoy's story, he rubbed his hands with delight and led Portnoy to the operating table, assuring the mother-ridden schlep that shortly he would be rid of his problem.

With great precision he ablated every one of the several thousand separate neurons and waited for Portnoy to recover. We must now conceive the interview in the recovery room.

"Portnoy?"

"Yeah."

"You remember your mother?"

"Huh?"

(Akakhi Akakhievitch can scarcely restrain himself. Dare he take Portnoy with him to Stockholm?)

"You remember your father?"

"Oh, sure."

"Who was your father married to?"

(Portnoy looks blank)

"You remember a red dress that walked around the house with slippers under it?"

"O Certainly."

"So who wore it?"

(Blank)

"You remember the blintzes you loved to eat every Thursday night?"

"They were wonderful."

"So who cooked them?"

(Blank)

"You remember being screamed at for dallying with shiksas?"

"God, that was awful."

"So who did the screaming?"

(Blank)

And so it went. . . . It made no difference—Portnoy had no mother. "Mother" he could conceive—it was generic. "My mother" he could not—it was specific. . . .

Akakhievitch then . . . went back to . . . "grandmother cells."

This parable is abridged from a letter Lettvin sent Horace Barlow in 1995 (Barlow 1995). Much earlier, Barlow (1953) had described cells in the frog's retina as "bug detectors," but little notice had been taken. In the late 1950s, Lettvin and his colleagues at M.I.T. (Lettvin and others 1959) were studying these and other complex cells in the frog, but again the mainstream neuroscience community had ignored their work. Thus, at a 1959 meeting on sensory communication at M.I.T., Barlow (respectable for other reasons by then) but not Lettvin had been invited; Barlow arranged for some of the participants to see experiments in Lettvin's lab. Subsequently, a paper by Lettvin and others was added to the end of the meeting proceedings (Rosenblith 1961), and eventually Lettvin and others "on what the frog's retina tells the frog's brain" became well known. Presumably, Lettvin's research on how the frog's retina codes complex stimuli (Lettvin and others 1959) was related to his story about mother and grandmother cells.

zation of the cerebral cortex in perception anticipated subsequent discoveries to an amazing degree. Konorski predicted the existence of single neurons sensitive to complex stimuli such as faces, hands, emotional expressions, animate objects, locations, and so on (see Fig. 2). He called them "gnostic" neurons, and

they were virtually identical to what later were called grandmother cells. He suggested that the gnostic neurons were organized into specific areas of the cerebral cortex he termed "gnostic fields." That is, he predicted (correctly in many cases) the existence of areas of the cortex devoted to the representations of

such things as faces, emotional expressions, places, and spatial relations. Destruction of a gnostic field would lead, he predicted, to what were later described as category-specific agnosias. Furthermore, he localized many of these gnostic fields, such as the face field in the ventral temporal cortex and the space

field in the posterior parietal cortex (Fig. 3). Overall, these gnostic fields and their locations are remarkably similar to contemporary views of the putative functions of extra-striate visual cortex based on monkey single neuron studies and human imaging experiments (e.g., Caramazza 2000; Martin and others 2000).

At the time of their publication, there was nothing in the literature like Konorski's ideas of highly specialized perceptual neurons in mammals or of areas of the cortex devoted to the representations of particular classes of visual stimuli. In retrospect, however, it is possible to delineate the origins of Konorski's speculations.

Konorski's views of the neural organization of perception were a synthesis and extension of three lines of work in the decade before the publication of his book. The first was Hubel and Wiesel's (1962, 1965) demonstration of the hierarchical processing of sensory information in the geniculo-striate system. In their schema, as one proceeds from center-surround to simple receptive fields and then to complex and then the (now revised) hypercomplex ones, both the selectivity of the cells and their ability to generalize across the retina increase. The possibility that this hierarchy of increasing stimulus specificity continues beyond V2 and V3 was made explicit in their 1965 article, which is repeatedly cited by Konorski. That article ends as follows:

How far such analysis can be carried is anyone's guess, but it is clear that the transformations occurring in these three cortical areas [V1, V2 and V3] go only a short way toward accounting for the perception of shapes encountered in everyday life. (p. 286)

A second line of inspiration for Konorski's ideas was the research by Karl Pribram and his students, particularly Mort Mishkin, on the cognitive effects of lesions on what was then called "association cortex" in monkeys. From his close association with Hal Rosvold and Mishkin (both commented on earlier drafts of the



Fig. 1. Jerzy Konorski in front of the Nencki Institute of Experimental Biology in Warsaw. (Photograph by the author in 1961.)

