heineken lectures 2000

The Heineken Lectures were presented during a special session of the Royal Netherlands Academy of Arts and Sciences on 2 October 2000 in Amsterdam.

Royal Netherlands Academy of Arts and Sciences

KNAW / Heineken Lectures

2000

Amsterdam, 2002

P.O. Box 19121, 1000 GC Amsterdam, the Netherlands
Telephone: 020-5510700
Fax: 020-6204941
E-mail: knaw@bureau.knaw.nl
WWW-address: http://www.knaw.nl

isbn 90-6984-351-x

The paper in this publication meets the requirements of ∞ iso-norm 9706 (1994) for permanence

contents

preface

It is quite surprising that the value of science for our knowledge-driven society is hardly recognized, neither by the public, using applications of science and new technologies daily, nor by our politicians.

Apparently, we as scientists are not able to convince our politicians of the importance to invest in new developments in science and technology now! Is our political system wrong or are we not using the right arguments? The latter may quite well be. As scientists we are very capable of demonstrating our *tour de forces*, but only a very few of us are taking a more structural approach in explaining the importance of science for our society and actually for our daily life.

Science is not attractive nowadays. It is very hard to enrol youngsters into science programs. We are not alone in this respect; many Western countries are facing this situation. In high schools the children are inadequately confronted with the attractiveness of science and the perspectives for a basic science career are bad, in both universities and industry. To improve the situation within universities a tenure track program has started in 2001, financially supported by the universities, the Netherlands Organization for Scientific Research (nwo), the Royal Netherlands Academy of Arts and Sciences and the Dutch Government.

A tremendous challenge for our universities, enforcing them to focus on their core business, that is the training of undergraduates and graduates and the creation of a stimulating environment for scientists. Universities should stay away from commercial activities and the instalment of non-scientific training programs for the only reason of attracting students.

With the Heineken Prizes we are honouring great scientists and a creative artist, who were able to do the things they wanted in relatively great freedom. The lectures in this volume show that those conditions create a situation in which science and culture flourish.

Robert S. Reneman
President of the Royal Netherlands Academy of Arts and Sciences

James E. Rothman

---

snare *proteins – The basis of cellular membrane fusion and its specificity*

The Dr H.P. Heineken Prize for Biochemistry and Biophysics 2000 was awarded to Dr James E. Rothman for clarifying the mechanism of intracellular membrane fusion.

Ladies and Gentlemen, if you were wise and ate a good lunch before this afternoon's lectures, right now your food is being digested. Glucose – blood sugar – is pouring into your bloodstream, and as a result insulin is being secreted by your pancreas into the bloodstream. This insulin is being detected by cells all over your body, signaling them to store the sugar away as carbohydrates.

Remarkably, these and many other physiological control processes utilize a common basic mechanism – membrane fusion – that is the subject of today's lecture.

How does this work? Insulin is stored in 'packets' inside the cells of your pancreas, and it is stored in little membrane envelope packets called 'secretory' or 'transport' vesicles. At the correct moment these vesicles merge with the outer surface of the cell – as their membranes fuse together – to release insulin into the blood. A few moments later, the insulin has circulated around the body, and signals cells throughout the body, via an insulin receptor to take up the sugar. This, too relies on membrane fusion. Most cells have a specialized protein in their outer (plasma) membrane that transports sugar into the cells. However, most of the glucose transporters are not on the outer surface of the cell, but rather they are inside the cell, once again stored in vesicular packets ready for delivery to the cell surface. So, when insulin signals a cell to take up more sugar, the cell responds by increasing the number of sugar transporters on the surface by fusing the transporter-containing vesicles with the surface membrane of the cell now in the walls. So, here are two cases where membrane fusion is critical for physiology – and there is a very long list.

Consider the synapse between nerve cells, which is the point at which information is passed, where memories are stored, and important computation is done. Synaptic communication involves the release of the substances called neurotransmitters – which are stored inside tiny sacs, little vesicles that merge with the surface membrane by membrane fusion at just the right moment, releasing neurotransmitters to traverse the synapse to signal the new neuron.

Membrane fusion, the merger of two lipid bilayers into one membrane, is analogous to the coalescence of two soap bubbles. Vesicles also fuse with intracellular membranes. A variety of different kinds of vesicles travel through the cytoplasm executing a complex pattern of protein traffic by delivering distinct groups of proteins and lipids to the different organelles they target for fusion. Intracellular transport is a universal and obligatory process for all eukaryotes. It is needed for cell division, since the surface and intracellular membrane systems must double in size to yield two daughter cells. It is also essential for homeostasis of the organism, producing both exocytic vesicles and their content. Vesicle

transport originating at the plasma membrane, termed endocytosis, is responsible for internalizing and distributing macromolecules and key nutrients such as vitamins, iron, and cholesterol. Endocytosis also allows the sensitivity of cells to external signals to be dynamically regulated by providing means to control the turnover of signaling receptors.

Remarkably, the underlying principle mediating membrane fusion events is the same for insulin release, neurotransmission, intracellular membrane assembly, and many other critical cellular and physiological processes. Of course, there are very significant differences in the regulation of these diverse processes – when to fuse, where to fuse – but the same core machinery is used, as I will describe to you today, an insight that has come from two decades of our work carried out successively at Stanford University, Princeton University and now at Sloan-Kettering.

The answer that has emerged could not be more intuitive (Figure 1). 'snare-pins' assemble to link a vesicle to its target membrane. This pin, not unlike a ladies' hairpin, has two ends. One is inserted into the vesicle, the other inserted into the target membrane. snare pins are assembled from two component parts. One part, the 'v-snare', is planted into the vesicle membrane. The other part, the 't-snare', is planted into the target membrane with which the vesicle is destined to fuse. The cytoplasmically-facing portions of the v-snare and the t-snare proteins link up and fold into a stable rod-like structure – the snarepin – which is now anchored in both membranes. Pinning the vesicle to the target membrane triggers the process of membrane fusion when the energy released as the snarepin assembles is used to 'shake-up' the nearby membrane lipids.

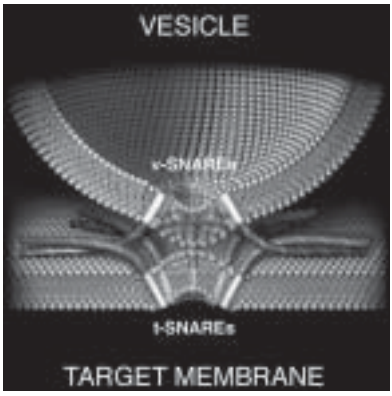


Figure 1. snarepins.

By the early 1960's, George Palade had captured and brilliantly interpreted images of 'zymogen granules' – vesicles storing digestive enzymes in the pancreas – with their membranes merging with the cell surface in the process of discharging their content. In the decade that followed, Palade discovered the secretory pathway and it became clear that membrane fusion must be highly specific to ensure accurate delivery within the cell (Palade, 1975). The fundamental mechanism of membrane fusion and its exquisite specificity remained a central mystery of cell biology whose solution would await reconstitution of transport and fusion and the discovery of responsible proteins, especially the snare complex.

Early experiments – reconstitution and NSF

The line of experiments revealing the fusion machinery stemmed from the cell-free reconstitution of transport involving vesicle budding and fusion in the Golgi (Fries and Rothman, 1980; 1981; Balch *et al*., 1984a,b). Figure 2 shows the assay for transport of vsv G protein within the Golgi. 15B mutant cho cells lack GlcNAc transferase I; thus, oligosaccharides linked to proteins in these cells never acquire GlcNAc, although they are transported normally through the secretory pathway. A Golgi-containing membrane fraction from vsv-infected 15B cells (donor) is incubated with the Golgi-containing fraction (acceptor) from uninfected wild-type cells (plus atp and cytosol). Acquisition of [^{3}H]GlcNAc by G protein measures transfer between donor and acceptor Golgi stacks. This transport is me-

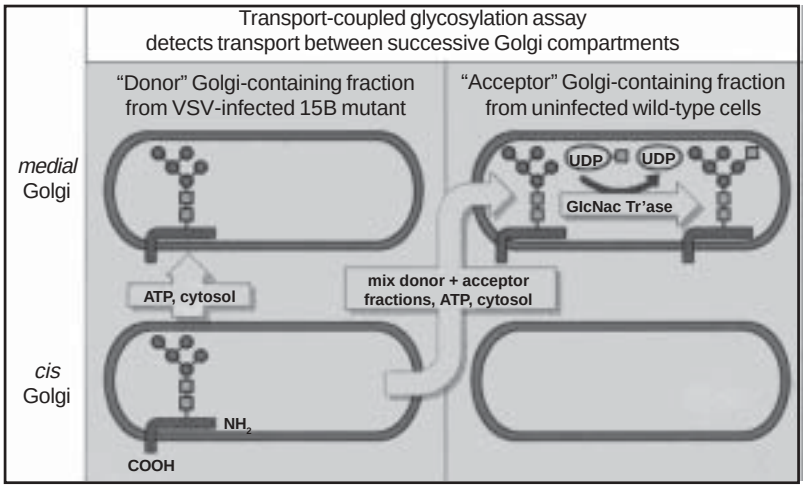


Figure 2. Assay for cell-free transport.

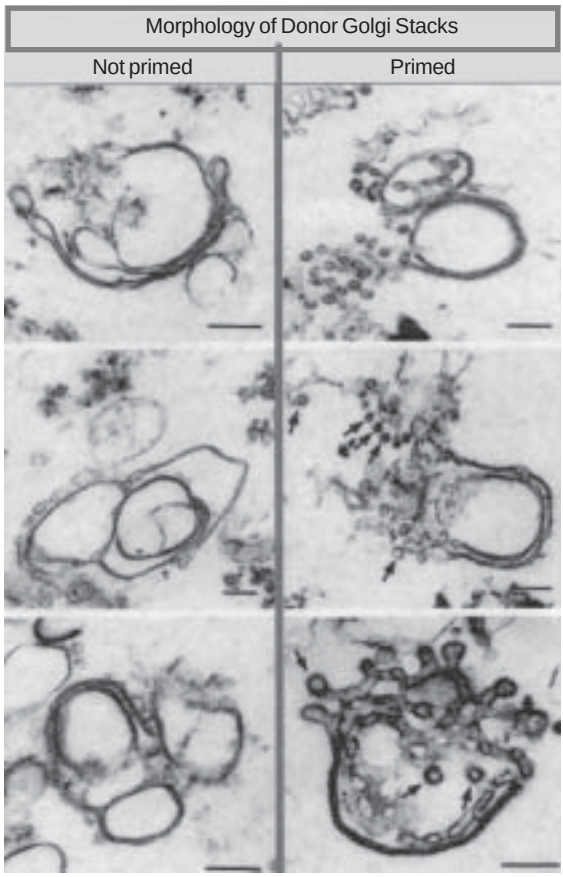


Figure 3. Transport is mediated by vesicles budding from Golgi stacks.

diated by vesicles budding and fusing between Golgi stacks. This was originally suggested by the electron micrographs in Figure 3 of Golgi fractions incubated with cytosol and atp for 15 minutes at 0°C ('not primed') or at 37°C ('primed'), respectively. Transport and vesicle budding are observed in the latter condition.

Transport was inhibited by the sulfhydryl alkylating reagent *N*-ethylmaleimide (nem), allowing the *N*-ethylmaleimide *S*ensitive *F*actor (nsf) to be purified from cytosol fractions based on its ability to restore transport (Block *et al*., 1988). Electron microscopy and other tests revealed that nsf is required for fusion since vesicles accumulate after nem inhibition (Malhotra *et al*., 1988).

We soon appreciated that nsf is an atpase required for vesicle fusion at many compartments in the cell, and that this mechanism is extremely well conserved in

evolution. nsf (Sec18p) from yeast will replace nsf in fusion with the Golgi of higher animals (Wilson *et al*., 1989), foreshadowing the universality of the fusion mechanism that became evident with the discovery of the snare complex.

SNAPS and SNAP receptors

The *S*oluble *N*sf *A*ttachment *P*rotein (snap) was purified according to its ability to bind nsf to Golgi membranes (Clary *et al*., 1990). snap binds to one or more saturable, high affinity '*SNAP RE*ceptors' (which we termed 'snares') and then binds nsf in its atp-bound form. This complex sediments as a 20S particle after extraction from membranes with mild detergents. When nsf hydrolyzes the atp (requiring magnesium ion) it releases itself from the complex and from the membrane (Wilson *et al*., 1992).

It seemed likely that snares were integral membrane proteins because membranes retained the ability to bind snap after alkali extraction (Weidman *et al*., 1989). That put purification of this membrane protein(s) at the very top of our agenda because of the expectation that lipid bilayer fusion would require membrane-anchored proteins, and the snare proteins were the prime candidates.

The meaning of the seemingly futile cycle of atpase-driven membrane binding and release of nsf was then unclear, we imagined that energy from hydrolysis of atp somehow activated the membrane-anchored snares to power fusion. However, the existence of the cycle had a huge impact on our strategy for identifying the snares, because the assembly and disassembly of 20S particles could be

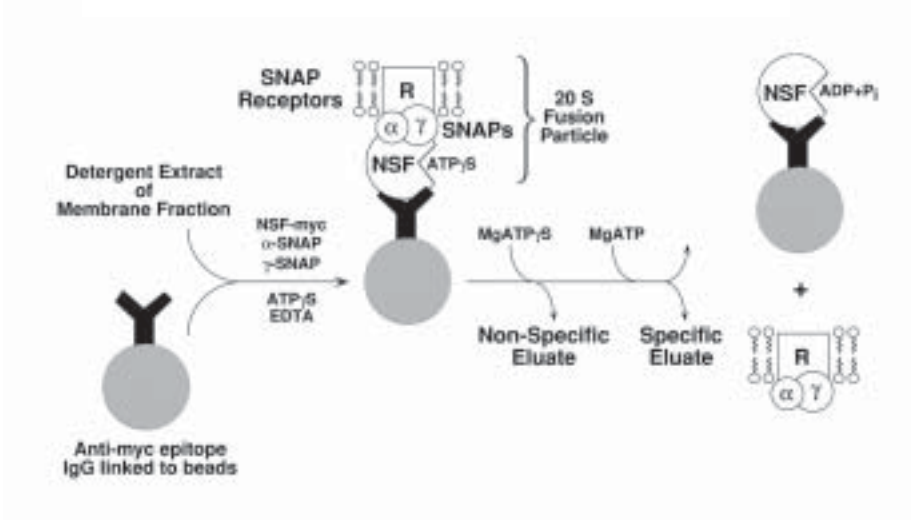


Figure 4. Affinity purification of snap receptors.

exploited as sequential affinity purification steps to isolate snares by binding and release from nsf and snap. Previous experiments had shown that standard chromatographic methods and a single affinity step were inadequate; the 20S cycle would add a second level of biological specificity.

The to-be-identified snares were sequestered from the bulk of membrane protein by incorporation into 20S complexes formed with exogenously added recombinant (bacterially-expressed) nsf and snap proteins. This incubation was done in the presence of atpgs (a non-hydrolyzable analogue of atp) and in the absence of free magnesium ion (Mg^{++} is required for hydrolysis of atp by nsf) to promote 20S particle assembly. The recombinant nsf was expressed with an epitope tag to allow the 20S particles to be isolated with a monoclonal antibody (immobilized on beads) directed against the short peptide epitope (Figure 4).

The second biologically-specific step recapitulated the disassembly of 20S particles. The to-be-identified snares are released when the beads are incubated with magnesium ion and atp to allow nsf to hydrolyze atp. Recombinant nsf remains bound to the beads by the antibody, but the recombinant snap proteins are released along with the to-be-identified snares from the brain membranes (Figure 4).

Because vesicle fusion occurs at many membrane compartments, we had suspected that cells would have a large family of snare proteins, related in sequence and differing in location. We were therefore surprised when the snare proteins derived from whole brain yielded a remarkably simple protein pattern (Figure 5) consisting of only four snare proteins, each present in the specific (MgATP) eluate and absent from the non-specific (MgATPgS) eluate.

Discovery of the SNARE complex

All four snares turned out to be proteins found in synapses (Figure 6). While they had all previously been cloned and sequenced, their function was still unknown. Two were isoforms of syntaxin, a plasma membrane protein identified by Richard Scheller (Bennett *et al*., 1992) based on its ability to bind synaptotagmin, a synaptic vesicle calcium sensor (Geppert *et al*., 1994). The third snare protein was snap-25, short for synaptosome-associated protein of 25 kDa, cloned by Michael Wilson (Oyler *et al*., 1989). snap-25 mainly resides in the plasma membrane and was originally identified because of its abundance in synapses. Its connection to syntaxin and to membrane fusion was a surprise, as was the coincidental relationship of its acronym to that of the soluble nsf attachment protein, snap.

vamp/Synaptobrevin was the last snare protein to emerge. But it was the

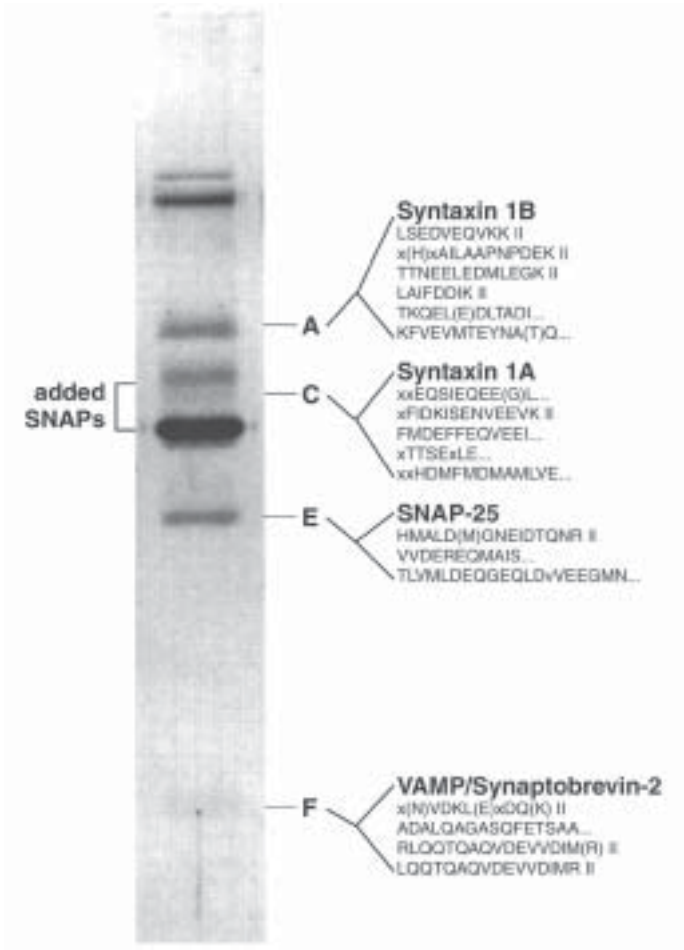


Figure 5. Identification of snare proteins.

lynchpin because in contrast to snap-25 and syntaxin, vamp resides mainly in synaptic vesicles, immediately suggesting how a complex of the snare proteins (perhaps with nsf and snap) could link the vesicle to the membrane and initiate fusion (Figure 6). vamp/Synaptobrevin had been cloned by DeCamilli and Jahn, and independently by Scheller (Baumert *et al*., 1989; Elferink *et al*., 1989).