book), Konorski was well aware that lesions of the inferior temporal cortex produced specific impairments in visual cognition in monkeys (Mishkin 1966) and that similar areas of association cortex existed for audition and somesthesia. Today, at the annual meeting of the Society for Neuroscience, there are multiple sessions on inferior temporal cortex under the general rubric of "Vision." However, at that time most visual neurophysiologists had never heard of this area and did not realize that it had visual functions, let alone that it sat at the top of a series of hierarchically arranged extra-striate visual areas. Indeed, although V2 and V3 had been described, no other extra-striate visual areas such as MT or V4 were known until 1971 (Allman and Kaas 1971). Citing Pribram and Mishkin (1955), Konorski (1967) wrote:

In monkeys the gnostic visual area seems to be localized in inferotemporal cortex, as judged from

numerous experimental results in which ablations of this region produced impairment of visual discrimination. (p. 123)

The third line of evidence for his theories of gnostic neurons and areas came from Konorski's familiarity with the various agnosias that follow cortical lesions in humans from his own clinical experience, from the Western neuropsychological literature, and from Luria's work in the Soviet Union. He was aware of both the symptomatic specificity of some cases of agnosia and their tendency to be localizable. Furthermore, unlike most contemporary neuropsychologists and neurophysiologists, he was aware of the similarity of human agnosias to the effects of experimental lesions in monkeys. For example, he directly related prosopagnosia or face agnosia after ventral temporal lesions in humans to the visual learning deficits in monkeys after inferior temporal lesions.