Supporting our simple model, vamp and syntaxin are membrane-anchored proteins with cytoplasmic domains, and snap-25 is anchored to the cytoplasmic side of the plasma membrane via covalently-attached fatty acids. Close to equimolar amounts of vamp, syntaxin (its two isoforms considered together) and

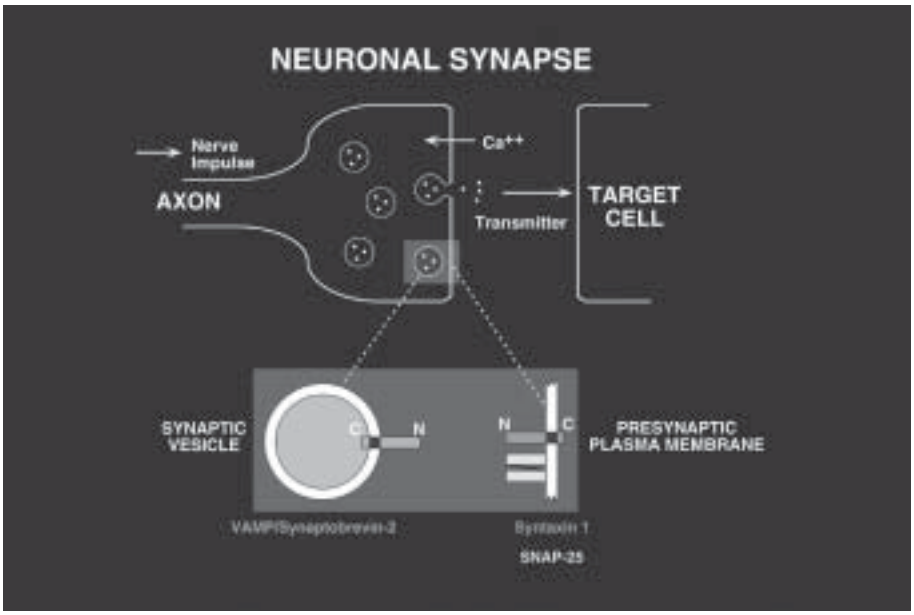


Figure 6. The synapse and snare proteins.

snap-25 were recovered in the isolated complexes. Further, the snare proteins were isolated because they function together with nsf and snap, known to function in fusion.

We interpreted the snare complex (Söllner *et al*., 1993) from the broad perspective of vesicle traffic as distinct from the specialist's point-of-view of synaptic exocytosis. First principles require that vesicles and targets are somehow marked to indicate which vesicles will fuse where. This, in turn, indicates that vesicle and target markers must be matched pairwise. We suggested that the simplest mechanism for matching is self-assembly, in which only matching pairs of 'cognate' vesicle ('v') and target ('t') markers bind each other between membranes, thereby forming a 'v-t' complex prerequisite for membrane fusion.

Based on our cognate vesicle and target marker concept, we proposed the 'snare *hypothesis*' in which the snares are the vesicle and target markers, which we termed v–snares and t-snares. vamp is the v-snare of the synaptic vesicle; syntaxin and snap-25 are the subunits of the cognate t-snare in the plasma membrane. The snare hypothesis provides the framework to generalize our results. We suggested that each vesicle would have its own characteristic v-snare, a homologue of vamp, and that each target membrane would be marked by a characteristic t-snare, homologues of syntaxin and snap-25.

SNARE proteins – minimal machinery for membrane fusion

We soon found that nsf and snap function to disrupt the snare complex using energy derived by atp hydrolysis. This critical step separates the v-snare from the t-snare only when they reside in the same bilayer (i.e., after fusion), but not when they are paired between bilayers (i.e., during fusion) (Weber *et al*., 2000). This allows nsf to recycle snare complexes after fusion while sparing fusion in progress. When it was shown that snap and nsf are not directly involved in bilayer fusion (Mayer *et al*., 1996), attention focused on the simplest remaining possibility, that the snare complex is all that is needed to mediate fusion. The extraordinary thermal stability of the complex (resisting) heat denaturation up to 90°c) was consistent with this source of energy for fusion (Hayashi *et al*., 1994) and its rod-like structure with all membrane anchors at one end, indicated that now the snare complex could bring the two membranes into close contact (Hanson *et al*., 1997).

The proof that the snare complex is the active principle of fusion could only come from demonstrating this function in the absence of any other proteins. Reconstituting recombinant exocytic/neuronal snares into liposomes established that the pairing of cognate snares between lipid bilayers – snarepins – indeed drives membrane fusion (Weber *et al*., 1998). Thus, when complementary v-snare and t-snare pairs engage, a productive fusion event is not only initiated – as we first imagined – but it is also completed (Figure 1).

The SNARE hypothesis – the specificity of membrane fusion

The snare hypothesis triggered extensive research that led to the identification of snares, which as predicted localize and function at compartments engaging in fusion (Bock *et al*., 2001). In the most direct test of the snare hypothesis to date, we established that the specificity for membrane fusion is encoded in the physical chemistry of the isolated snare proteins (McNew *et al*., 2000). Almost without exception, fusion only takes place with the rare combinations of v- and t-snares that correspond to flow patterns occurring in the living cell.

Matching snares establish which fusion events can potentially occur in a cell. Because t-snares are intrinsically auto-inhibited (Parlati *et al*., 1999) they can be locally activated, thereby allowing vesicles to be targeted to a distinct region within a membrane (such as the leading edge of plasma membrane of a moving cell) and also adding a layer of specificity. Vesicles are initially captured by flexible protein tethers and they are 'tethered' before they can be 'snared' (Mellman and Warren, 2000). Current research focuses on how the snare complex is regulated

by additional proteins to permit membrane fusion to be spatially and temporally controlled.

Acknowledgements

The text of this Heineken Lecture is adapted from a discovery of the snare complex published on the website ergito.com. Figures 3 and 4 are from *Nature* 362, 318-324 (1993). Figure 6 is adapted from a cover of *Cell* 92 (1998). Figures 1 and 2 were adapted from *Cell* 39, 525-536 (1984) by ergito.com.

This occasion provides the chance to thank the many co-workers who have contributed to the understanding of vesicle transport and membrane fusion over the past 23 years. Especially, I would like to express my gratitude to two long-term collaborators who made special contributions, Dr. Lelio Orci and Dr. Thomas Söllner, and to the generosity of and superb environments of Stanford University, Princeton University and Sloan-Kettering Institute. Finally, I want to acknowledge the important influence of and inspiration by the two men who most impacted me as a scientist – Eugene P. Kennedy of Harvard (my Ph.D. thesis advisor) and Arthur Korn of Stanford (in whose department this body of work began when I joined as an Assistant Professor in 1978).

References

Balch, W., Dunphy, W., Braell, W., and Rothman, J.E. (1984a). Reconstitution of the transport of protein between successive compartments of the Golgi measured by the coupled incorporation of N-acetylglucosamine. *Cell* 39, 405-16.

Balch, W., Glick, B., and Rothman, J.E. (1984b). Sequential intermediates in the pathway of intercompartmental transport in a cell-free system. *Cell* 39, 525-36.

Baumert, M., Maycox, P.R., Navone, F., De Camilli, P., and Jahn, R. (1989). Synaptobrevin: an integral membrane protein of 18,000 daltons present in small synaptic vesicles of rat brain. *EMBO J.* 8, 379-84.

Bennett, M.K., Calakos, N., and Scheller, R.H. (1992). Syntaxin: a synaptic protein implicated in docking of synaptic vesicles at presynaptic active zones. *Science* 257, 255-9.

Block, M.R., Glick, B.S., Wilcox, C.A., Wieland, F.T., and Rothman, J.E. (1988). Purification of an N-ethylmaleimide-sensitive protein catalyzing vesicular transport. *Proc. Natl. Acad. Sci. USA* 85, 7852-6.

Bock, J.B., Matern, H.T., Peden, A.A., and Scheller, R.H. (2001). A genomic perspective on membrane compartment organization. *Nature* 409, 839-41.

Clary, D.O., Griff, I.C., and Rothman, J.E. (1990). snaps, a family of nsf attachment proteins involved in intracellular membrane fusion in animals and yeast. *Cell* 61, 709-21.

Elferink, L.A., Trimble, W.S., and Scheller, R.H. (1989). Two vesicle-associated membrane protein genes are differentially expressed in the rat central nervous system. *J. Biol. Chem.* 264, 11061-4.

Fries, E., and Rothman, J.E. (1980). Transport of vesicular stomatitis virus glycoprotein in a cell-free extract. *Proc. Natl. Acad. Sci. USA 77*, 3870-4.

Fries, E. and Rothman, J.E. (1981). Transient activity of Golgi-like membranes as donors of vesicular stomatitis viral glycoprotein in vitro. *J. Cell Biol.* 90, 697-704.

Geppert, M., Goda, Y., Hammer, R.E., Li, C., Rosahl, T. W., Stevens, C.F., and Sudhof, T.C. (1994). Synaptotagmin i: a major Ca2+ sensor for transmitter release at a central synapse. *Cell* 79, 717-27.

Hanson, P.I., Roth, R., Morisaki, H., Jahn, R., and Heuser, J.E. (1997). Structure and conformational changes in nsf and its membrane receptor complexes visualized by quick-freeze/deep-etch electron microscopy. *Cell* 90, 523-35.

Hayashi, T., McMahon, H., Yamasaki, S., Binz, T., Hata, Y., Sudhof, T.C., and Niemann, H. (1994). Synaptic vesicle membrane fusion complex: action of clostridial neurotoxins on assembly. *EMBO J.* 13, 5051-61.

Malhotra, V., Orci, L., Glick, B.S., Block, M.R., and Rothman, J.E. (1988). Role of an N-ethylmaleimide-sensitive transport component in promoting fusion of transport vesicles with cisternae of the Golgi stack. *Cell* 54, 221-7.

Mayer, A., Wickner, W., and Haas, A. (1996). Sec18p (nsf)-driven release of Sec17p (a-snap) can precede docking and fusion of yeast vacuoles. *Cell* 85, 83-94.

McNew, J.A., Parlati, F., Fukuda, R., Johnston, R.J., Paz, K., Paumet, F., Söllner, T.H., and Rothman, J.E. (2000). Compartmental specificity of cellular membrane fusion encoded in snare proteins. *Nature* 407, 153-9.

Mellman, I., and Warren, G. (2000). The road taken: past and future foundations of membrane traffic. *Cell* 100, 99-112.

Oyler, G. A., Higgins, G. A., Hart, R. A., Battenberg, E., Billingsley, M., Bloom, F.E., and Wilson, M.C. (1989). The identification of a novel synaptosomal-associated protein, snap-25, differentially expressed by neuronal subpopulations. *J. Cell. Biol.* 109, 3039-52.

Palade, G. (1975). Intracellular aspects of the process of protein synthesis. *Science* 189, 347-58.

Parlati, F., Weber, T., McNew, J.A., Westermann, B., Söllner, T.H., and Rothman, J.E. (1999). Rapid and efficient fusion of phospholipid vesicles by the a-helical core of a snare complex in the absence of an N-terminal regulatory domain. *Proc. Natl. Acad. Sci. USA* 96, 12565-70.

Söllner, T., Whiteheart, S.W., Brunner, M., Erdjument-Bromage, H., Geromanos, S., Tempst, P., and Rothman, J.E. (1993). snap receptors implicated in vesicle targeting and fusion. *Nature* 362, 318-24.

Weber, T., Parlati, F., McNew, J.A., Johnston, R.J., Westermann, B., Söllner, T. H., and Rothman, J.E. (2000). snarepins are functionally resistant to disruption by nsf and asnap. *J. Cell Biol.* 149, 1063-72.

Weber, T., Zemelman, B.V., McNew, J.A., Westermann, B., Gmachl, M., Parlati, F., Söllner, T.H., and Rothman, J.E. (1998). snarepins: minimal machinery for membrane fusion. *Cell* 92, 759-72.

Weidman, P.J., Melancon, P., Block, M.R., and Rothman, J.E. (1989). Binding of an N-ethylmaleimide-sensitive fusion protein to Golgi membranes requires both a soluble protein(s) and an integral membrane receptor. *J. Cell Biol.* 108, 1589-96.

Wilson, D.W., Whiteheart, S.W., Wiedmann, M., Brunner, M., and Rothman, J.E. (1992). A multisubunit particle implicated in membrane fusion. *J. Cell Biol.* 117, 531-8.

Wilson, D.W., Wilcox, C.A., Flynn, G.C., Chen, E., Kuang, W.J., Henzel, W.J., Block, M.R., Ullrich, A., and Rothman, J.E. (1989). A fusion protein required for vesicle-mediated transport in both mammalian cells and yeast. *Nature* 339, 355-9.

Eric R. Kandel

Genes, synapses and memory storage

Professor Eric R. Kandel received the Dr A.H. Heineken Prize for Medicine 2000 for his pioneering research on the molecular mechanisms underlying learning processes and memory.

In the last several decades of this century we have witnessed a remarkable increase in the explanatory power of biology, including the biology of the brain. The ability of biology to address central issues of brain function will only increase in the 21st century and will likely have a broad impact on many aspects of our lives. As a result when the intellectual historians look back on these decades, they are likely to acknowledge that the deepest insights into the nature of mental processes will not come from the disciplines traditionally concerned with mind. They will have come not from philosophy, from the arts, or even from psychology of psychoanalysis, but from biology. This is because in the last two decades biology has participated not simply in one but in two major unifications of thought which bear on our understanding of mind.

First, there has been a remarkable scientific synthesis, achieved through molecular biology, that has brought together the disciplines of cell biology, biochemistry, developmental biology, and oncogenesis. This unification derives from major advances in our understanding of the gene which have revealed how its structure determines heredity and how its regulation determines development and function. These remarkable insights have given us a marvelous sense of the conservation between different cells in any one organism and different organisms across evolution.

Second, there has been a parallel unification between neurobiology, the science of the brain, and cognitive psychology, the science of the mind. This second unification is far less mature than that brought about by molecular biology, but it is potentially equally profound, for it promises to provide us with a new framework for the analysis of a variety of mental functions, such as perception, action, language and memory.

These two independent unifications stand at the extremes of the biological sciences: the one at the interface between biology and chemistry; the other at the interface between biology and psychology. This raises a question: to what degree can these two disparate strands be brought together? Can molecular biology enlighten the study of mental processes as it has enlightened the study of development and oncogenesis? Can we anticipate an even broader synthesis in the 21st century, a synthesis ranging from molecules to mind? In this essay I outline the possibility of a *molecular biology of cognition*, and suggest that it will occupy center stage in the early part of the 21st century. I outline this development by using as an example the study of memory.

We begin with a historical perspective because many of the themes that dominate current research on memory – including the distinction between systems and molecular approaches to memory – emery best in a historical overview. We then

will describe more recent molecular biological investigations of memory. Here we will focus in particular on one component of the memory process: the switch from short- to long-term memory. This component can in fact now be analyzed by combining cognitive psychology with modern molecular biology. Finally, we will suggest some possible future directions in memory research, focusing on the drawing together of strands of research that have historically been separated.

The problem of memory has a systems component and a molecular component

The work of Ramon y Cajal at the beginning of the century (Cajal, 1893) and of Donald Hebb in 1949 (Hebb, 1949) established a useful conceptual framework for the study of memory, based on the idea that memory is stored as changes in the strength of specific synaptic connections. This framework divides the study of memory into two components: the *systems problem* and the *molecular problem*. The *systems problem* of memory is concerned with where in the brain memory is stored and how neural circuits work together to create, process, and recall memories. The *molecular problem* of memory is concerned with the mechanisms whereby synapses change and information is stored. Most early word on memory focused on the systems problem, focusing on the question, 'Where is memory stored?' We therefore begin our historical perspective with this question.

Broca, Wernicke and the localization of language

In 1861, Pierre-Paul Broca, a French neurologist, described the first of nine patients who suffered form a language impairment (now known as *Broca aphasia*) in which they could understand language but could not speak fluently. These defects were specific to the expression of language rather than to motor control of the vocal tract, as the patients could hum or whistle a tune but could not write fluently. Post-mortem examination of the patient's brains revealed in each case a lesion in the posterior region of the frontal lobe (a region of cortex now called *Broca's area*. Thus, for the fist time, Broca was able to assign a well-defined higher function to a specific region of cortex (Broca, 1865) Since all of these lesions were in the left hemisphere, Broca was also able to establish that the two hemispheres, although apparently symmetrical, have slightly different functions.

A decade later, in 1876, Carl Wernicke, a German neurologist, described a second type of aphasia (*Wernicke's aphasia*) that is in a way the opposite of Broca's: an impairment not of the production of speech but of comprehension. Wernicke found that this syndrome was caused by a lesion in the posterior superior portion of the temporal lobe of the left hemisphere, a lesion distinct from that described

by Broca in an area we now call *Wernicke's area* (Wernicke, 1908).

Taking his findings together with those of Broca, Wernicke put forward a theory of how cortex is organized for language that, although simpler than our current understanding, is still central to how we now view the brain. Wernicke proposed that any complex behavior requires the activity not of one but of a number of different brain areas, and that these areas are interconnected in various ways. Mental activity is not unitary or seamless as might intuitively appear to be the case, bus can be broken down into multiple components, and each component can be assigned to a more or less specific brain region, much as the organologists had insisted. However, these different specialized areas do not function by themselves bus as part of large, interconnected networks. By application of this model Wernicke predicted, correctly, the possibility of a third kind of aphasia, *conduction aphasia*, resulting from lesion not of Broca's area or Wernicke's area but of the fibers (the arcuate fasciculus) passing between the two. Thus, while specific functions are localized as Gall had insisted, the function of the brain as a whole requires distributed processing somewhat reminiscent of that propounded by Flourens. Wernicke's model of a distributed network of specialized areas has emerged as a dominant theme in the study of the brain.

Can memory storage be localized to specific regions of the brain?

The finding that the language can be localized within the brain led to the hunt for other areas concerned with specific higher functions. Areas concerned with motor control and with each of the senses were soon identified. It was only a matter of time before efforts to localize cognitive function would turn to memory (Ferrier, 1890; Jackson, 1884). However, attempts in the first half of the twentieth century to localize memory failed. The dominant figure in the period was Karl Lashley, Professor of Psychology at Harvard. Lashley began the experimental search for the locus of memory storage by training rats on specific memory tasks, systematically removing portions of cortex, and then testing them for recall. He repeatedly failed to find any particular brain region that was special to or necessary for the storage of memory. On the basis of these findings, Lashley formulated the law of mass action, according to which the extent of a memory deficit is correlated with the size of a cortical lesion but not with the specific site of that lesion (Lashley, 1929). This law was reminiscent of the views of Flourens a century earlier.

The first clear evidence for the localization of memory came not from experimental animals but from clinical studies. In 1938, the neurosurgeon Wilder Penfield, working at the Montreal Neurological Institute, developed methods for

the surgical treatment of focal epilepsy, a form of epilepsy in which seizure is restricted to a relatively small region of the cortex. To functionally map the areas surroundings the epileptic center so as to avoid later damage to critical areas, such as Broca's and Wernicke's, Penfield electrically stimulated the cortical surface. Because the brain contains no pain receptors, the patients could remain unanaesthetized and could report what they experienced during stimulation. In this way Penfield studied over one thousand patients and mapped out in each of them most of the exposed cortical surface. On rare occasions during such mappings, Penfield found a region of temporal cortex where stimulation gave rise to specific *experiential responses*, memory-like perceptions that the patients could describe. Penfield concluded that portions of the temporal lobe were specifically involved in memory (Penfield and Perot, 1963).

More conclusive evidence for the involvement of temporal lobe structures in memory came in 1957, when William Scoville, a neurosurgeon influenced by Penfield, and Brenda Milner, a psychologist and long-term collaborator of Penfield's, reported the now famous case of H.M. (Scoville and Milner, 1957). At age 9, H.M. sustained a head injury after being hit by a bicycle; over the next 18 years he suffered progressive seizures until he was completely incapacitated. As a last resort, H.M. underwent complete bilateral removal of the medial temporal lobe (where his seizures initiated). The surgery relieved his epilepsy, but he was left with a profound memory deficit: from the time of his surgery until this day he has been unable to form any new memories of people, facts, or events.

Brenda Milner studied H.M. and demonstrated that structures in the medial temporal lobe that Scoville had removed are specialized for memory. Her further studies with H.M. not only controverted Lashley's influential views but also cast new light on the systems problem of memory: there are, within the brain, multiple, functionally specialized memory systems.

Memory is not unitary faculty of mind: there are multiple memory systems

The idea that there may be multiple memory systems is old, but it did not enter mainstream psychological thinking until Milner's work in the 1960s. In the early parts of the nineteen century, the French philosopher Maine de Biran argued that memory can be subdivided into different systems for ideas, feelings and habits (Copplestone, 1977). In the 20th century, William James emphasized the idea that memory has distinct temporal phases. Henri Bergson developed the distinction between conscious memory and habit (Bergson, 1913) In 1949 the British philosopher Gilbert Ryle proposed a similar distinction between 'knowing that' (con-

scious recall of knowledge for facts and events) and 'knowing how' (knowledge of performance or skills without recourse to conscious awareness) (Ryle, 1949). A similar distinction was made by the psychologist Jerome Bruner, who termed 'knowing that' a *memory with record* and 'knowing how' a *memory without record* (Milner *et al*., 1998). The defining characteristic of memories with record is the ability to summon up a more or less detailed conscious recollection of facts and events about persons, places, and objects. The defining characteristic of a memory without record, by contrast, is a change in the way an organism responds to a situation or a stimulus, without access to the specific circumstance under which the memory was formed. The idea of distinct memory systems is in a sense implicit in Freud's psychoanalytic writing. Central to Freud's view of the brain is the distinction between conscious and unconscious memories.

Thus, even prior tot 1960 a fractionation of memory had already been proposed on the basis of content, function, and temporal profile. Nevertheless, the concept of multiple memory systems only drew the attention of the scientific community with the studies of H.M. After the profound nature of his memory deficit was recognized, Milner made the surprising further discovery that that despite his impairment he could learn a surprising amount of new information. First, H.M. was found to have perfectly good short-term memory: he could accurately repeat back a telephone number, and he can carry on a normal conversation. It is only when he is distracted from the topic or tasks at hand that his memory deficit reveals itself. Thus, the temporal lobe structure, which H.M. lacks, is not required for short-term (or '*working*') memory. This finding validated the early distinction between short and long-term memory.

Second, Milner found that H.M. has reasonably good long-term memory for events prior to his operation. He maintains his overall intelligence and has good command of English. He remembers events from his childhood and adult life before his surgery. There is a period of retrograde amnesia for events shortly before the surgery, but for the most part H.M.s symptoms revealed that the medial temporal lobe is not the ultimate storage site for previously acquired knowledge. This finding supports the idea that knowledge is ultimately stored in whatever area of the cortex processes the relevant sort of information (see for example Zeki, 1993).

Third, and most surprising, H.M. is able to form certain types of long-term memory. In 1962 Milner and the psychologist Suzanne Corkin found that H.M. was able to acquire new motor skills (specifically, the ability to trace a complex figure in a mirror) (Corkin, 1965). When asked, he would deny that he had encountered or practiced the task before; but his performance showed unequivocal

improvement over time. This finding showed that learning of this skill is preserved after severe temporal lobe damage and in the presence of profound amnesia for facts and events, and thus it demonstrated for the first time a fractionation of memory on the basis of content rather than just duration. Milner and Corkin thus validated the distinction between conscious memory and habit propounded by Bergson 50 years earlier and by Ryle in 1949.

The learning tasks that amnesia patients like H.M. are capable of mastering have several things in common. They have an automatic quality, and the formation or expression of the memories is not dependent on awareness or cognitive processes such as comparison and evaluation. This type of memory typically builds up slowly over many trails and is expressed primarily by improved performance on certain tasks. The psychologist Lawrence Weiskrantz has noted that the spared learning skills are *reflexive* rather than *reflective* – typically the patient need only produce a physical response tot a stimulus or cue.

This distinction was soon validated on normal subjects. Larry Squire has framed the distinction particularly well by emphasizing the ability of humans to report verbally the contents of explicit memory but not of implicit memory: explicit memory is thus *declarative* whereas implicit memory is *non-declarative* (Squire and Zola-Morgan, 1991). Daniel Schachter framed Bruner's distinction using the terms *implicit* for 'knowing how' and *explicit* for 'knowing that' (Schachter, 1996). These are the most widely used terms, and while specific definitions of these various terms can differ in different contexts, all these authors are describing the basis distinction that was revealed in the studies of H.M.

The molecular component of memory storage

An examination of memory storage at the systems level revealed that memory is not a unitary faculty but has at least two forms: implicit (non-declarative) and explicit (declarative). This distinction raises the question: What are the cell and molecular mechanisms by which implicit and explicit memories are stored in the brain? Are the molecular storage mechanisms as different as is the logic of the explicit and implicit memory systems?

These two questions (1) the nature of storage mechanisms and (2) the generality and specificity of these mechanisms that have fascinated me most of my scientific career.

I began my studies of memory storage in 1957 by focusing a cell biological approach on the hippocampus, a structure critical to explicit forms of memory storage. But although we were fortunate in making an exciting start on the hippo-

campus, I soon realized that explicit memory storage was exceedingly complex. Without good plastic and pharmacological tools no progress could be made in relating cellular changes to organismic behavior. To achieve an intellectually satisfying understanding of memory storage one needed to take a completely different approach to a memory – one needed to take a more radical reductionist approach. One needed to take an approach to mental processes such as memory similar to that used in other core areas of biology – an approach similar to that used, for example, by Thomas Hunt Morgan, Max Delbruck, and Salvador Luria in the early studies of that nature of the gene in *Drosophila* and bacterial phages. One needed to study not the most complex case but the simplest possible case of memory storage and to study it in the simplest and technically most advantageous animal available so that one could drive it into the ground.

After an extensive search, I focused for reasons I will tell you in a minute on the marine snail *Aplysia* where I have studied a very simple procedural memory. Within a few years after I began to study simple forms of procedural memory, we made the surprising discovery that this simple case of procedural memory shared certain key mechanisms with the most complex forms of explicit declarative memory.

The first clue to shared mechanisms came form the behavioral study of stages in memory storage. We found that this simple instance of procedural memory storage had stages that were surprisingly similar to those described earlier for declarative memory including human declarative memory. There was in each case a short-term memory that was labile and lasted only minutes and a long-term memory that was stable and self-maintained and can lest for days, weeks, or even years. We next found that in implicit, as well as in explicit memory repetition is responsible for converting short- to long-term memory. It is practice that makes perfect. Finally, we found that for implicit memory, as had several years earlier been found for explicit memory that long-term memory, but not short-term, requires new protein synthesis.

This suggested to me the possibility that the requirement for protein synthesis during the switch form short tot long-term memory is very general. *It was evolutionarily conserved and was evident not only in vertebrates but ales in the invertebrate. Moreover, it held for both implicit and explicit forms of learning.*

This generality suggest the possibility that some of the key proteins that make up the switch might also be conserved and used for both explicit and implicit forms of memory storage. If that were so the identifying the relevant proteins in any one system might provide molecular insight that are generally important for understanding other systems. Moreover, the information emerging in different

systems might be sufficiently complementary so as to provide a coherent outline, in molecular detail, of one important issue in the biology of mental process – how a transient short-term memory is converted into stable long-term memory. So with this logic as a background we focused on a molecular approach in Aplysia and then in mice. Before I tell you something about the molecular nature of this switch and complex forms of memory, let me introduce you to Aplysia.

Short-term storage for implicit memory involves functional changes in the strength of pre-existing synaptic connections

Implicit memory refers to memory about perceptual and motor skills. The simplest instances of such storage are the elementary forms of non associative and associative memory. These first emerged with particular clarity as distinctive forms of implicit memory from studies of the family of learning processes related to classical conditioning.

Learning refers to the acquisition of new information about the world and memory refers to the retention of that information over time. Classical conditioning is a form of learning in which an animal learns to associate two sensory stimuli, a neutral initiating sensory stimulus (called *the to be conditioned stimulus* or cs by behaviorist psychologists) that initially produces a weak reflex response, and a highly effective sensory stimulus (called *the unconditioned stimulus* or us by behaviorist psychologists) that produces an inborn, unlearned, reflex response (called the *unconditioned response or* ur*).* As a result of pairing these two sensory stimuli, (the cs and us) the animal learns to strengthen its pre-existing response to the neutral sensory stimulus (the cs) or to develop a completely new response to the cs.

In the course of studying classical conditioning, Ivan Pavlov and others discovered that when each of these two stimuli were repeatedly presented alone, they each gave rise to distinctive forms of learning and memory storage. Thus, repeatedly presenting a neutral stimulus (cs) by itself gives rise to a form of learning called *habituation,* whereby the animal learned to recognize a stimulus as innocuous and comes to ignore it. By contrast, when the animal is presented with an aversive (us) stimulus, it gives rise to *sensitization,* a form of learning whereby the animal recognizes the stimulus as being highly noxious and learns to enhance its defensive and escape responses. Thus, simple forms of learning take one of two forms involve (1) in non-associative learning such as habituation and sensitization the animal learns about the properties of a *single* stimulus, (2) in associative learning the animal learns about the *relation* between two stimuli (Squire and Kandel, 1999).

What are the cellular changes that, result in the brain when animals learn these simple tasks? Initial insights into the cell biological nature of each of these three forms of memory storage first came from studies of the marine snail *Aplysia*. *Aplysia* has lent itself to the study of implicit memory storage because it offers a number of technical advantages for this study. First, the animal is quite smart; it can learn a number of different tasks and store them in both short- and long-term memory. Second, the animal has a relatively simple central nervous system, consisting of only about 20,000 neurons. Third, the neurons of *Aplysia* are particularly large, which allows them to be uniquely identified, so that one can return to the same cell in every animal of the species. Fourth, it is possible to map in detail the synaptic connections between individual cells and between a given cell and the sensory and motor periphery. As a result of these advantages, it is possible to work out significant parts of the neural circuitry of a given behavior – such as the gill and siphon withdrawal reflexes – in terms of uniquely identifiable cells and their pattern of interconnections (Kandel, 1976). Finally, one can culture *Aplysia* neurons and construct with them *in vitro* microcircuits of components of behavior such as the gill withdrawal reflex in ways that are not yet possible in other systems.

Studies of memory in *Aplysia* also first illustrated the advantages that accrue from using elementary forms of nonassociative and associative learning for studying memory storage. These simple, implicit forms of memory offer the advantage that memory can be tested (retrieved) at any time after learning by simply examining the time locked reflex response to a sensory test stimulus (the cs) (Kandel *et al.*, 1995). This feature specifies that the key requirement for a neurobiological analysis of memory is to work out, in cellular detail, the neural circuitry of the behavior being modified by learning. Since each of these simple forms of learning leads to a change in the response of the reflex to the initiating sensory stimulus, one needs in particular to work out the pathway whereby the sensory stimulus of the reflex leads to a behavioral response. As a corollary, in learning tasks in which there is a clearly defined alteration in the behavioral response to the cs, all the important learning-related changes are contained within the circuit of the conditioned stimulus itself.

In Aplysia most work on implicit memory has been carried out on the withdrawal reflex of the gill and the siphon to a weak tactile stimulus applied to the siphon. This withdrawal reflex is mediated by both monosynaptic and polysynaptic connections. The sensory neuron that innervates the siphon make direct monosynaptic connections to the motor neuron that withdraws the gill. In addition, the sensory neurons also make polysynaptic connections to motor neurons

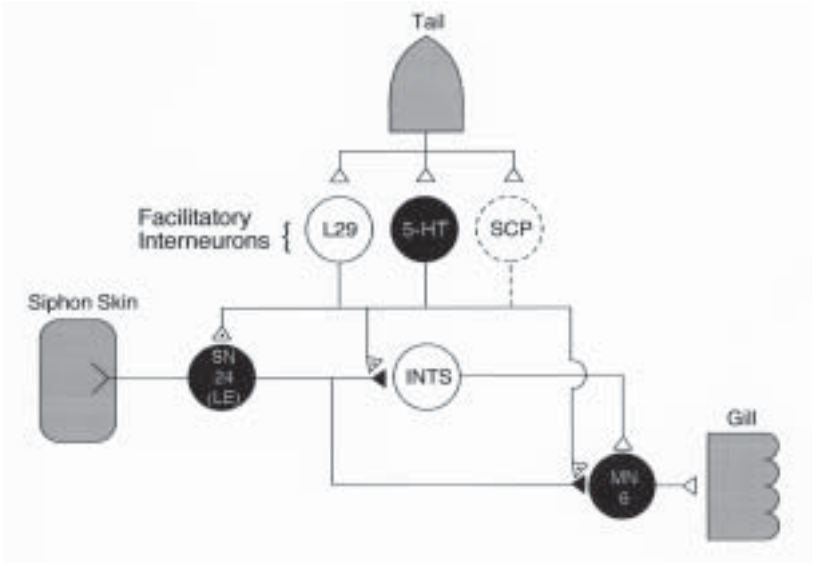


Figure 1. The gill withdrawal reflex in *Aplysia* is mediated by a simple circuit. Approximately 50% of the learning observed in sensitization of this reflex by tail shock results from potentiation by serotonin of the direct synapse between the sensory neuron innervating the mantle shelf and the motor neuron innervating the gill. Individual identified neurons are indicated.

via interneurons (Figure 1.). While the polysynaptic connections contribute importantly to both the basal reflex and to learning, most studies have concentrated on the monosynaptic connections. These connections (1) form a significant component of the behavior, (2) their electrophysiological properties recapitulate the basic properties of both short- and long-term memory for various forms of learning, (3) this component can be cultured and therefore studied in great detail morphologically and biochemically.

Although the gill and siphon withdrawal reflex is quite simple, it exhibits all three simple forms of learning. With each of these forms the animal learns to alter its behavioral response to the tactile stimulus to the siphon (which is the cs for these forms of learning), and in each case repetition converts a short-term form to a long-term form of implicit memory.

With habituation of the gill and siphon withdrawal reflex, the animal learns about the properties of a *single*, novel stimulus, a weak tactile stimulus to the siphon. When this stimulus is first presented, the animal perceives it as novel and responds to it with a brisk reflex response. But when the same weak stimulus is repeated, the animal comes to recognize the stimulus as trivial and gradually

learns to ignore the stimulus and to stop responding to it. As a result the same weak siphon stimulus that once produced a brisk response now produces little or no response at all. This progressive decrease in the response to the weak siphon stimulus is reflected in the neural circuit of the gill-withdrawal reflex, as a *weakening* of the synaptic connections of the sensory input pathway, the synaptic connections between the siphon sensory neurons and their central target cells: the interneurons and motor neurons of the reflex (Figure 1). This weakening in synaptic strength of the sensory input pathway of the cs results from a decrease in the amount of the transmitter glutamate released from the presynaptic terminals of the sensory neurons.

With sensitization, Aplysia learns about the properties of an important aversive stimulus, a noxious shock to the tail (an *unconditioned stimulus).* The animal recognizes the stimulus as aversive and learns to enhance its gill and siphon withdrawal responses to the cs, to the weak touch to the siphon. Sensitization is reflected in the neural circuit as an *increase* in synaptic strength in the input connections of the reflex, the pathway between the siphon sensory neurons end their targeted cells. This strengthening is due to an *increase* in the release of glutamate from, he terminals of the sensory neurons (the cs pathway) (Kandel, 1976; Byrne and Kandel, 1996; Carew and Sahley, 1986; Hawkins *et al.*, 1993).

Aplysia also can learn to associate these two stimuli; it can learn classical conditioning. When a weak (cs) stimulus to siphon is repeatedly paired with a shock to the tail (the us), the reflex response to the siphon stimulus will be enhanced, and this enhancement of the response to the cs with classical conditioning is substantially greater than with sensitization when the weak siphon stimulus (cs) and the tail shock (us) are not paired. This classical conditioning is reflected in the neural circuitry as a greatly enhanced strengthening in the input connections of the sensory neurons to their target cells, an enhancement that is greater than that of sensitization. In addition to a presynaptic mechanism, there is with classical conditioning a postsynaptic mechanism that comes into play (Squire and Kandel, 1999).

There is now a reasonably good understanding of how this association is achieved on the -cellular level (Glanzman, 1995; Squire and Kandel, 1999). As we have seen the sensory neurons release glutamate from their presynaptic terminals. Glutamate acts on two types of postsynaptic receptors: an ampa receptor and an nmda receptor. Under normal circumstances and with nonassociative learning such as habituation and sensitization, only the ampa receptors are utilized because the mouth of the nmda receptor channel is blocked by Mg^{2+}. To remove the Mg^{2+} block and activate this channel, two events need to happen simultaneously:

glutamate needs to bind to the postsynaptic nmda receptor, and the postsynaptic membrane needs to be depolarized substantially so as to extrude Mg^{2+} out of the nmda receptor channel mouth. This coincident activation of the nmda receptor and postsynaptic depolarization only occur when the weak siphon stimulus (cs) and the strong tail shock (us), are paired together. Only then is each of the postsynaptic motor cells sufficiently depolarized to activate the nnwa receptors (Lin and Glanzman, 1994).

Detailed analyses of the distinctive role of the nnwa receptor in the postsynaptic cell were first carried out in the hippocampus where, as we shall learn below, it was found that the extrusion of Mg^{2+} allows Ca^{2+} to flow into the postsynaptic cell. The Ca^{2+} influx in turn activates a signaling cascade in the postsynaptic cell. One component of long-term potentiation *in Aplysia* is a retro-grade signal that is generated in the postsynaptic cell and sent back from it to the presynaptic neuron where the signal acts to enhance presynaptic transmitter release to that it is even greater than occurs with sensitization alone. Thus, in the case of *Aplysia*, the facilitation of the connections between the sensory and motor neurons that occurs with classical conditioning has an additional, associative component superimposed on the facilitation produced by sensitization (Squire and Kandel, 1999).