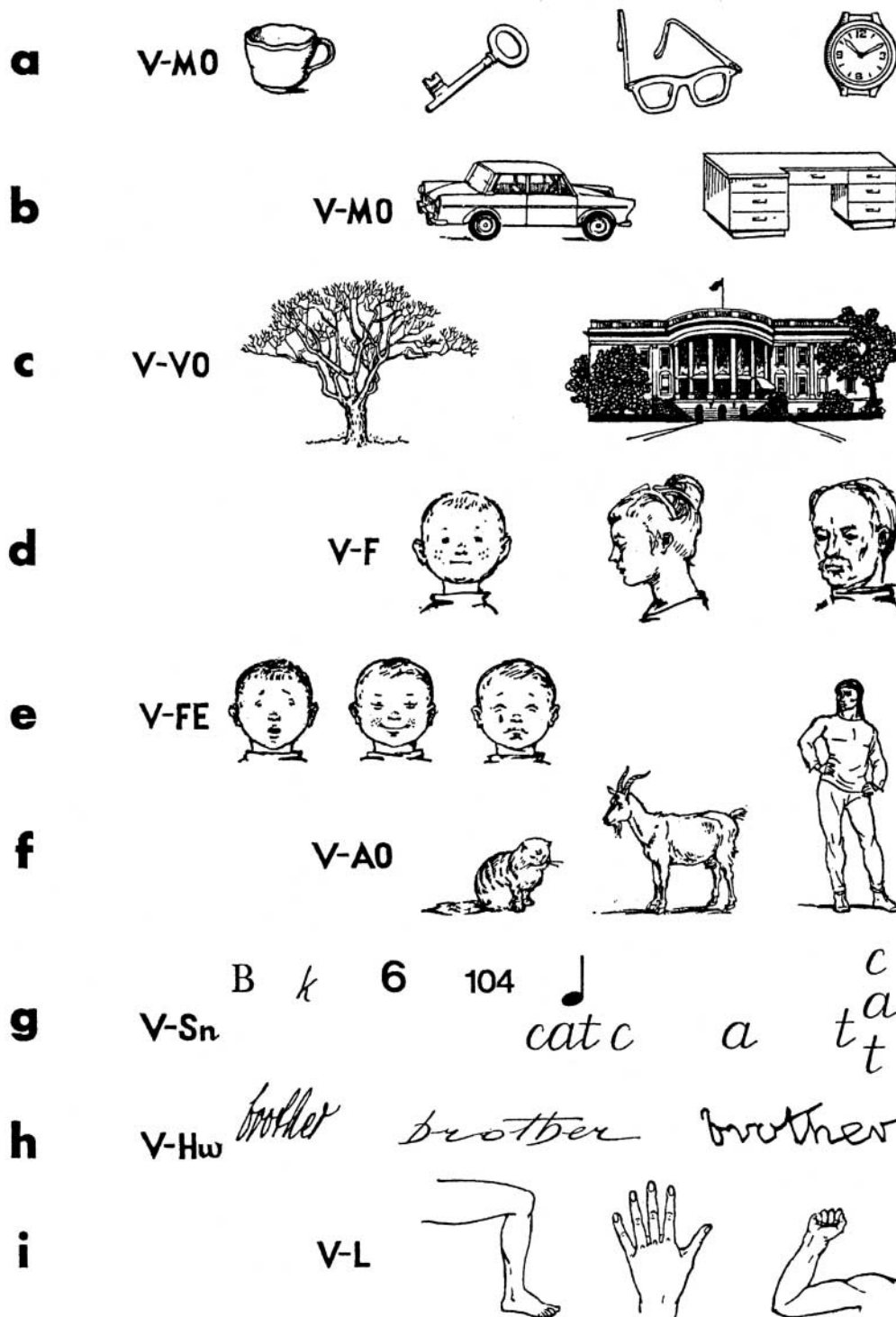


Fig. 2. "Particular categories of visual stimulus-objects probably represented in different gnostic fields" (Konorski 1967).

In summary, Konorski's prophetic ideas on gnostic neurons and gnostic fields came from a bold extension of Hubel and Weisel's findings to account for specific cognitive effects of specific lesions in monkeys and humans.

Konorski's book received a long and laudatory review in *Science*

(Gross 1968). However, for at least the next decade virtually all the many citations to the book were to the parts concerned with learning rather than perception; learning theory still dominated American psychology. As described in the next section, the ideas on gnostic neurons did influ-

ence one laboratory, namely, the laboratory that first reported (the predicted) neurons in the IT cortex that selectively respond to faces and hands.

In the last decade, gnostic cells have begun to be commonly mentioned in textbooks and in the vision

Box 2. Jerzy Konorski (1903–1973)

Konorski's (1967) speculations about gnostic cells came at the end of a long and distinguished career studying the brain and behavior (Fonberg 1974; Konorski 1974). As medical students in Warsaw, he and Stefan Miller discovered that there was another type of conditioned reflex other than the one discovered by Pavlov, namely, one under the control of reward. They called it Type II to distinguish it from Pavlov's which they called Type I. Subsequently and independently, Skinner made this same distinction and Konorski and Miller's Type II conditioning became known as operant or instrumental conditioning.

After a few years as a psychiatrist, Konorski spent 2 years in Pavlov's laboratory in Leningrad but never convinced the master that there really were two types of conditioned reflexes. Konorski then returned to Warsaw and set up a conditioning laboratory in the Nencki Institute of Experimental Biology. He also married and collaborated with Dr. Liliana Lubinska who had studied neurophysiology in Paris and through her became familiar with Western and particularly Sherringtonian neurophysiology. When the war started, Konorski was extraordinarily fortunate to be able to escape Poland. (His colleague Miller committed suicide when the Nazis arrived.) His Russian friends got him appointed head of the famous primate laboratory at Sukhumi on the Black Sea. The laboratory eventually moved to Tbilisi as the Germans approached. It was still near the front, and Konorski had a great deal of experience treating head wounds in a nearby army hospital. At the end of the war, he returned to Poland and played a major role in reconstructing Polish neuroscience as head of the Department of Neuro-physiology at the Nencki Institute.

In 1948, Cambridge University Press published his *Conditioned Reflexes and Neuron Organization*, which was an attempt to bring Pavlovian reflexology in line with Sherringtonian neurophysiology. In the 1960s, there were close collaborative relations between Konorski's laboratory and the Laboratory of Neuropsychology at N.I.H. In addition to the concepts of gnostic units and gnostic fields, Konorski's *Integrative Activity of the Brain* (1967) contains many important and influential ideas about learning and memory.

strong selective pressure for a monkey to distinguish individual bananas, it would probably have ensembles for doing so that were made up of cells selective for bananas in general, but that showed graded response to different characteristics of bananas.

Labeled Lines and Hierarchies

Two central characteristics of grandmother/gnostic cells have a long history. The first is that they are examples of labeled line coding, and the second is that they are at the top of a hierarchy of increasing convergence.

Labeled line coding refers to activity in a neuron coding a particular stimulus property, such as line orientation or a grandmother. This specificity derives from the connec-

tions of the neuron, not from the pattern of the neuron's firing as is the case for various temporal codes such as rate, latency, or phase locking. Perhaps the earliest notion of a labeled line was in Galen's distinction between sensory and motor nerves in the second century (Gross 1998a, 1998b). On the basis of his experience as physician to the gladiatorial school in Pergamon, he realized that section of some nerves results in a specific sensory loss and section of others results in motor loss. He thought the distinction derived from the connections of the nerves to specific regions of the brain.