These several studies in *Aplysia* revealed three insights into short-term memory storage that have proven quite general and apply as well to explicit as to implicit memory storage. First, these studies showed that learning can lead to alterations in synaptic strength and that the persistence of these synaptic changes in both the mono- and polysynaptic component of the neural circuit of the reflex represent the cellular storage mechanisms for memory. Second, a single synaptic connection can participate in, and be modified in different ways by, different forms of learning and participate in different types of short-term memory storage. Finally, each of these three simple forms of learning – habituation, sensitization, and classical conditioning - gives rise to both a short- and a long-term memory depending upon number of repetitions. Each of the long-term forms of memory is associated with a long-term change in synaptic strength in the monosynaptic connection between the sensory and motor neuron of the reflex. Thus, a single synaptic connection can not only participate in different types of short-term memory storage but the same connection can also be the site of both short- and long-term memory storage.

Long-term storage for implicit memory involves the synthesis of new protein and the growth of new connections

Given that a single connection can participate in both short and long-term memory, what are the molecular mechanisms for these different phases of memory storage? In particular, how is short-term memory converted to long-term memory? Can molecular biology with its ability to reveal homology relationship delineate commonalties in the mechanism of storage that might again encompass explicit as well as implicit storage?

The first insight into the molecular mechanisms of memory storage come from the discovery that there were phases of memory storage. The study of memory phases dates to 1885 and the work of Herman Ebbinghaus. By forcing himself to memorize lists of nonsense syllabus and then testing his own recall, Ebbinghaus determined that there are at least two phases to memory: a short term phase which contains much information but is transient, lasting minutes, and a long-term phase which is far more stable (Ebbinghaus, 1913). This is consistent with our everyday experience: we have access to far more of the information of the recent past (say, the last few minutes) than we will be able to remember a few hours hence. Ebbinghaus' distinction is also consistent with observations of victims of injury: a person who is struck on the head or shocked will typically lose memory for events that occurred shortly before the insult but not for more remote events (e.g., Zubin and Barrera, 1941). Modern work showed that the distinction between short- and long-term memory applies to both implicit and explicit memory.

In the 1960s Louis Flexner and his colleagues first found that short- and long-term memory are not only distinguished by their time course but also by their biochemical mechanism. Long-term memory differs from short-term memory in requiring the synthesis of new protein (Flexner *et al*., 1965; Agranoff, 1976). This requirement of protein synthesis for long-term memory has a specific time window, during and shortly after training, which is called *the consolidation phase.* Blockade of protein synthesis during the consolidation phase will disrupt long-term memory, but blockade before or after will have no effect (Bourtchouladze *et al*., 1998; Freeman *et al*., 1995).

The requirement for protein synthesis for long-term but not short-term memory, and the existence of a consolidation window during which memories are sensitive to disruption, has proven to be very general. lt has been demonstrated in explicit as well as implicit storage and in different vertebrates as well as in invertebrates. This conservation in turn suggests that the proteins involved in the switch to long-term memory may also be conserved. If that were true, then a detailed

study of the molecules involved in the switch in any given memory storage process in any animal is likely to yield proteins that are of general importance. Moreover, a molecular study of several different instances of memory storage are likely to reveal the general nature of a cognitive process: the switch whereby a transient short-term memory is converted to a persistent, self-main-tained long-ten-n memory (Pittenger and Kandel, 1998). During the last decade, studies in *Aplysia, Drosophila,* and mice have begun to reveal some of the proteins essential for this switch.

Aplysia *and* Drosophila *share some of the same genes and proteins for converting short- to long-term memory*

The initial molecular insights into long-term storage of implicit memory came from studies of sensitization in *Aplysia.* As in other forms of learning, repetition of the sensitizing training in *Aplysia* increases the duration of the memory for sensitization. Thus, one tail shock produces an enhancement of the withdrawal response that lasts a few minutes. Five shocks produce an enhancement that lasts several days, and further training gives rise to memory that lasts weeks. This long-lasting sensitization requires protein synthesis during a critical time window, whereas short-term sensitization does not (Montarolo *et al.,* 1986). Here, then, is a very simple model system of long-ten-n memory that shares some of the key mechanistic properties of more complex vertebrate systems.

Sensitizing tail stimuli activate three different classes of modulatory inter-neurons that synapse on the axon terminals of the siphon sensory neurons; all three have similar actions (Figure 1). Of the three, the interneurons that release serotonin (or 5-hydroxytryptamine, 5-ht) are thought to be particularly important. In the intact animal, in reduced preparations, and even in a dissociated cell culture consisting of a single sensory neuron, a single motor neuron and a single serotonergic facilitatory neuron (Figure 2, see next page), serotonin acts on the sensory neurons by increasing the intracellular concentration of the second messenger camp and by activating the camp dependent protein kinase (pka) and protein kinase C (Braha *et al*., 1990; Brunelli *et al*., 1976). Transient activation of these intracellu-lar signaling cascades by one tail shock or one pulse of 5-ht leads to a transient strengthening, or facilitation, of the synapse between the sensory and motor neurons by increasing the amount of neurotransmitter (glutamate) released by the sensory cell onto the motor cell when the siphon is touched. Repeated training or of five pulses of 5-ht produces long-lasting facilitation that can persist for 72 hr or more and is accompanied by the growth of new connections (Castellucci *et al.,* 1986; Montarolo *et al.,* 1986; Bailey and Kandel, 1993).

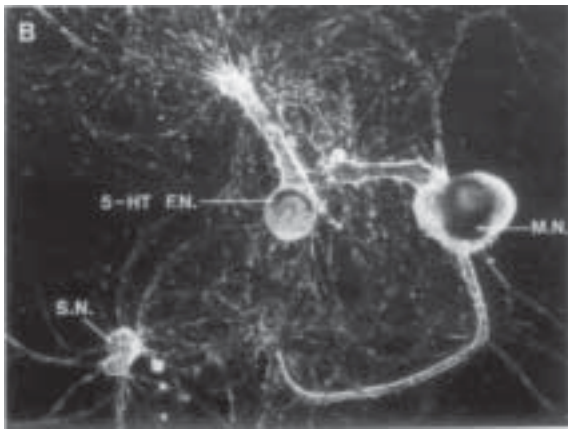


Figure 2. The circuit controlling this reflex can now be studied in reconstituted cell culture. (S.N. – sensory neuron. 5-ht F.N. – facilitatory interneuron. M.N. – motorneuron.)

Whereas a single pulse of 5-ht activates the kinases pka and pkc transiently, five repeated pulses lead to a persistent activation of pk and to the recruitment of the map kinase signal transduction pathway (Figure 3). Both pka (Bacskai *et al*., 1993) and map kinase (Martin *et al*., 1997b) then translocate to the nucleus, where they activate the transcriptional activator crebla (Bartsch *et al.*, 1998) and inactivate the transcriptional repressor creb2 (Bartsch *et al.*, 1995). The creb family of transcriptional activators have been implicated in plasticity in many systems, as we will discuss below, and may be one of the most conserved molecular components of the mechanisms for switching on long-term plasticity. Once crebla is activated and the repressive action of creb2 is removed, a set of immediate-early genes is activated, of which two – the N-terminal ubiquitin hydrolase and the transcriptional factor ApC/EBP (Alberini *et al*., 1994) – have been well characterized. ApC/EBP, forms both homodimers and heterodimers with another factor (activating factor l). The homodimers and heterodimers act different on downstream genes that lead to the growth of new synaptic connections. This structural change, which represents the long-term, self-maintained, stable form of memory, is associated with a rearrangement of structural proteins such as the cell adhesion molecule ApCAM (Bailey and Kandel, 1993). This internalization of ApCAM, for example, is thought to be a necessary prerequisite for the growth of neuronal processes (Figure 3).

The requirement for transcription in long-term facilitation in *Aplysia* explains why long-term memory requires the synthesis of new protein. However, this

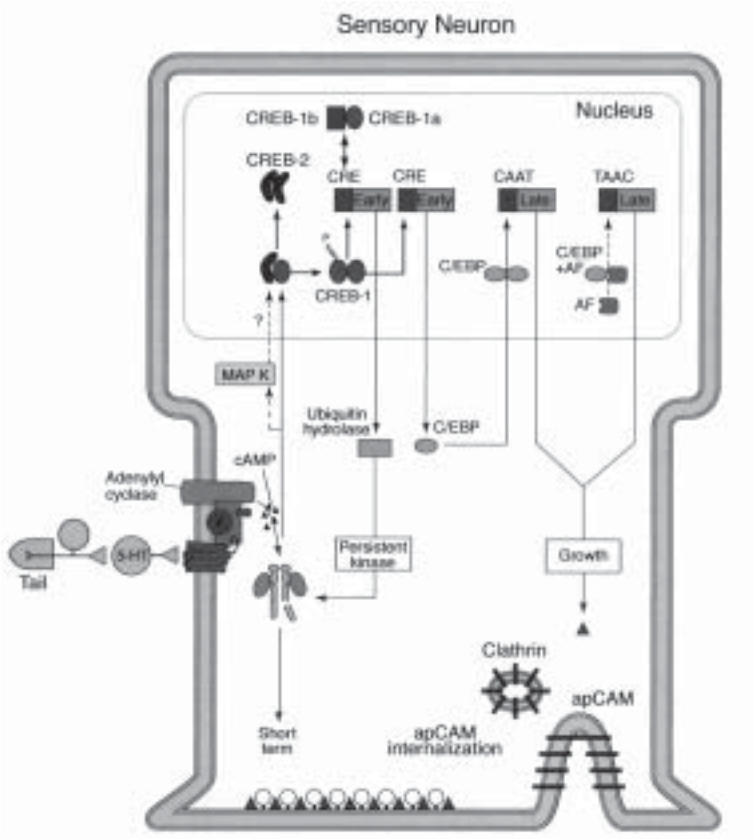


Figure 3. Multiple molecular pathways are involved both in short-term facilitation and in long-term facilitation of the sensory neuron-motor neuron synapse in *Aplysia.*

requirement poses a cell biological puzzle: since long-term plasticity relies on the activation of genes in the nucleus, one might expect that long-lasting changes in the connectivity of the neuron would be cell-wide. Recent experiments have revealed that each synapse or group of synapses can be modified independently. This spatially restricted plasticity requires the activity of crebi in the nucleus as well as local protein synthesis in the processes that were modulated by 5-ht. The mechanism of this specificity is revealed by a second phenomenon: *synaptic capture.* When synapse- specific long-term facilitation is initiated at one of two branches, a single pulse of 5-ht – which alone is able to induce only a transient facilitation lasting minutes – is thereafter able to induce long-term plastic changes including the growth of new synaptic connections when applied to a second

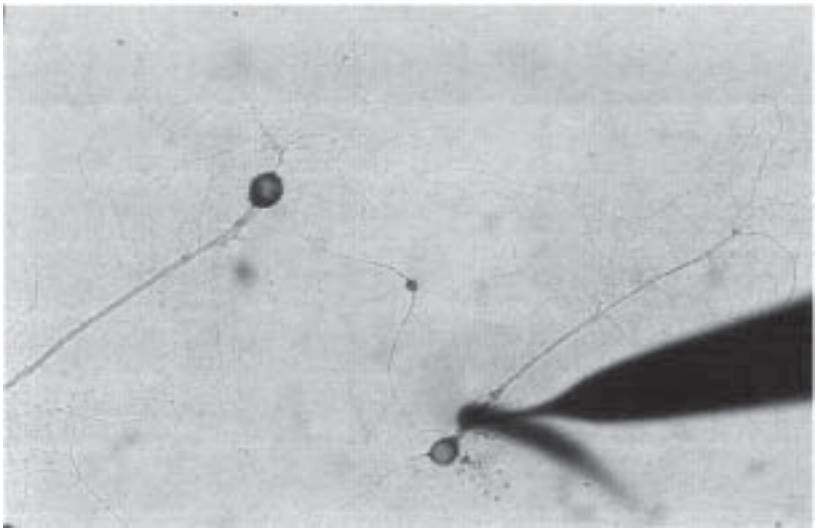


Figure 4. The ability to study two synapses from the same sensory neuron onto widely separated motor neurons in a modified *Aplysia cell* culture system allows synaptic tagging and synaptic capture to be explored for the first time.

branch (Figure 4) (Martin *et al*., 1997a). This phenomenon suggests that the new genes that are being activated in the nucleus have their products distributed widely, but that the products only persistently strengthen those synapses which have been somehow marked by short-term facilitation. This phenomenon is also evident in the vertebrate brain (Frey and Morris, 1997).

A similar set of genes important in the switch from short- to long-term memory has also emerged from studies of *Drosophila.* As an experimental system, *Drosophila* is in many ways the complement of *Aplysia.* The great advantage of *Aplysia* is that it is tractable in a top down approach for cell biological studies of learning, and allows in the limit the reconstitution of synaptic circuits in cell culture. Mutational (forward) genetics, however, is almost impossible in *Aplysia.* In *Drosophila,* on the other hand, cell biological and electrophysiological studies are difficult due to the tiny size of the neurons, but the genetics are extremely tractable and mature – more so than in any other multicellular organism. The great strength of the *Drosophila* work in learning, predictably, is therefore in its genetic analysis. Both genetic screens and reverse genetic analysis, in which specific genes have been disrupted to investigate their function, have been fruitful.

The pioneering work of Seymour Benzer, and the subsequent studies of Quinn, Tully, and Yin, has led to the identification of a number of genes required for memory storage (Weiner, 1999). Many of the genes identified in this way are the same as those implicated in plasticity in *Aplysia,* hinting at a dramatic conservation of the mechanisms of synaptic plasticity over evolution. For example, the *Drosophila genes dunce, rutabaga, and amnesiac* all encode components of the camp-pka cascade (Figure 5). Other genes identified encode participants in other signal transduction cascades (for example, the gene *leonardo;* Skoulakis and Davis, 1998) or cell-cell adhesion molecules similar to ApCAM (Grotewiel *et al.*, 1998). Based on the genes that have been identified, there seems to be a clear

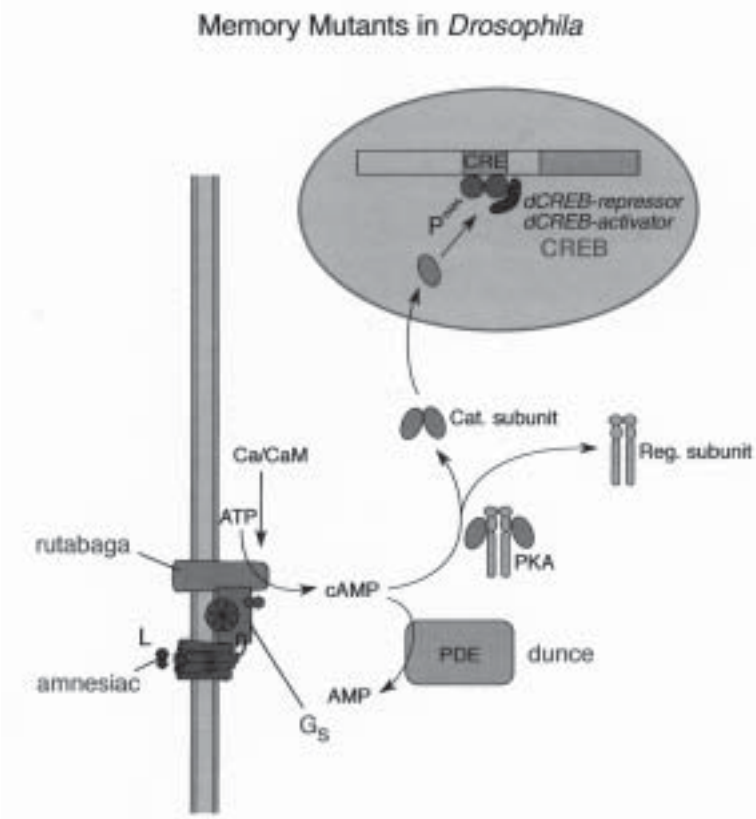


Figure 5. Many of the learning-related genes identified by screens in *Drosophila,* including dunce, *rutabaga, and amnesiac,* participate in the cyclic amp-pka pathway that is also implicated in *Aplysia.*

conservation of mechanism between *Drosophila* and *Aplysia.* For example, a protein synthesis-dependent phase of learning has been described by Tully and his colleagues, and Yin and Tully have shown that, as in *Aplysia*, creb has a critical role in the induction of long-term memory. They found that a dominant negative creb allele blocked the formation of long-term memories (Yin *et al.*, 1994), whereas overexpression of the wild type allele enhances long-term memory stabilization (Yin *et al.*, 1995). This work shows that creb has a role in learning in *Drosophila* that is similar or identical to its role in *Aplysia*, demonstrating striking evolutionary conservation.

Finally, Corey Goodman and his colleagues have used the neuromuscular junction of *Drosophila* to examine the developmental plasticity of nerve-muscle synapse. Goodman and colleagues have found that at this synapse, creb also is required for the induction of transcription-dependent plasticity. But whereas creb is required for functional plasticity at this synapse, it is not sufficient for morphological plasticity. For morphological plasticity, the loss of the cell adhesion molecule Fasl, a homolog of ApCAM, is required, much like the internalization of ApCAM is required for learning-related growth in *Aplysia* (Schuster *et al.*, 1996).

Explicit memory storage in the mammalian brain: spatial memory, long-term potentiation, and the role of CREB in mice

What about explicit memory storage? As we have seen, explicit memory involves the conscious recall of facts and events. It is therefore more complex than implicit memory storage for two reasons. First, explicit memory involves conscious participating in memory recall whereas implicit recall is unconscious. Second, one cannot with explicit memory readily define a simple cs pathway; rather, explicit memory involves the integration of several sensory cues. Nevertheless as the technology for manipulating the mouse genome improved, it became clear that mice are likely to be powerful systems for the study of explicit forms of memory. As is the case with humans, mice and other rodents require the hippocampus for spatial memory and navigation (O'Keefe and Nadel, 1978) as well as for object recognition (Gaffan, 1998). Spatial memory seems a good model for explicit memory in humans, because both require for its formation not a simple cs but an arbitrary association between several different sensory cues characteristic of explicit memory. Finally, the hippocampus, which is critically involved in spatial and other forms of explicit memory, exhibits several well-studied forms of synaptic plasticity.

The basic hippocampal circuit consists of a three-synapse loop – the perforant pathway, the mossy fiber pathway, and the Schaffer collateral pathway - which

runs from the entorhinal cortex to the cai region of the hippocampus. In groundbreaking work, Bliss and Lomo (1973) found in 1973 that when the input from the entorhinal cortex to the dentate gyrus (the first cell field of the hippocampal formation) is stimulated repetitively at high frequency, the synapse is persistently strengthened. They termed this phenomenon long-term potentiation, or ltp. Subsequent work has shown that each of the three synapses in the hippocampal loop exhibits long-term potentiation, and that at each synapse ltp has an early and a late phase which can be induced by different stimulus protocols (Figure 6, see next page). In each case, the late phase differs from the early phase in that it is blocked by drugs that inhibit protein or mrna synthesis. Moreover, the induction of this late phase requires camp and pka at all three synapses. The similarities to the molecular mechanisms involved in the late phase of ltp and in *Aplysia* provided the first suggestion of conserved mechanisms for converting short-term memory to long-term memory in mammals (Abel and Kandel, 1998; Huang *et al*., 1994).

Most work has focused on the third synapse in the loop from the entorhinal cortex through the hippocampal formation, the Schaffer collateral synapse, for several reasons. First, it is technically a particularly easy synapse to study. Second, studies in humans by Squire and colleagues have found that a patient with a hypoxic lesion limited to the cai field (the location of the Schaffer collateral synapse) had a significant amnesia, suggesting that damage restricted to this area is sufficient to disrupt memory formation (Zola-Morgan *et al*., 1986). Third, whereas dissociations between synaptic plasticity and behavior in mice have been described at other synapses in the hippocampal formation (sec, for example, Huang *et al*., 1995), the efforts of many labs have built up a long list of correlations between disruptions of Schaffer collateral ltp and disruptions of hippocampus-dependent memory. Although there are occasional dissociations, even at this synapse (see for example Zamanillo *et al*., 1999), this synapse nevertheless seems a good place to start in disentangling the role of plasticity in the complex hippocampal circuit.

There is consensus that induction of ltp in the Schaffer collateral pathway involves activation of the nmda receptors (Nicoll and Malenka, 1999). As we have seen earlier, unlike other glutamate receptors, which are activated whenever the presynaptic neuron releases glutamate onto the postsynaptic membrane, the nmda receptor is only activated when the release of glutamate by the presynaptic cell is accompanied by substantial depolarization of the postsynaptic cells, as would occur by coincident firing of the pre- and postsynaptic neuron (Hebb, 1949). The nmda receptor is therefore ideally suited to a role in initiating plasti-

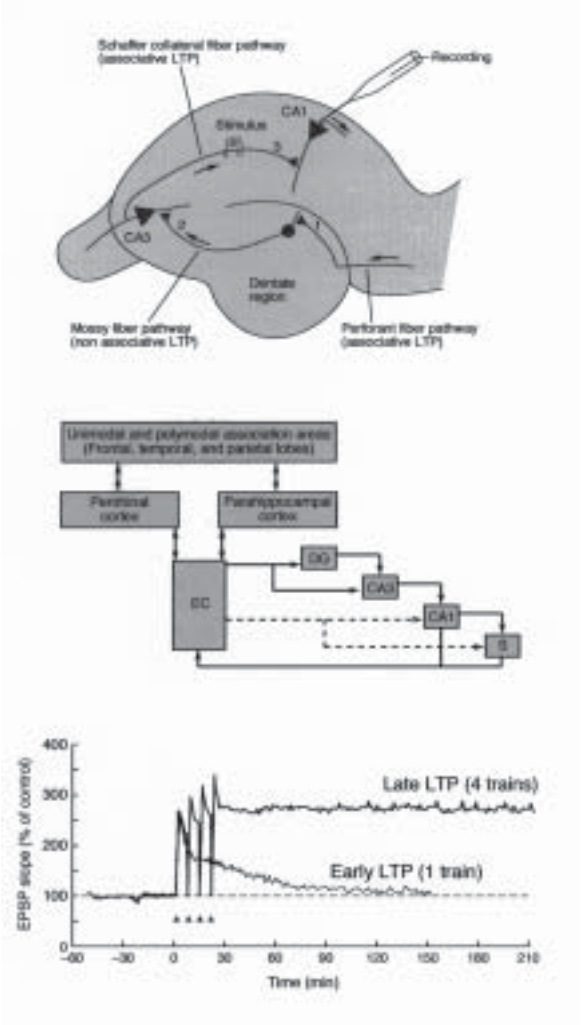


Figure 6.

A. In the mammalian brain, plasticity has been most intensively studied in the hippocampus, and particularly at the Schaffer collateral synapse between ca3 and ca1 pyramidal cell fields.

B. The hippocampus consists of a series of interconnected cell fields. Plasticity has been described at all synapses in this circuit.

C. ltp at the Schaffer collateral synapse has an early and late phase. The early phase, induced by a single high-frequency train of stimuli, lasts 60-90 minutes and is independent of macromolecular synthesis. The late phase, induced by repeated high-frequency trains, lasts many hours and depends on gene induction and the synthesis of new proteins.

city, where synaptic strengthening results from coincident firing (Hebb, 1949). Activation of the nmda receptor leads to influx of Ca^{2+} into the postsynaptic cell and activation of CaMKII. Despite much work, however, there continues to be significant controversy as to the subsequent events. Whether the expression of early ltp is post- or presynaptic remains unclear. We will again focus on the late, transcription-dependent phase.

Whereas a single tetanus activates CaMKII which induces early-phase ltp, multiple tetani increase intracellular camp and thus also recruit pka and map kinase, which induce nuclear (Figure 7) (Abel *et al*., 1997; Impey *et al*., 1998a).

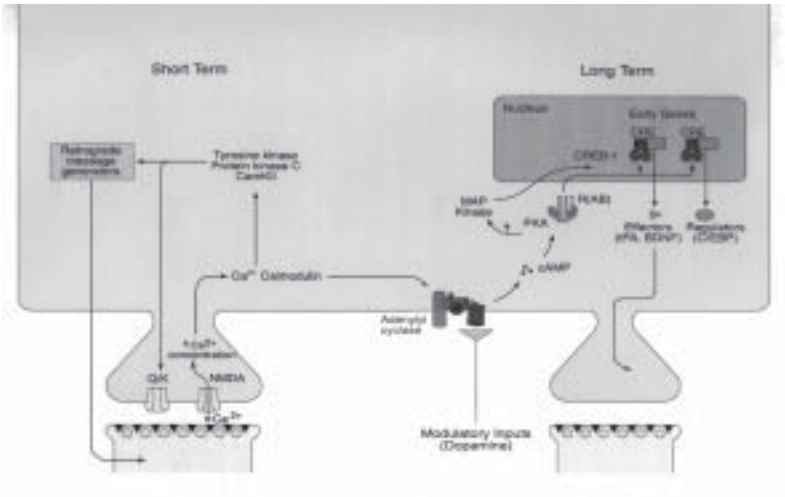


Figure 7. Schematic view of short- and long-term mechanisms for hippocampal E-ltp and L-ltp in the Schaffer collateral pathway of the hippocampus.

Paralle-ling the finding that mapk is required for long-ten-n facilitation in *Aplysia,* David Sweatt and his colleagues have found that mapk is activated both *in vitro* upon induction of ltp (English and Sweatt, 1997) and *in vivo* after certain forms of learning (Atkins *et al*., 1998). Although the mechanism remains obscure, there is currently a consensus that pka activates the mapkcascade in neurons and that mapk is particularly important in the activation of nuclear factors required for the late phase.

The sine qua non of L-ltp in mice is, of course, the requirement of transcription and induction of genes in the potentiated cell. In another example of evolutionary conservation, creb is once again implicated. This was first suggested in a study

by Silva and colleagues, who found L-ltp and hippocampus-dependent learning to be disrupted in a mouse with a knockout of the two most prevalent forms of creb (Bourtchouladze *et al.*, 1994). However, this study is complicated by the nature of the knockout, which only eliminated some creb alleles (Blendy *et al.*, 1996), and by the complex nature of creb-regulated transcription in mammals. Indeed, a complete knockout of all mouse creb isoforms is embryonically lethal (Rudolph *et al.*, 1998). Furthermore, genes related to creb are altered in the partial creb knockout mice in compensation for the missing creb (Hummler *et al.*, 1994).

Despite these complications, however, recent work from Daniel Storm and his colleagues has provided strong, albeit still circumstantial, evidence that creb is indeed involved in mouse plasticity. They produced a transgenic mouse in which a *lacZ* reporter gene is activated by a creb-responsive promoter, and they found that this reporter is activated both by L-ltp *in vitro* (Impey *et al.*, 1996) and by certain forms of hippocampus-dependent learning *in vivo* (Impey *et al.*, 1998b). This demonstrates that creb (or creb-like transcription factors) is in fact induced under circumstances that lead to plasticity and suggests a causal role similar to that seen in *Aplysia.*

As we have noted above, the transcriptional regulation of genes required for ltf in *Aplysia* is more complicated than the regulation of a single transcription factor, involving several other elements. Substantially more complex interactions are likely to be encountered in mice. To work out in detail the molecular switch for converting short-term to long-term facilitation in the Schaffer collateral pathway, much more work will be required.

Structural changes and the biological basis of individuality

As we have seen, structural changes activated by creb-1 are the defining features of long-term memory storage in invertebrates. There is beginning evidence as well for structural changes with ltp in the hippocampus. Thus, several studies suggest that a stably potentiated synapse can release multiple packets of neurotransmitter (or quanta), whereas an unpotentiated synapse releases either zero or one (Bolshakov *et al.*, 1997; Bolshakov and Siegelbaum, 1995). Consistent with this, some ultrastructural studies have found that in tissue in which potentiation has been induced, synapses can have multiple active release zones, or even be split into two. Given the widely held belief that each active zone can release at most one quantum of neurotransmitter, this sort of splitting is precisely what may be needed to explain the observed release of multiple quanta (see, however, Sorra and Harris, 1998). Finally, recent studies indicate that ltp leads to the outgrown

of dendritic processes that appear to be percursors of dendritic spines (Maletic-Savatic *et al*., 1999).

Structural plasticity has been best studied in development, and in sensory rewiring in the adult. Developmental plasticity has been studied in the visual system of frogs, cats, ferrets, and monkeys and in the whisker-dedicated portion of the somatosensory cortex, or *barrel cortex, in* rodents. Most often studies involve systemically perturbing the relevant sensory input, such as by blinding one eye or trimming a subset of whiskers, and observing the results. Such studies in kittens have revealed that plasticity in the developing visual system, like adult plasticity at the Schaffer collateral synapse, requires activation of the nmda receptor (Gu *et al*., 1989). Furthermore, in the mouse, monocular removal of visual input upregulates expression of a creb-driven reporter gene in the portion of visual cortex where synaptic plasticity is occurring (Pham *et al*., 1999).

Cortical plasticity in adults has been demonstrated in both sensory and motor cortices in several systems. The clearest demonstrations come from work by Merzenich and Kaas in the somatosensory cortex of the monkey. When somatosensory input is altered by nerve section (Kaas *et al*., 1983) or by amputation (Merzenich *et al*., 1984), a portion of the corresponding primary somatosensory cortex is deprived of its input. Over time the surrounding regions, which still receive normal input, expand into the deprived area, so that the somatotopic map over the surface of the cortex is altered. In the converse experiment, Merzenich and colleagues showed that when a normal monkey is trained to preferentially use only some fingers, the cortical representation of those fingers expands (Jenkins *et al*., 1990).

Dramatic evidence for cortical reorganization in normal humans has been provided by studies from Taub and colleagues. They scanned the brains of string instrument players. During performance, string players are continuously engaged in skillful hand movement. The second to fifth fingers of the left hand, which contact the strings, are manipulated individually, while the fingers of the right hand, which move the bow, do not express as much patterned, differentiated movement. Brain images of these musicians revealed that their brains were different from the brains of non-musicians: the cortical representation of the fingers of the left hand, but not of the right, was larger in the musicians (Elbert *et al*., 1995). Such structural changes are more readily achieved in the early years of life. Indeed, Taub and his colleagues found that musicians who learned to play their instruments by the age of 12 years had a larger representation of the fingers of the left hand than did those who started later in life.

These several studies suggest that long-term memory storage lead to anatomical

changes in the mammalian and even the human brain much as it does in *Aplysia*. The finding of the anatomical changes are potentially significant for understanding the biological basis of individuality for they suggest that even identical twins that share an identical genome are likely to have somewhat different brains because they are certain to have somewhat different life experiences.

Place cells as a test of models of memory in the hippocampus

How can we bridge the gap between synaptic plasticity and spatial memory? In spatial learning, the animal is thought to use the sensory input – visual, proprioceptive, vestibular, olfactory – to develop an internal representation of its spatial environment and then to use that representation for spatial navigation. Thus, one step to working out the neural circuitry of the sensory stimuli that converge to produce the representation of space is to study how the hippocampus represents space. This is thought to occur by the ability of the animal to use the pyramidal cells of the hippocampus to form a spatial map of the environment.

The hippocampal pyramidal cells are place cells. They fire when an animal occupies a specific area in its environment (this area is called the *place field* for a specific cell). Any pyramidal cell can function as a place cell and, in any given environment, about half of the cells in the hippocampus function as place cells. These place cells form a 'cognitive map' of the environment, and learning a new environment and storing that new representation in the hippocampus involves the creation and stabilization of a new cognitive map (O'Keefe and Nadel, 1978). In a new environment the firing field of a place cell forms within a period of minutes (comparable to the acquisition of a learning task), and once formed can be stable for months (Muller *et al*., 1987), comparable to long-term memory. The map is environment-specific, in that the place field in one environment does not in any way predict the place field in another environment (Muller and Kubie, 1987). A variety of manipulations of the animal can lead to a *remapping*, in which the place fields of all cells change. Such a remapping may correspond to disorientation, in which the animal no longer recognizes that it is in a familiar environment.

Research on how the network properties of place cells in the hippocampus are remapped in a new environment by learning, represents perhaps the best example of research on the plasticity in an important cellular network involved in explicit memory storage. An analysis of plasticity in a system of place cells during spatial learning represent an important advance in the molecular biology of cognition that we advocate for the next century. With such research the top down and bottom up modes of analysis begin to converge (Rolls and Treves, 1998).

To bring this fine-grained systems level analysis together with a molecular one

it requires investigating the mechanisms of plasticity in a population of place cells that may be involved in the learning of a new environment. There is now evidence based on such an approach that hippocampal ltp may be involved in the stabilization of place field as an environment becomes familiar. When an nmda blocker (the drug cpp) was given to rats during exposure to a novel environment, the animals formed normal place fields, and (surprisingly) those fields were stable for an hour. However when the animals were returned to the environment after 24 hours, a complete remapping occurred: the place cell map of the environment was not stable. Importantly, the place cells in a control environment with which the animal had been familiarized before drug treatment remained stable (Figure 8) (Kentros *et al.*, 1998). This study demonstrates that pharmacological treatment that blocks hippocampal ltp can disrupt the stability of hippocampal place cells and supports the idea that an ltp-like phenomenon underlies spatial learning in rodents.

Studies with genetically modified animals also support this general conclusion.

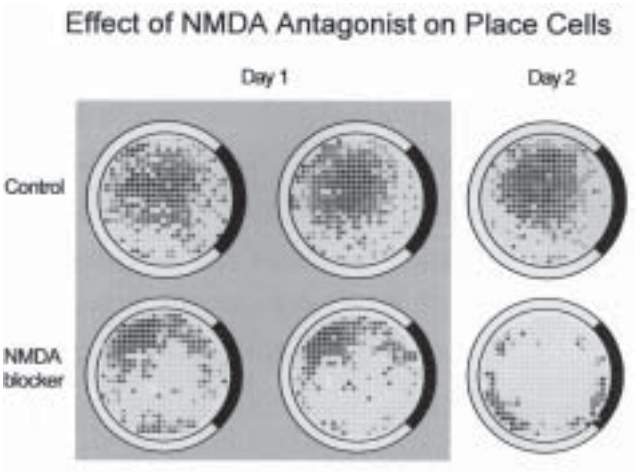


Figure 8. Place cells in the hippocampus represent the most advanced current example of a fusion of systems and molecular investigations of memory. Firing rates of two place cells are shown here. Each circle corresponds to an environment the animal explored; colored pixels correspond to firing rate at different points in the environment. Firing rate is highest in a particular region of the environment, the place field. In the upper cell, from a control rat, the place field is stable across three recording sessions on two days. In the lower cell, from a rat treated with the nmda blocker cpp, the place field is stable over two training sessions on the first day (separated by an hour) but is dramatically altered 24 hours later. Adapted from Kentros *et al.*, 1998.

Genetic manipulations of the nmda receptor (McHugh *et al*., 1996), CaMKII (Rotenberg *et al*., 1996), and pka (Rotenberg *et al*., 1997) all show that place fields form fairly normally without nmda-dependent plasticity, but are not stable in the long term. These observations amount to significant evidence (albeit still correlational) for a connection between the molecular mechanisms of synaptic plasticity, place cell stability, and spatial learning as assessed by behavioral measures. If this correlation is substantiated it will represent a major step towards a unification of molecular and systems approaches to the study of spatial memory formation in rodents, and a major step towards the molecular cognition of the future.

We have here only considered one modal point in the medial temporal lobe circuit with the hippocampus. One will need to move backwards from the cai region to work out the nature of the sensory input into the hippocampus, that gives rise to a spatial map, how the sensory information are in the cortex? How is it transformed at various relays in the medial temporal lobe, and how is that transformed information used by the motor system for spatial navigation. As these arguments illustrate, this analysis which is extremely ambitious and will occupy us well into the distant future. It illustrates one of the key task that confronts the analysis of the systems components of memory storage to which we now turn.

Conclusion

The 1990s were proclaimed as the Decade of the Brain in acknowledgement of both the unique challenges posed by the human nervous system and the rapid pace of discovery in neuroscience. The last ten years we have indeed experienced an accelerating rate of discovery in many fields of neuroscience, including in memory research. The gradual unification of neurobiology with cognitive psychology and the subsequent emergence of a molecular biology of cognition have been very fruitful and has increased substantially our understanding of the mechanisms of memory formation.

We now have reached the end of both the 20th century and of the Decade of the Brain. We have at this point a clearer understanding of biologically meaningful subdivisions of memory storage and clearer understanding in outline of some molecular mechanisms of storage relevant to each of these subdivisions. Most impressive is the finding that explicit and implicit storage seems to use a common and limited set of mechanisms to convert short- to long-term memory.

Whereas satisfactory insight into even the details of the storage mechanisms are in sight, the systems problems are much more difficult and will continue to occupy us for many decades. This is because the anatomical systems that store explicit

memory are complex as is the multivariant nature of the memory that is stored. Moreover, explicit memory requires for its recall conscious recollection, an area which still elude us. To dissect the complexities of explicit memory will therefore require a substantial period of time – probably another century. It will also require the development of much further methodology, multiple unit recordings, better cellular imaging, as well as better genetic methodology for studying cognition. It therefore will be advisable to continue to analyze instances of implicit memory storage, including the simple implicit memory systems of invertebrates and vertebrates, and to use them as prototypes for understanding the more complex explicit systems. Because explicit memory storage is so deeply embedded in perception, action and consciousness, its future is the future of neuroscience. As a result, insofar as understanding explicit memory is predicated on understanding much of the brain, the mass action arguments of Flourens and Lashley, and the localization arguments of Gall, Penfield and Milner' will, in the long run, be joined.