The first modern labeled line theory of vision was Thomas Young's trichromatic theory of color (Boring 1950). Johannes Muller (1838/1965) then generalized this idea to all the

senses in his Doctrine of Specific Nerve Energies. In that doctrine, when a given nerve type (or nerve energy, in his terms) was excited, the same type of experience is produced independent of the cause of the excitation. The first example of labeled line coding by single neuron activity was probably Adrian and Matthews' (1927) finding that action potentials in a given optic nerve fiber of the conger eel signaled the photic stimulation of a specific part of the eel's retina.

Turning to the other property of grandmother cells, convergence, the most extreme example of neural convergence is William James's concept of a "pontifical cell" whose activity is identical to consciousness as in this passage from his *Principles of Psychology* (1890).

There is, however among the cells one central or pontifical one to which our consciousness is attached. But the events of all the other cells physically influence this arch-cell; and through producing their joint effects may be said to 'combine.' (p. 179)

C. S. Sherrington in his classic *Man on His Nature* (1940) took up James's ideas that there might be "convergence . . . of the nervous system . . . onto one ultimate 'pontifical nerve-cell.'" He then rejected pontifical cells in favor of an ensemble cell theory of consciousness as "a million-fold democracy whose each unit is a cell."

Barlow (1972) thought that the proposal of grandmother cells was inadequate because the multidimensional aspects of visual percepts could not be represented by a single individual cell. Rather, he proposed that a small number of cells would be needed to represent a percept. He named such cells "cardinal" cells because cardinals are lower in the hierarchy than popes and there are more of them.

Concluding Comment

The idea that there might be convergence of neural input onto a single cell, which would provide that cell with the ability to represent a com-

plex and specific percept, seems to have arisen independently several times, first as the gnostic cells elaborated in detail by Konorski and then as the grandmother cells deriving from Lettvin's parable. Cells with properties that are similar to those of gnostic and grandmother cells have been found in both the inferior temporal cortex and the hippocampus. Grandmothers (and other complex objects) may be represented by ensembles of "grandmother" cells, which vary in their responses to different aspects of the stimulus. Finally, contemporary human brain imaging studies have yielded specialized regions of the cortex that closely resemble the gnostic fields proposed by Konorski.