Acknowledgements

This lecture is a shortened form of a paper published in the Centennial Issue of the Royal Society. I am indebted to Chris Pittenger who collaborated with me on that original manuscript. I also am greatly indebted to the students and post-doctoral fellows with whom I have been privileged to collaborate and who have contributed so importantly to the work I here report. I have tried to cite their individual contributions throughout this manuscript. My research has been supported by the Howard Hughes Medical Institute and by the N.I.H.

References

Abel, T. and Kandel, E. (1998) Positive and negative regulatory mechanisms that mediate long-term memory storage. *Brain Res. Rev.* 26:360-378.

Abel, T., Nguyen, P.V., Barad, M., Deuel, T.A.S., Kandel, E.R., and Bourtchouladze, R. (1997) Genetic demonstration of a role for pka in the late phase of ltp and in hippocampus- based long-term memory. *Cell* 88:615-626

Agranoff, B.W. (1976) Biochemical concomitants of the storage of behavioral information. *Biochem. Sensory Funct.* pp. 597-623.

Alberini, C.M., Ghirardi, M., Metz, R., and Kandel, E.R. (1994) ciebp is an immediate-early gene required for the consolidation of long-term facilitation in *Aplysia*. *Cell* 76:1099-1114.

Atkins, C.M., Selcher, J.C., Petraitis, J.J., Trzaskos, J.M., and Sweatt, J.D. (1998) The mapk cascade is required for mammalian associative learning. *Nat. Neuro-sci.* 1:602-609.

Bacskai, B.J., Hochner, B., Mahaut-Smith, M., Adams, S.R., Kaang, B.K., Kandel, E.R., and Tsien, R.Y. (1993) Spatially resolved dynamics of camp and protein kinase A subunits *in Aplysia* sensory neurons. *Science* 260:222-226

Bailey, C.H., and Kandel, E.R. (1993) Structural changes accompanying memory storage. *Annu. Rev. Physiol.* 55:397-426.

Bartsch, D., Casadio, A., Karl, K.A, Serodio, P., and Kandel, E.R. (1998) *CREBI* encodes a nuclear activator, a repressor, and a cytoplasmic modulator that form a regulatory unit critical for long-term facilitation. *Cell* 95:211-223.

Bartsch, D., Ghirardi, M., Skehel, P.A., Karl, K.A., Herder, S.P., Chen, M., Bailey, C.H., and Kandel, E.R. (1995) *Aplysia* creb2 represses long-term facilitation: relief of repression converts transient facilitation into long-term functional and structural change. *Cell* 83:979-992.

Blendy, J.A., Kaestner, K.H., Schmid, W., Gass, P., and Schutz, G. (1996) Targeting of the creb gene leads to up-regulation of a novel creb mrna isoform. *Embo J. Is*: 1098- 1106.

Bliss, T.V. and Lomo, T. (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol. (Lond.)* 232:331-356.

Bolshakov, V.Y., Golan, H., Kandel, E.R., and Siegelbaum, S.A. (1997) Recruitment of new sites of synaptic transmission during the camp-dependent late phase of ltp at ca3-cai synapses in the hippocampus. *Neuron* 19:635-651.

Bourtchouladze, R., Abel, T., Berman, N., Gordon, R., Lapidus, K., and Kandel, E.R. (1998) Different training procedures recruit one or two critical periods for context memory consolidation, each of which requires protein synthesis and pka. *Learning & Memory* 5:365-374.

Bourtchouladze, R., Frenguelli, B., Blendy, J., Cioffi, D., Schutz, G., and Silva, A.J. (1994) Deficient long-term memory in mice with a targeted mutation of the camp-responsive element-binding protein. *Cell* 79:59-68.

Braha, O., Dale, N., Hochner, B., Klein, M., Abrams, T.W., and Kandel, E.R. (1990) Second messengers involved in the two processes of presynaptic facilitation that contribute to sensitization and dishabituation in Aplysia sensory neurons. *Proc. Natl. Acad. Sci. USA* 87:2040-2044.

Broca, P. (1865) Sur le siège de la faculté du language articulé. *Bull. Soc. Anthropol.* 6:377-393.

Brunelli, M., Castellucci, V., and Kandel, E.R. (1976) Synaptic facilitation and behavioral sensitization in *Aplysia:* possible role of serotonin and cyclinc amp. *Science* 194:1178-1181.

Byrne, J.H. and Kandel, E.R. (1996) Presynaptic facilitation revisited: State and time dependence. *J. Neurosci.* 16:425-435.

Carew, T.J. and Sahley, C.L. (1986) Invertebrate learning and memory: From behavior to molecules. *Annu. Rev. Neurosci.* 9:435-487.

Castellucci, V., and Kandel, E. (1976) Presynaptic facilitation as a mechanism for behavioral sensitization in *Aplysia. Science* 194:1176-1178.

Castellucci, V.F., Frost, W.N., Goelet, P., Montarolo, P.G., Schacher, S., Morgan, J.A., Blumenfeld, H., and Kandel, E.R. (1986) Cell and molecular analysis of long-term sensitization *in Aplysia. J. Physiol.* 81(4):349-357.

Copplestone, F. (1977) *A History of Philosophy:* ix. *Modern Philosophy.* New York: Doubleday.

Corkin, S. (1965) Tactually-guided maze-learning in man: effects of unilateral cortical excisions and bilateral hippocampal lesions. *Neuropsychologia 3:339-35* 1.

Ebbinghaus, H. (1913) *Memory: A contribution to Experimental Psychology.* New York: Columbia University Teachers College.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., and Taub, E. (1995) Increased cortical representation of the fingers of the left hand in string players. *Science* 270:305-307.

English, J.D. and Sweatt, J.D. (1997) A requirement for the mitogen-activated protein kinase cascade in hippocampal long-term potentiation. *J. Biol. Chem.* 272:19103-19106.

Ferrier, D. (1890*) The Croonian Lectures on Cerebral Localisation.* London: Smith, Elder.

Flexner, L.B., Flexner, J.B., De La Haba, G., and Roberts, R.B. (1965) Loss of memory as related to inhibition of cerebral protein synthesis. *J. Neurochem.* 12:535-541.

Freeman, F.M., Rose, S.P., and Scholey, A.B. (1995) Two time windows of anisomycin- induced amnesia for passive avoidance training in the day-old chick. *Neurobiol. Learn. Mem.* 63:291-295.

Frey, U. and Morris, R.G. (1997) Synaptic tagging and long-term potentiation. *Nature* 385:533-536.

Gaffan, D. (1998) Idiothetic input into object-place configuration as the contribution to memory of the monkey and human hippocampus: A review. *Exp. Brain Res.* 123:201-209.

Grotewiel, M.S., Beck, C.D., Wu, K.H., Zhu, X.R., and Davis, R.L. (1998) Integrin-mediated short-term memory in *Drosophila. Nature* 391:455-456.

Gu, Q.A., Bear, M.F., and Singer, W. (1989) Blockade of nmda-receptors prevents ocularity changes in kitten visual cortex after reversed monocular deprivation. *Brain Res. Dev. Brain Res.* 47:281-288.

Hawkins, R.D., Kandel, E.R., and Siegelbaum, S.A. (1993) Learning to modulate transmitter release: Themes and variations in synaptic plasticity. *Annu. Rev. Neurosci.* 16:625-665.

Hebb, D.O. (1949) *The Organization of Behavior*. Wiley: New York

Huang, Y.Y., Kandel, E.R., Varshavsky, L., Brandon, E.P., Qi, M., Idzerda, R.L., McKnight, G.S., and Bourtchouladze, R. (1995) A genetic test of the effects of mutations in pka on mossy fiber ltp and its relation to spatial and contextual learning. *Cell* 83:1211-1222.

Huang, Y.-Y., Li, X.-C., and Kandel, E.R. (1994) camp contributes to mossy fiber ltp by initiating both a covalently mediated early phase and macromolecular synthesis- dependent late phase. *Cell* 79:69-79.

Hummler, E., Cole, T.J., Blendy, J.A., Ganss, R., Aguzzi, A., Schmid, W., Beermann, F., and Schutz, G. (1994) Targeted mutation of the creb gene: compensation within the crebiatf family of transcription factors. *Proc. Natl. Acad. Sci. USA*, 91:5647-5651.

Impey, S., Obrietan, K., Wong, S.T., Poser, S., Yano, S., Wayman, G., Deloulme, J.C., Chan, G., and Storm, D.R. (1998a) Cross talk between erk and pka is required for Ca?+ stimulation of creb-dependent transcription and erk nuclear translocation. *Neuron* 21:869-883.

Impey, S., Smith, D.M., Obrietan, K., Donahue, R., Wade, C., and Storm, D.R. (1998b) Stimulation of camp response element (cre)-mediated transcription during contextual learning. *Nat. Neurosci.* 1:595-601.

Jackson, J.H. (1884) The Croonian Lectures on Evolution and Dissolution of the Nervous System. *Br. Med. J.* 1:591-593; 660-663; 703-707.

Jenkins, W.M., Merzenich, M.M., Ochs, M.T., Allard, T., and Guic-Robles, E. (1990) Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation. *J. Neurophysiol.* 63:82-104.

Kaas, J.H., Merzenich, M.M., and Killackey, H.P. (1983) The reorganization of somatosensory cortex following peripheral nerve damage in adult and developing mammals. *Annu. Rev. Neurosci.* 6:325-656.

Kandel, E.R. (1976) *Cellular Basis of Behavior: An Introduction to Behavioral Neurobiology*. San Francisco: Freeman and Company

Kandel, E.R. and Pittenger, C. (1999) The past, the future and the biology of memory storage. *Phil. Trans. R. Lond. B.* 354:2027-2052.

Kandel, E.R., Schwartz, J.H., and Jessell, T.M. (2000) *Principles of Neural Science, Edition.* New York: McGraw-Hill.

Kentros, C., Hargreaves, E., Hawkins, R.D., Kandel, E.R., Shapiro, M., and Muller, R.V. (1998) Abolition of long-term stability of new hippocampal place cell maps by nmda receptor blockade. *Science* 280:2121-2126.

Lashley, K.S. (1929) *Brain Mechanisms and Intelligence: A Quantitative Study of Injuries to the Brain.* Chicago: Chicago University Press.

Lin, X.-Y. and Glanzman, D.L. (1994). Hebbian induction of long-term potentiation of *Aplysia* sensorimotor synapses: Partial requirement for activation of an nmda-related receptor. *Proc. R. Soc. Lond. B. Biol. Sci.* 255:215-221.

Maletic-Savatic, M., Malinow, R., and Svoboda, K. (1999) Rapid dendritic morphogenesis in cai hippocampal dendrites induced by synaptic activity. *Science* 283:1923-1927.

Martin, K.C., Casadio, A., Zhu, H., E, Y., Rose, J.C., Chen, M., Bailey, C.H., and Kandel, E.R. (1997a) Synapse-specific, long-term facilitation of *Aplysia* sensory to motor synapses: a function for local protein synthesis in memory storage. *Cell* 91:927-938.

Martin, K.C., Michael, D., Rose, J.C., Barad, M., Casadio, A., Zhu, H., and Kandel, E.R. (1997b) map kinase translocates into the nucleus of the presynaptic cell and is required for long-term facilitation in *Aplysia. Neuron* 18:899-912.

McHugh, T.J., Blum, K.I., Tsien, J.Z., Tonegawa, S., and Wilson, M.A. (1996) Impaired hippocampal representation of space in cai-specific ~ari knockout mice. *Cell* 87:1339-1349.

Merzenich, M.M., Nelson, R.J., Stryker, M.P., Cynader, M.S., Schoppmann, A., and Zook, J.M. (1984) Somatosensory cortical map changes following digit amputation in adult monkeys. *J. Comp. Neurol.* 224:591-605.

Milner, B., Squire, L.R., and Kandel, E.R. (1998) Cognitive neuroscience and the study of memory. *Neuron* 20:445-468.

Montarolo, P.G., Goelet, P., Castellucci, V.F., Morgan, J., Kandel, E.R., and Schacher, S. (1986) A critical period for macromolecular synthesis in long-term heterosynaptic facilitation in *Aplysia. Science* 234:1249-1254.

Muller, R.U. and Kubie, J.L. (1987) The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *J. Neurosci.* 7:1951-1968.

Muller, R.U., Kubie, J.L., and Ranck, J.B. Jr (1987) Spatial firing patterns of hippocampal complex-spike cells in a fixed environment. *J. Neurosci.* 7:1935-1950.

Nicoll, R.A. and Malenka, R.C. (1999) Expression mechanisms underlying ~A receptor- dependent long-term potentiation. *Ann. N. Y. Acad. Sci.* 868:515-525.

O'Keefe, J. and Nadel, L. (1978) *The Hippocampus As A Cognitive Map.* Clarendon: Oxford.

Penfield, W. and Perot, P. (1963) The brain's record of auditory and visual experience. *Brain* 86:595-696.

Pham, T.A., Impey, S., Storm, D.R., and Stryker, M.P. (1999) cre-mediated gene transcription in neocortical neuronal plasticity during the developmental critical period. *Neuron 1999* 22:63-72.

Pittenger, C. and Kandel, E. (1998) A genetic switch for long-term memory. *C. R. Acad. Sci.* 321:91-96.

Ramon y Cajal, S. (1893) Neue Darstellung vom Histologischen Bau des Centralnervensystem. *Arch. Anat. Entwick.,* 319-428

Rolls, E.T. and Treves, A. (1998) *Neural Networks and Brain Function.* Oxford: New York.

Rotenberg, A., Mayford, M., Hawkins, R.D., Kandel, E.R., and Muller, R.U. (1996) Mice expressing activated Camkii lack low frequency ltp and do not form stable place cells in the cai region of the hippocampus. *Cell* 1996 87:1351-1361.

Rudolph, D., Tafuri, A., Gass, P., Hammerling, G.J., Arnold, B., and Schutz, G. (1998) Impaired fetal T cell development and perinatal lethality in mice lacking the camp response element binding protein. *Proc. Natl. Acad. Sci. USA,* 95:4481-4486.

Ryle, G. (1949) *The Concept of Mind.* Chicago: University of Chicago Press

Schacter, D.L. (1996) *Searching for Memory: The Brain, The Mind, and The Past.* Basic Books: New York.

Schuster, C.M., Davis, G.W., Fetter, R.D., and Goodman, C.S. (1996) Genetic dissection of structural and functional components of synaptic plasticity. ii. Fasciclin ii controls presynaptic structural plasticity. *Neuron* 17:655-667.

Scoville, W.B. and Milner, B. (1957) Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry* 20:11-21.

Skoulakis, E.M. and Davis, R.L. (1996) Olfactory learning deficits in mutants for leonardo, a *Drosophila* gene encoding a 14-3-3 protein. *Neuron* 17:931-944.

Sorra, K.E. and Harris, K.M. (1 998) Stability in synapse number and size at 2 hr after long-term potentiation in hippocampal area cai. *J. Neurosci.* 18:658-671.

Squire, L. and Kandel, E.R. (1999) *Memory: From Mind to Molecules.* New York: Scientific American Books.

Squire, L.R. and Zola-Morgan, M. (1991) The medial temporal lobe memory system. *Science* 253:1380-1386.

Weiner, J. (1999) *Time, Love, Memory: A Great Biologist and His Quest for the Origins of Behavior.* New York: Knopf.

Wernicke, K. (1908) The symptom-complex of aphasia. In A. Church (ed.), *Diseases of the Nervous System.* New York: Appleton, pp. 265-324.

Yin, J.C., Del Vecchio, M., Zhou, H., and Tully, T. (1995) creb as a memory modulator: induced expression of a dcreb2 activator isoform enhances long-term memory in *Drosophila. Cell* 81:107-115.

Yin, J.C., Wallach, J.S., Del Vecchio, M., Wilder, E.L., Zhou, H., Quinn, W.G., and Tully, T. (1994) Induction of a dominant negative creb transgene specifically blocks long-term memory in Drosophila. *Cell* 79:49-58.

Zamanillo, D., Sprengel, R., Hvalby, O., Jensen, V., Burnashev, N., Rozov, A., Kaisser, K.M.M., Köster, H.J., Borchardt, T., Worley, P., Lübke, J., Frotscher, M., Kelly, P.H., Sommer, B., Andersen, P., Seeburg, P.H., and Sakmann, B. (1999) Importance of ampa receptors for hippocampal synaptic plasticity but not for spatial learning. *Science* 284:1805-1811.

Zeki, S. (1993) 4 *Vision of the Brain.* Oxford: Boston.

Zola-Morgan, S., Squire, L.R., and Amaral, D.G. (1986) Human amnesia and the medial temporal region: enduring memory impairment following a bilateral lesion limited to field cai of the hippocampus. *J. Neurosci.* 6:2950-2967.

Zubin, J. and Barrera, S.E. (1941) Effect of electric convulsive therapy on memory. *Proceedings of the Society of Experimental Biology and Medicine* 48:596-597.

Poul Harremoës

Scientific incertitude in environmental analysis and decision making

Professor Poul Harremoës received the Dr A.H. Heineken Prize for Environmental Sciences 2000 for his contributions to the theory of biofilm kinetics in relation to biological waste water treatment and for his successful organisation of the international scientific community in water pollution research and control.

Introduction

During the last century there have been too many 'surprises', situations where detrimental impacts on the environment occurred due to disregarded knowledge or ignorance with respect to the cause effect relationship between the pressures on the environment and consequences to society. It is essential to distinguish between the different levels of incertitude: Determinism, risk, uncertainty, ignorance and indeterminacy. There is need for a change of paradigm from an elitist, narrow approach to integrated environmental assessment and risk assessment. The incertitude has to be accounted for in order to prevent surprises. In case of recognised incertitude, solutions have to be flexible and robust, especially in situations involving irreversibility of the consequences of the decision. Participatory approaches should have a larger role in the procedural and institutional aspects of framing and implementation of the choices for environmental abatement.

The 'DPSIR' approach

In The European Environmental Agency, the problems related to the environment are fitted to a formal structure called the dpsir-approach to integrated environmental assessment, European Environment Agency (1998). The structure is a formalisation of the interplay between the society and the environment with respect to *Driving forces*, by which *Pressures* are exerted on the environment. The changes to *State* of the environment is the basis for evaluating the *Impacts* due to the Pressures. That leads to *Responses* by society to the impacts, which requires identification of objectives and choice of tools by which to curb the pressures. From pressures to impacts is the cause-effect relationship, based on physical, chemical and biological mechanisms, describing the relationship between pressures and impacts. This environmental cause-effect relationship involves the natural sciences, while the driving forces and the responses involve the political, judicial and the social sciences. The role of social sciences is as important as the role of the natural sciences – a fact, frequently ignored, because there is a general favour for natural sciences as being the exact, factual science (frequently an incorrect assumption), as opposed to the more ambiguous social sciences.

Alternative approaches

There are two approaches to design and operation of technology:

– The empirical iterative approach, also called 'trial and error'.
– The deterministic predictive approach, also called 'design'.

Remarkable engineering accomplishments have been achieved by the empirical iterative approach: Roman aqueducts still standing, drainage in ancient Athens, middle age gothic cathedrals and many other famous structures. Many mistakes paved the road to these successes: the tower in Pisa is still (just) standing. The bridge over the Sct. Lawrence River in Quebec was designed by scaling up the bridge in Scotland across the river Forth, but columns cannot be scaled up proportionally and the bridge fell down! The empirical iterative approach is still a valid approach. Those favouring this approach tend to take pride in being called: 'practitioners'.

In the other end of the scale is the deterministic predictive approach, which favours to develop an understanding of all elements in the structure, so that the performance of the structure can be predicted. On that basis the structure can be designed to meet predetermined requirements. The enthusiasts of this approach will be called 'theoretists'.

There has always been a schism between the practitioners and the theoretists, frequently in relation to value judgement on contributions to international conferences, at which it is difficult to satisfy both. The fact is that there is little justification for this schism, because there is need for both, e.g. expressed by the joke: 'There is nothing more practical than a good theory'.

The schism between the two approaches can also be illustrated by the development of engineering education, which has become still more theoretical – leaving the student to achieve know-how from practise after graduation. It is a valid question to pose, whether the tendency to educate by exemplified application of models is a good approach. Do the students get a basic understanding of theory and/or practise, or do they get entangled in the mechanisms of futile handling of hard- and software, which will be obsolete before the students leave the university?

Determinism

Determinism is based on a reductionist philosophy. It is based on the concept that a full understanding can be achieved by identification and description of all underlying physical, chemical and biological laws of nature that govern the engineering application. On reflection, there is reason the question the validity of this basic axiom.

We have many indications from practise that this is a reasonable axiom, but philosophically we have counter-indications that the validity is limited:

– Moving from macrophysics to microphysics, the Bohr-model of the atom is not deterministic and Heisenberg's uncertainty theorem quantifies the limitation.

– Theory of 'chaos' showed us that there might not be unique solutions to even simple differential equations. In fact, at university we are taught only the soluble problems, which turns out to be a small fraction of the universe of problems.
– Neural networks illustrate a totally different set of learning and 'understanding': pattern recognition – the holistic approach. Are we back to the empirical-iterative approach?

The university graduated engineer go into practise with an exaggerated faith in determinism, while the practitioners, who – after years in practise – have forgotten the theoretical background, defensively show disrespect to theoretical approaches. The reality is somewhere in-between – but where?

The deterministic model

The deterministic model is a mathematical description of the physical, chemical and biological phenomena involved in the engineering problem of concern. They can be very simple and they can be very complicated. However, it must be emphasised as a philosophical fact that no model is anything but an approximation to the reality. It may be a good approximation or a poor one, but always an approximation.

In the engineering application, the issue is not to develop an ever 'better' description. That is frequently the model developers apparent goal. The issue is to identify and develop the model that fits the engineering problem to which it shall be applied. There is no one solution to this. To a multitude of engineering problems fit a multitude of models. Flexibility is a virtue.

From the last twenty years of model development, one gets the impression that 'better' models are interpreted as more detailed models. The more that can be described, the better. That is not necessarily the case, as will be discussed in the following. Parsimony is a principle introduced by Ockham (Eng. philosopher, 1290-1349). Parsimony means that the best approach is the simplest that fits the purpose of the application. Again, flexibility of model features is a virtue that fits with parsimony.

A good example of the need for parsimony is model reduction for application in real time control. Due to the limited time available for reaching a solution within a time step, the very complicated models are too time consuming for practical application. In a few years time, faster computers might overcome that. However, the complicated models may be difficult to keep stable. It may be difficult to analyse the reason for an error, because the model complexity complicates the understanding of performance.

The key point is that an engineering model has to model the essential features that are important to the resulting design or operation (Harremoës and Madsen, 1999). All other details just obscure the picture and hamper engineering application. The slide-rule engineers of the past had to simplify down to the bare essentials, a few combinations of key parameters, and to incorporate the rest in safety factors. That is (was?) an engineering art. Models and computers have just shifted the level of comprehension; but the art of choosing a perceivable and manageable structure and a corresponding set of parameters is the same at another level. Is the art about to be forgotten?

Determinism end incertitude

How predictable is the performance of the technologies associated with environmental abatement? How well can scientific understanding and/or experience provide sufficient assurance of reliability.

In the development of natural sciences since Galilæi, Kepler, Newton, etc., the deterministic description of cause-effect relationships has been the core of development. It is axiomatically assumed that there is a unique relationship between the action taken and the effects (in the environment or on human health). However, during the last century it has been realised that there are both inherited and practical uncertainties associated with this relationship and more recently questions have been raised as to whether there is such an identifiable, unique relationship at all. The precautionary principle is in fact a response to that realisation.

The cause-effect relationship can be written as follows:

$e = f(i_1,i_2,\ldots i_n; p_1, p_2,\ldots.p_n) + e$

Where:	e	is the effect
	f	is a functional relationship
	i	are input variables
	p	are parameters
	e	is incertitude

The functional relationship may be empirical in the form of correlations or theoretical in the form of generic relationships, based on a-priori knowledge of the phenomena involved. In relation to environmental cause-effect relationships the phenomena are physical, chemical and biological.

The input consists of the variables associated with natural phenomena and the anthropocentric pressures on the environment. The parameters characterise the functional relationship. In some cases such parameters are very well known from

a-priori scientific knowledge, in many cases they have to be determined in each individual case by experiments and analysis of the phenomena involved.

The incertitude describes the extent to which it has not been possible to predict the effect on the basis of a deterministic functional relationship. That includes both the statistical uncertainty in cases where sufficient experimental data provide information by which to estimate the statistical error, but it also include the incertitude associated with not knowing the essential phenomena and the lack of data by which to estimate the incertitude.

This has also been expressed by the following words: *Determinism, Risk, Uncertainty, Ignorance, Indeterminacy* (Wynne, 1992).

Determinism is an ideal that is never achieved. However, history has demonstrated beyond any doubt that determinism is worth striving for. Even with the best of determinism, the input variables show statistical properties. However, we have statistical instruments with which to handle variable input data and *risk* expressed statistically is a rational approach to description of variation. There is always statistical uncertainty involved, due to the mere fact that all relationships have to be calibrated with data. With a well known functional relationship and an adequate combination of number of parameters plus number and character of data, the uncertainty can be expressed statistically and can be incorporated in the risk analysis.

Uncertainty beyond the statistical uncertainty is experienced when we know the range of outcome, but not the statistics of it. That is where we can use scenarios as an approach to analysis, because we can describe a set of outcomes to be expected, but we cannot associate probabilities.

Ignorance applies when we do not know essential functional relationships. They may become known later due to research and development on the issue, but they may not be known at the time, when far-reaching decisions have to be made.

Practical indeterminacy is the situation where the functional relationships are so complicated and the number of parameters so large that neither determinism nor stochasticity is within reach. The functions and the parameters become unidentifiable. *Theoretical indeterminacy* is the situation where the relationships are inherently unidentifiable, e.g. due to chaotic properties that make predictions impossible.

The precautionary principle has been introduced as a way in which to formu-

late an approach to the situations where ignorance and indeterminacy dominate the cause-effect relationships. The importance is due to the realisation that many of the cause-effect relationships between the pressures from development of society and the environmental impacts are less than well known.

The burden of proof

It is a frequent interpretation of the precautionary principle that the burden of proof shall be shifted from the environmental stakeholder to the polluter. It is not that simple (Harremoës, 1998).

Imagine a situation associated with risk assessment of a chemical, interpreted according to Popper (1965).

Imagine the claim that the chemical does cause harmful effects in the environment or to human health. If those adverse effects are well defined, the burden of proof can in fact be lifted. One example will prove the claim. It may be quite com plicated to do so, but in principle it is feasible.

The opposing claim that the chemical does not cause harmful effects in the environment or to human health, cannot be lifted. There will always be situations, which have not been covered with adequate investigations. The claim can only be evaluated by induction. The better and the more comprehensive the investigation, the more likely the claim is; but there can never be certainty. Just one example would falsify the claim.

The conclusion is that environmental and human health acceptance can only be demonstrated by induction and only by likelihood. In the end, it becomes a question of confidence. That confidence is based on procedure: Have the authorities or the polluter performed adequate investigations in order to establish a reasonable likelihood that the chemical is harmless according to acceptable criteria?

One interpretation of the precautionary principle is that it is the polluter who shall demonstrate by induction based on acceptable procedures that the chemical is acceptable according to agreed criteria.

Pro et con

Where adequate information on cause-effect relationships and acceptable effects is available based on determinism and statistical uncertainty analysis, the political choice becomes a balancing act between opposing concerns: the benefit to society versus the acceptance of a level of effects in the environment and/or effects to human health.

It is frequently postulated that the public does not have a rational approach to

the acceptable levels of effects, because it is a well known fact that people accept harmful effects with probabilities orders of magnitude apart (e.g. traffic, alcohol). However, people have a much better perception of the pro et con than recognised by the specialists in risk assessment. It is when adequate information on the pro et con is unknown to the public that the reaction is fear and unrealistic wishes for action. Risk assessment is performed by an elitist profession. Public participation and stakeholder information has to be expanded and improved in quality, in order to improve the confidence in the system.

In relation to the 'pro et con' in evaluation of risk assessment, the precautionary principle is an expressed concern regarding uncertainties in favour of the environment and human health.

The risk of being wrong

In the weighting of concerns, the precautionary principle can be interpreted as an urge to evaluate the risk of being wrong. What are the consequences of not doing anything, in case the harmful effects turn out to be severe? The occupational health problems associated with asbestos, is a striking historical example. In relation to environment, the problem with mercury is a similar example. What are the consequences of an action, in case the action turns out to be not necessary or non-proportional?

Many issues fall into this category of concern for being wrong. At present, in Denmark, the hottest issue is the regulation, ultimately ban, of pesticides in agriculture. The very same issues will be faced with the introduction of water reuse – plus many more.

Example: MTBE

The example chosen to illustrate the problem is the introduction of mtbe as a substitute to the use of methyl-lead as an anti-knocking constituent in gasoline. The toxicity of lead and methyl-lead has been known for a long time, but it was not until the early 1980´s that this fact got recognition to the extent that an alternative was introduced. The choice as a substitute was mtbe, which combine a number of advantageous properties, including also to serve as an oxygenate for the reduction of air pollution. In usa the addition of mtbe was initiated in the late 1980´s. In eu a policy on gasoline composition was adopted in 1997, with the consequence that mtbe could be the massively produced compound that reduce the air pollution from car engines. However, already since 1990 papers were published in the scientific community that mtbe might be a serious groundwater contaminant. That became obvious in 1997 when severe groundwater pollution

was detected in Santa Monica, California, from leaking gasoline stations. Similar incidents during the late 1990´s have been revealed in Denmark. In March 2000, a ban of the use of mtbe was announced in usa.

This story calls for analysis of the story behind, in an attempt to learn from a development that could have been avoided, Krayer von Krauss and Harremoes, 2000. Looking into the scientific background it can be revealed that the detrimental properties of mtbe in groundwater was known in the 1950´s and could be found in the textbooks in the 1960´s. mtbe has a high solubility in water. It does not adsorb to the soil structure. It does not degrade in groundwater. It gives an unpleasant taste and odour to water at low concentrations. Leaking tanks have made water supply wells unusable. eu is now under pressure to ban mtbe. The whole traffic paradigm is under pressure to change approach, because the traditional engine designs are lacking alternative that satisfy modern air and water quality standards. There is room for a lot of innovations on the path to solutions in the traffic sector, which has not yet been environmentally decoupled from economic growth.

This is a case of 'reducible ignorance'. The information was there, but it was either unknown to the decision makers or ignored. In both cases, there is reason to wonder how a chemical with such known properties can be made one of the greatest mass produced compounds in society.

The European Environmental Agency (2000) has launched a project looking into the history of such surprises, derived from ignorance or lack of attention to existing knowledge in environmental decision making. The list of cases examined is:

– Antibiotic as animal growth promoters
– Hormones as animal growth promoters
– bse, Mad Cow Disease
– Hormones in fertility treatment
– Spectrum of persistent pollutants
– Asbestos
– SO_2 and acid rain
– cfcs and the ozone layer
– pcbs
– Benzene
– mtbe as a substitute anti-knocking agent
– Radiation
– Marine overfishing

While mtbe is a case of ignoring/disregarding existing knowledge, cfc and pcb is a case of shear ignorance. At introduction, the concept was that these compounds were ideal due to being inert and apparently without detrimental effects.

The no-know situation

The real importance of the precautionary principle is in the situation where either ignorance or indeterminacy prevails. However, still it is quite unexpected that actions should be taken in situations of no knowledge of a cause-effect relationship at all. There has to be some reason on which to argue a reasonable suspicion, based on an expected cause-effect relationship – however uncertain. Otherwise the arguments change from a rational utilitarian concept to deontological dogmas or just superstition or non-structured empiricism.
This situation is in fact not unfamiliar from engineering approaches. This can be looked at from two extreme points of view, described in the introduction:

– The empirical, iterative approach
– The predictive, scientific approach

The empirical, iterative approach was the way in which craftsmanship developed before modern science provided much better means for design. That empiricism did in fact thrive is visible from antic and middle-age structures, e.g. cathedrals. However, it was the Greek heritage of logic and scientific development that created the basis for modern technology after the renaissance. This is in particular visible by comparison between two cultural developments in parallel: the European development on the basis of science versus the Chinese development on the basis of an empirical correlation philosophy and technology, which worked extremely well by comparison until the scientific break-through in Europe. In fact, the Chinese had cast iron, pumps, suspension bridges, china, etc. long before Europe (Temple, 1998).

In case of ignorance and indeterminacy as described above, the precautionary principle comes in as an approach, which combines with an empirical, iterative approach, where a predictive, scientific approach fails or suffers from a predominance of incertitude. In relation to risk assessment, the chemical properties to be concerned with are:

– Persistency in the environment
– Bio-accumulation in the environment
– Severity of toxicity to humans
– Irreversibility of consequences

The greater the suspicion of harmful effects to the environment and/or to human

health is, the more pre-emptive measures are called for. In view of the level of ignorance and indeterminacy, it is important to make development in the right direction in a stepwise fashion, because the risk of being wrong must be counter balanced by the advantages of a new technology or chemical. At the same time, it is urgent to increase monitoring and research on the issues of concern. It must be in everybody's interest to improve, where possible, the knowledge of the cause-effect relationship (remove the uncertainty and ignorance) and to decrease the uncertainty (provision of more and better data and establishment of safe procedures for implementation) associated with design and operation of technologies.

There has to be a much greater attention to irreversibility. In such cases, the primary demand is for thorough investigation of existing knowledge, investigation of potential consequences, analysis of alternatives, assessment of risk perception and intensified monitoring of the environment and initiation of research on identified suspicion. The precautionary approach is not a rationale by which to tighten the regulatory screw on well known issues. It is an integrated attempt to avoid scientific surprises in the future.

Participatory versus modelling approaches

The approach of the recent decades has been to rely on the knowledge on the cause-effect relationship as the means for predicting results of optional policies and to ignore the uncertainty/ignorance associated. The models have become more and more complex and less and less interpretable to the non-expert. This in combination with surprises/mistakes of the past has increased the mistrust on the part of the politicians and the public. The way forward is not to curtail development in fear of new failures, but to involve all stakeholders in the process of formulating, assessing and choosing policies, Stirling (1999). Quotation from Stockholm Water Prize laureate year 2000 (siwi, 2000): Kader Asmal, Minister of South Africa: 'When moving forward we shall make mistakes, but it is wise to do so with open eyes and with a democratic consensus behind us'. Do present elitist science/engineering procedures for integrated environmental assessment and risk assessment live up to such ideals?

New paradigm

In the context of scientific development there is a need for change of paradigm. According to Kyhn (1970) science moves in stages. A fundamentally new concept/approach, called a new paradigm revolutionise science at certain intervals. In the interim period 'normal science' works to improve knowledge within the frame of the prevailing paradigm. Slowly, the fundamental concept is threatened

by mismatches to reality. A new paradigm emerges as a fundamentally new approach to the framing, structuring, analysis and interpretation to the craftmanship of science – much to the dislike of the establishment. The question is whether we are approaching an new paradigm, where incertitude (uncertainty, ignorance and indeterminacy) become an accepted fact on both side of the borderline between the scientists on the one side and the public and the politicians on the other side. That will require a change of attitude on both sides: The politicians have to accept that fuzzy answers may be the best expression of expertise. The scientists have to learn that identification of the fuzzy borderline between knowledge and ignorance may be the sign of real competence.

Literature

Harremoës, P. (1998): Can risk analysis be applied in connection with the precautionary principle? In: *The Precautionary Principle. Extracts and summary from the Danish Environmental Protection Agency's Conference on the precautionary principle, Eigtveds Pakhus, Copenhagen 29 May*, pp. 30-35. Danish Environmental Protection Agency, Copenhagen. (Environmental News 35).

Harremoës, P. & Madsen, H. (1999): Fiction and reality in the modelling world Balance between simplicity and complexity, calibration and identifiability, verification and falsification. *Water Science and Technology*, 39, (9), 1-8.

Kyhn T.S. (1970): *The Structure of Scientific Revolutions*, 2nd. Chicago, The University of Chicago Press.

European Environment Agency (1998): *Europe's Environment: The Second Assessment*, 293 pp.

Krayer von Krauss and Harremoës (2001): *European Environment Agency: Late Lessons from Early Warnings: the precautionary principle 1896-2000.*

Popper K.R.(1965): *Conjectures and Refudiations*, 2nd. Ed. London, Routledge and Kegan Paul.

siwi (2000): *Our Struggle for Water*, ed. P. Hanneberg, Stockholm International Water Institute. isbn 91-973359-4-0.

Stirling, A. (1999): *On 'Precautionary' and 'Science Based' Approaches to Risk Assessment and Ennvironmental Appraisal*, spru, University of Sussex, 15 pp.

Temple, R. (1998) *The Genius of China – 3000 years of Science, Discovery and Invention*, Simon & Schuster Inc.

Wynne, B. (1992): *Uncertainty and Environmental Learning, Global Environmental Change*, Butterworth-Heinemann Ltd. pp. 111-127.

Dr A.H. Heineken Prize for History

Jan de Vries

Towards a history that counts

The Dr A.H. Heineken Prize for History 2000 was awarded to professor Jan de Vries for his pioneering research into the development of the European economy between 1500 and 1800.

Permit me to preface my lecture with some reflection on the academic discipline of history, for services to which the Royal Academy jury awards the Heineken prize. Having so recently accepted this prize, I can hardly now deny that I am a member of this academic tribe. But the sort I do is, of course, economic history. It strives to be an interdisciplinary field of study, and while some such fields seek the interstices of established disciplines in order to escape from *all* discipline, mine expressly subjects itself to both – both history an economics – with the aspiration of contributing to not one but two disciplines.