References

- Adrian ED, Matthews B. 1927. The action of light on the eye. Part 1. The discharge of impulses in the optic nerve and its relation to the electric changes in the retina. *J Physiol* 97:378–414.
- Allman JM, Kaas JH. 1971. A representation of the visual field in the caudal third of the middle temporal gyrus of the owl monkey (*Aotus trivirgatus*). *Brain Res* 31:85–101.
- Anstis SM. 1975. What does visual perception tell us about visual coding? In: Gazzaniga MS, Blakemore C, editors. *Handbook of psychobiology*. New York: Academic Press. p 269–323.
- Barlow HB. 1953. Summation and inhibition in the frog's retina. *J Physiol* 119:69–88.
- Barlow HB. 1972. Single units and sensation: a neuron doctrine for perceptual psychology. *Perception* 1:371–94.
- Barlow HB. 1995. The neuron in perception. In: Gazzaniga MS, editor. *The cognitive neurosciences*. Cambridge (MA): MIT Press. p 415–34.
- Blakemore C. 1973. The language of vision. *New Scientist* 58:674–7.
- Boring EG. 1950. A history of experimental psychology. 2nd ed. New York: Appleton-Century-Crofts.
- Caramazza A. 2000. The organization of conceptual knowledge in the brain. In: Gazzaniga MS, editor-in-chief. *The new cognitive neurosciences*. Cambridge (MA): MIT Press. p 1037–46.
- Churchland PS. 1986. *Neurophilosophy: toward a unified science of the mind-brain*. Cambridge (MA): MIT Press.
- Cowey A. 1994. Cortical visual areas and the neurobiology of higher visual processes. In: Farah MJ, Ratcliff G, editors. *The neuropsychology of high-level vision: collected tutorial essays*. Hillsdale (NJ): Lawrence Erlbaum. p 3–31.
- Desimone R. 1991. Face-selective cells in the temporal cortex of monkeys. *J Cognit Neurosci* 3:1–8.
- Fonberg E. 1974. Professor Jerzy Konorski. *Acta Neurobiol Exp* 34:655–64.
- Frisby JP. 1980. *Seeing*. Oxford (UK): Oxford University Press.
- Gazzaniga MS, Ivry RB, Mangun GR. 1998. *Cognitive neuroscience: the biology of the mind*. New York: W. W. Norton.
- Gross CG. 1968. Review of J. Konorski, *Integrative activity of the brain* (1967). *Science* 160:652–3.
- Gross CG. 1992. Representation of visual stimuli in inferior temporal cortex. *Phil Trans R Soc Lond* 335:3–10.
- Gross CG. 1998a. Brain, vision, memory: tales in the history of neuroscience. Cambridge (MA): MIT Press.
- Gross CG. 1998b. Galen and the squealing pig. *The Neuroscientist* 4:216–21.
- Gross CG. 2000. Coding for visual categories in the human brain. *Nat Neurosci* 3:855–6.
- Gross CG, Bender DB, Rocha-Miranda CE. 1969. Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science* 166:1303–6.
- Gross CG, Rocha-Miranda CE, Bender DB. 1972. Visual properties of neurons in inferotemporal cortex of the macaque. *J Neurophysiol* 35:96–111.
- Hubel DH, Wiesel TN. 1962. Receptive fields, binocular interaction and functional architecture in cats visual cortex. *J Physiol Lond* 160:106–54.
- Hubel DH, Wiesel TN. 1965. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *J Neurophysiol* 28:229–89.
- James W. 1890. *The principles of psychology*. New York: Dover.
- Konorski J. 1948. *Conditioned reflexes and neuron organization*. Cambridge (UK): Cambridge University Press.
- Konorski J. 1967. *Integrative activity of the brain; an interdisciplinary approach*. Chicago: University of Chicago Press.
- Konorski J. 1974. Jerzy Konorski. In: Lindzey G, editor. *A history of psychology in autobiography*. Englewood Cliffs (NJ): Prentice Hall.
- Kreiman G, Fried I, Koch C. 2001. Single neuron responses in humans during binocular rivalry and flash suppression. *Abstr Soc Neurosci* 27.
- Kreiman G, Koch C, Fried I. 2000. Category-specific visual responses of single neurons in the human medial temporal lobe. *Nat Neurosci* 3:946–53.
- Lettvin JY, Maturana HR, McCulloch WS, Pitts WH. 1959. What the frog's eye tells the frog's brain. *Proc Inst Radio Engin* 47:1940–51.
- Lettvin JY, Maturana HR, McCulloch WS, Pitts WH. 1961. Two remarks on the visual system of the frog. In: Rosenblith WA, editor. *Symposium on principles of sensory communication*. Cambridge (MA): MIT Press. p 757–76.
- Logotheitis NK, Sheinberg DL. 1996. Visual object recognition. *Annu Rev Neurosci* 19:577–621.
- Marr D. 1982. *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman.
- Martin A, Underleider LG, Haxby JV. 2000. Category specificity and the brain: the sensory/motor model of semantic representations of objects. In: Gazzaniga MS, editor-in-chief. *The new cognitive neurosciences*. Cambridge (MA): MIT Press. p 1023–36.
- Mishkin M. 1966. Visual mechanisms beyond the striate cortex. In: Russel RW, editor. *Frontiers in physiological psychology*. New York: Academic Press. p 93–119.
- Muller JM. 1838 [1965]. On the specific energies of nerves. In: Herrnstein R, Boring E, editors. *A sourcebook of the history of psychology*. Cambridge (MA): Harvard University Press.
- Perrett DI, Rolls ET, Caan W. 1982. Visual neurons responsive to faces in the monkey temporal cortex. *Exp Brain Res* 47:329–42.
- Pribram KH, Mishkin M. 1955. Simultaneous and successive visual discrimination by monkeys with inferotemporal lesions. *J Comp Physiol Psychol* 48:198–202.
- Rolls ET. 1984. Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Hum Neurobiol* 3:209–22.
- Rosenblith WA, editor. 1961. *Symposium on principles of sensory communication*. Cambridge (MA): MIT Press.
- Rozenzweig MR, Leiman AL, Breedlove SM. 1999. *Biological psychology: an introduction to behavioral, cognitive, and clinical neuroscience*. 2nd ed. Sunderland (MA): Sinauer Associates.
- Roth P. 1969. *Portnoy's complaint*. New York: Random House.
- Sherrington CS. 1940. *Man on his nature*. Cambridge (UK): Cambridge University Press.
- Tanaka K. 1996. Inferotemporal cortex and object vision. *Annu Rev Neurosci* 19:109–39.
- Yamane S, Kaji S, Kawano K. 1988. What facial features activate face neurons in the inferotemporal cortex of the monkey? *Exp Brain Res* 73:209–14.