And, they are very different disciplines. Karl Popper 40 years ago pronounced economics the first of the social sciences to have 'gone through its Newtonian revolution'.[1] We might imagine these evolved economists as the first amphibians among the social scientist fish, laboriously learning to use their stubby little fins to scramble upon the shore. By now they have moved well inland into the Newtonian forest.

Meanwhile, history as a discipline has moved further out to sea. The prospect of evolving as 'scientists'[2] lost its appeal to the historian fish. Instead, historians in the past two decades have shed several of their functional limbs to become rather simpler organism, organisms that attract attention increasingly by their bright colour, and exotic and diverse forms.

Perhaps I can now set this conceit aside to observe simply that the gap between the deductive and the inductive discipline, between the nomothetic (predictive, law-giving) economics and the idiographic (unique, descriptive) history is now much wider than it was a generation ago, placing us economic historians before some difficult choices.

Some responded by developing a more social scientific history. In the 1960s there was a sharp increase in interest, certainly within the American historical profession, in the theories and methodologies of the social sciences. This new interest was motivated primarily by a desire for a more objective historical scholarship, one with standards sufficiently rigorous to withstand ideological manipulation. This interest was widely shared among historians, but the search for rigor was pushed furthest in economic history. 'There,' as Naomi Lamoreoux observed, 'a small group of economists launched a veritable revolution, seizing control of the discipline's organizations and using them to build a coherent and uniquely American body of scholarships based on the study of the past.'[3] This New Economic History became almost an applied field of economics, using neoclassical theory to test hypotheses concerning historical questions. It had, and continues to have, success in addressing issues of market performance and of rational behaviour, and its influence has spread far beyond North America. But its re-

liance on an a historical theory imposes obvious limitations, and it has had little to say concerning the larger questions of long-term change and development that long had been the distinguishing objective of the field. This New Economic History formed the intellectual climate in which I entered the discipline. I have struggled with its strengths and its limitations for many years now.

Others responded differently to the challenge of interdisciplinarity, by advancing the audacious claim that rather than making history social scientific, social science should be made historical. To those seized with the vision, the first step was to renovate history as a discipline, to allow it to take its place as the 'queen of the social sciences.' This was the stated objective of the French Annalistes.[4] The most influential of them, Fernand Braudel, argued that to become the unifying center of all the social sciences, the historical profession had to overcome its continued allegiance to a 'pernicious humanism' (what he later describes as a 'militant anthropocentrism').'[5] Furnished with a 'clear awareness of the plurality of social time,' and capable of incorporating the dialectics of both time and space, history – and history alone – could reconstitute the global nature of human phenomena.

This spacious conception of the historian's enterprise has always attracted me, even though the exponents of historical social science more often led by inspiration than by practical example. Not being myself of a philosophical frame of mind, I have applied myself to learning and teaching by doing, attempting to find practical ways to combine theory and history. But, as the two disciplines moved further apart, my discomfort grew. Indeed, with a foot in both camps, as it were, my posture must have appeared increasingly undignified, and this knowledge only heightened my surprise when Professor Reneman called with the news that I had been chosen to receive the Heineken Prize in history.

I do not want to let this opportunity pass to state my continuing conviction that to save history from its sometimes brilliant irrelevancy and the social sciences from their often sophisticated superficiality requires a history that counts; that is, a history with which we can take accounts of our pasts with both the rigor to understand and the scope to comprehend. Such a history needs to incorporate the multiple layers of duration and their interactions that are at the heart of social reality and to introduce them to the variously static, linear and a historical models that have for so long dominated the social sciences.

Today, history is as far as ever from becoming 'the queen of the social sciences', but this does not mean that the aspirations for a reformed history have been abandoned. There are within the social sciences themselves serious movements aspiring to develop theories in which duration and space – history and geography

– are central dimensions. The general climate of thought, which has scattered historical research to the four winds, has set at least some social scientists to exploring non-linear dynamics and the so-called 'sciences of complexity'. Economists in particular, are now – some of them – leaving the Newtonian forest fort the biological highlands, exploring the implications of bounded and myopic rationality, increasing returns to scale and 'spillover effects' – all of which accord new significance to the details of historical sequence and spatial relationships.

In economics, these new developments challenge from within the neoclassical tendency to abstract from time and space: to assume that however important events and institutions might appear to be in the short run, their influence is erased by equilibrium-seeking processes, making them irrelevant to ultimate outcomes. Te erosion of this confidence has led to the development of a 'new institutional economics' (including the new organizational and informational economics)[6], the new growth theory[7], and the concept of 'path dependence'[8], all of which seek to convince the economist that, in some sense, 'history matters'.

Is there a comparable impulse within the historical profession to renew the dialogue with the social sciences? In the United States the most notable initiatives appear to come primarily from non-Western historiography, where macrohistorical and comparative questions continue to motivate research. But in the historical 'heartlands' of western history the belief that knowledge is cumulative is so little acknowledged and, consequently, the impulse to empirical refinement is so weakened, that there is little motivation to maintain a broadly-based discourse, and this leaves historians isolated, talking only to their closest soul-mates. At any rate, the leaders in the construction of a historical social science are now to be found within the social science disciplines themselves.

In my own work I have often found myself working a margin between the prosaic, everyday market behaviour of ordinary people (using simple concepts drawn from economics and statistics), and the large, patterned and structured webs that channel and constrain such behaviour (using classical theory, geography, and historical models that incorporate durational plurality). I have worked this margin to explore the character of an historical epoch that itself stands at the margins of historical periodization and therefore is full of questions about its true identity: the early modern era of European history. Just what sort of modernity is early modernity? This question stands behind a good deal of my work.

Some historians have a gift for uncovering the greatness of even ordinary people. My effort have been directed more to emphasizing the ordinariness of even great people. It is the ordinary, everyday, repeated behaviour – whether of farmers or landscape painters – that lends itself to quantitative analysis, compa-

rison and modelling. And it is these techniques that allow for the description and assessment of the institutions and market structures that influence those behaviours, causing, for example, some ordinary cultivators to be 'peasants' and others to become 'farmers'.[9]

That was the first historical question I tackled, by searching, as it were, for a tipping mechanism that could lead similarly endowed households to seek and accept opportunities for specialization in agricultural production or to resist such opportunities as either unattractive or too risky.

In the case of painters, the question that first intrigued me was the relationship between quantity an quality. A confirmed humanist might supposes creative artistic gifts to be inherently rare and, at any rate, inexplicable. With this assumption, a rise in the volume of artistic production would be expected to reduce the average level of quality. Yet seventeenth-century Dutch painting appears to offer an example where high average quality was directly and positively correlated to the achievement of an extraordinarily high level of production as measured by both the number of painters and their output.[10]

This could be so because institutions for the training of artists and innovations in new products (types of painting) and new processes (technique of painting) were supported by the volume of production – and would disappear when that volume could no longer be sustained. It appeared that the techniques used to study innovation within a conventional industry offered insights into the dynamics of an important phenomenon in art history.[11] Here the economic historian can go only so far, of course, before enlisting the aid of the art historian, since the enduring value of an insight derived from the theory of industrial organization depends on al full contextualization.

One of the chief tasks of the historian – of whatever stripe – is to study a phenomenon in context – in a full context. In contemplating the present our eyes are inevitably clouded by the smoke of explosive events. Distinguishing those that can endure to shape or channel future actions from those that disappear with yesterday's newspapers requires some means of clearing that smoke. Indeed, it is the possibility of wholistic study that makes historical work of such interest and value to students of contemporary society.[12] Should the historian, then, not make a special effort to establish contexts that aspire to capture the complex social reality? Does this task not require moving beyond the banality of the simple chronological narrative that has recaptured the allegiance of most of the profession, at least in the United States?

The individual face to face with laws, rules, institutions, market forces, etc. is a common starting point in social scientific research, but individuals are themselves

implicated in larger social units and all of these relate to each other in complex ways that form networks and hierarchies of trade, finance, information, redistribution and technology standardizations. The form these networks take – how extensive, how costly to use, how centralized – have powerful and lasting effects. Indeed, here we reach historical phenomena that follow a different, slower, rhythm than individual events. Once a durable form is established, it tends to persist, like putty turning to clay, channelling the frenetic actions of individuals into rigid paths. A history that can examine the evolution of these phenomena must necessarily cover long time spans, and this I have tried to do, for example, in examining the emergence, persistence, and ultimate demise of the Dutch system of *trekschuit* (passenger barge) transportation. This system, emerging from many independent decisions in the seventeenth century, is long gone. Today, only the trained eye can detect its earlier existence. Yet it cast a shadow forward in time, shaping the form of much of the rail network that succeeded it, and determining features of the urban structure of the Netherlands even in the present age.

Later I studied European urban history from this network perspective. Here, too, the whole (the urban system) appeared to be more than the sum of its parts (the cities of Europe), and the whole took a shape recognizable and influential today in an unexpected period. That is, in the seventeenth century, a time of little overall urbanization and without dramatic technological change, a major spatial reorganization, and functional specialization emerged that continued to determine much of the shape of urban Europe throughout the era of industrialization. The most basic characteristics of modern urbanization emerged neither along a linear path nor as a result of the industrial technologies that attract our attention by dint of their novelty. Indeed, judging technologies by their novelty rather than by their economic impact is a common pitfall in both historical and contemporary analysis.

The seventeenth century saw no dramatic change in the potential for urbanization, but an interactive combination of commercial and political development brought about the realization of a new international urban hierarchy.

Historical research of this type, mine and that of many others, has led me to a topic at a yet higher level of abstraction. From microeconomic studies of behaviour and macroeconomic investigations of structures and patterns, I have been led to historiographical issues of periodization. Over the years Heineken prize juries seem to have been attracted to – perhaps, they have had a weakness for – historians whose have proposed to blur, or reset, the venerable lines of demarcation that divide western history into its fields of study. Peter Brown audaciously and brilliantly transgressed the boundaries between late ancient and early Christian

and medieval history, while Heiko Oberman led in the integration of late medieval and early modern history into a single, seamless period of study. Debates over periodization among professional historians may appear as minor or technical matters to the outsider. But, in fact, movement along these frontiers is a bit like change in the networked structures I discussed earlier. They yield only to great and persistent force, but when they do, the historiographical landscape changes in profound ways. New questions can be asked, and old verities questioned.

My demolition efforts have been directed at the high walls buttressed by the French and Industrial Revolutions that separate early modern from late modern (or 'contemporary') history, or, if you will, traditional from modern society. In the case of economic history, this dividing line also affects the theory of modern economic growth, for this economic concept, however a historical is its practical execution, is anchored in a specific historical understanding of the profound difference between th cyclical character of economic life before and the dynamic character after the events we know as the Industrial Revolution. A world held in the grip of Malthus was liberated – Prometius was unbound – in a brief period around 1800. Both the English and Dutch titles of my recent book with Ad van der Woude reaveal a criticism of this conventional historiography. Our study of Dutch economic history in the period 1500-1815 – the early modern period – claimed that this was already, before the Industrial Revolution, an essentially modern economy – the First Modern Economy – understandable with the concepts thought appropriate to the post-1800 western world. At the same time, the Dutch subtitle claimed that this three hundred year period was the 'eerste ronde', the first round, of modern growth, implying that later growth also has a cyclical character and should not be cast as a stable, linear process. That is, the titles cast doubt both on our understanding of the origins of a modern economy and on our assumptions about what this modernity really is – or was.

These issues, both in theory and in their application to the Dutch past, are the stuff of ongoing debate. What I have called the 'revolt of the early modernists' is a body of scholarship questioning the assumptions under girding conventional periodization.[13] It has no single leader and is, in fact, barely a conscious movement. But, as ever more historians ponder the striking contemporary fact that the future is no longer what the past several generations have thought it would be, I expect that further questions will inevitably be raised about the conventions of periodization used by historians. All of the themes I have introduced so far – the microeconomic study of individuals interacting with institutions and markets, the role of networks that coordinate complex systems, and the study of historical

developments on both sides of the late-eighteenth century revolutionary divide – play a role in my current research. Briefly stated, it is an exploration of the role of consumer demand as an active force in long-term economic development. While Keynesianism has long accustomed us to accept that consumption levels are a critical component of modern economic prosperity, economists ordinarily ascribe importance to consumption only in the short run. In the long run, it is supply rather than demand that governs the growth of economies. But, when individuals are placed in the families and households in which they live and work, matters are not so simple. Household decisions about how to satisfy its needs and define its wants are directly related to decisions made within the household about work versus leisure, market labour versus household work, and consumption versus investment (such as investment in the health and education of children). My thesis is that an industrious revolution of more and harder market-oriented work arose among growing numbers of household in north-western Europe and the American colonies from the late seventeenth through the early nineteenth centuries, and that this industrious behaviour, increasing marketed output, was elicited by new consumer aspirations, by bundles of new consumer products that fashioned new lifestyles. These goods were rarely the products of technologically new industries; they were, taken separately, conventionally produced goods brought together by commerce and organization, but they were given new, compelling meaning and value by the imaginative acts of consumers. A new consumer society was built on novel commodities such as coffee, tea, tobacco and sugar, but it also featured new commercialised ways of consuming long established goods such as beer and bread. It featured new standards of domestic comfort, privacy and refinement and also a shift toward 'breakability', toward goods that depreciate quickly because of their materials and their embodied fashions. In all these ways and more, consumer expenditures rose, dependence on retail shops rose, and family members devoted more of their time to market-oriented work. Ironically, it would be the rise of modern industrialization in the nineteenth century which would establish the material basis for a new form of household – where mother stayed at home, producing broad range of home goods and services, while a male breadwinner became the sole contact with the market. This so called 'traditional family' was a creation of the century after 1850; however, rather than being the starting point of modernity it was its product.

These glimpses into my historical research and theoretical concerns are, perhaps, too brief to demonstrate anything, but, I hope, they render plausible my conviction that social reality moves in indirect but not in capricious ways. History will always retain the capacity to surprise us, but it is not fiction.

Notes

[1] Karl Popper, *The Poverty of Historicism* (London, 1960) p. 60n.

[2] The limitations of the English language prevent exact expression here. Science, in English usage, is generally confined to the exact sciences, while the scholarship of other disciplines – even those that call themselves 'social sciences' – is denied access to this term. Indeed, most English-speaking historians would respond with considerable hostility to the proposition that there is such a thing as 'historical science.'

[3] Naomi Lamoreaux, 'Economic History and the Cliometric Revolution', in A. Molho and G. Wood, eds., *Imagined Histories: American Historians Interpret the Past* (Princeton, 1998), p. 59.

[4] A movement of French historians who took their name from their journal, *Annales (Economes, Sociétés, Civilisations)*, which, since 1994 has taken the name, *Annales (Histoire, Sciences Sociales)*.

[5] Fernand Braudel, 'Histoire et Sciences Sociales: la Longue Durée', *Annales E.S.C.* 13 (1958), 725-53.

[6] The 'old' institutionalism takes us back to the German Historical School. The leading exponent of the 'new' institutionalism is Douglas North, *Institutions, Institutional Change, and Economic Performance* (Cambridge, 1990); 'Economic Performance Through Time', *American Economic Review* 84 (1993), 359-68.

[7] The new growth theory breaks with the neoclassical model in positing pervasive increasing returns to scale and the stressing the importance of (necessarily localized) spillover effects of investment in new technologies. For an overview of what is now a large literature, see Paul Romer, *et al*., 'Symposium on New Growth Theory', *Journal of Economic Perspectives* 8 (1994), 3-72. The spatial is emphasized in: Paul Krugman, *Geography and Trade* (Cambridge, Mass., 1991).

[8] The concept of path dependence breaks with the neoclassical assumption that in the long run, all 'accidental' forces are shaken off in the process of achieving a unique equilibrium value. Rather, path dependence affirms the existence of positive feedback effects and the consequent possibility of multiple equilibria, including sub-optimal 'pools of local attraction'. See: Paul A. David, 'Clio and the Economics of qwerty', *American Economic Review* 75 (1985), 332-37; Paul A. David, 'Historical Economics in the Long Run: Some Implications of Path Dependence', in G.D. Snooks, ed., *Historical Analysis in Economics* (London, 1993), pp. 29-40; Brian Arthur, 'Positive Feedback in Economics', *Scientific American* 262 (1990), 92-99.

[9] This was the first question I addressed, seeing it als one with both historical and contemporary (developmental) importance. Jan de Vries, *The Dutch Rural Economy in the Golden Age, 1500-1700* (New Haven, 1974) pp. 1-21. I have returned to this issue in 'De Boer', in H.M. Beliën, A.Th. van Deursen, and G.J. van Setten, eds., *Gestalten van de Gouden Eeuw. Een Hollands groepsportret* (Amsterdam, 1995), pp. 281-314; and 'De economische ontwikkeling van Friesland na het einde van de Friese Vrijheid', in J. Frieswijk, *et al.*, eds., *Fryslân. Staat en macht 1450-1650* (Hilversum and Leeuwarden, 1999), pp. 140-157.

[10] Ad van der Woude, 'The Volume and Value of Paintings in Holland and the Time of the Dutch Republic', in Jan de Vries and David Freeberg, eds., *Art in History, History in Art* (Santa Monica, California, 1991), pp. 285-330.
[11] Jan de Vries, 'Art History', in De Vries and Freedberg, *Art in History*, pp. 264-271.
[12] When a new institutional economist such as my Berkely colleague Oliver Williamson describes the first step in economics as the study of 'embeddedness' what he has in mind is full context, or wholistic social analysis. Oliver Williamson, 'The New Institutional Economics: Taking Stock, Looking Ahead', *Journal of Economic Literature* 38 (2000), 595-613.
[13] Jan de Vries, 'The Industrial Revolution and the Industrious Revolution', *Journal of Economic History* 54 (1994), 252-55.

Guido Geelen

A good work of art lifts you out of your earthly existence

The Dr A.H. Heineken Prize for Art 2000 was awarded to Guido Geelen, creator of sculptures, for the unorthodox and innovative way in which he uses the traditional material of clay.

Why this title? Because for me, good art is a necessity in my life, as important as food and drink. In addition, the existence of art has enriched my life enormously. So it's unimaginable for me that I should miss out on the pleasures of enjoying art.

Art: more than 'just messing about'

The word 'art' is often used much too lightly. To give an example, people are quickly inclined to describe a drawing by a five-year-old as art. I think they're way off the mark. Because a work of art is most definitely something more than just a thoughtless daub of paint on paper, or, as the Dutch artist Karel Appel famously

said: ‘just messing about a bit’. Creating a work of art is a very deliberate form of ‘messing about’, based on an inner need and vision.

All too often, people make quick judgments about art: a work of art is attractive or ugly, fantastic or worthless. In reality, however, appreciating and judging the quality of art requires some background knowledge. Similarly, an English poem cannot be fully appreciated and understood without some command of the English language.

In the past, as a child, I had no idea what art was, and in a way you could say that for me everything was art. But that’s a bit too easy, because it is definitely possible to draw a distinction. There are the traditional forms of art such as a painting, a sculpture, a piece of music, a building, and so on. But today, the representations of a road network on large motorway signs is also regarded as art.

With all this in mind, within the whole arsenal of what we call art, I would like to apply the distinctive characteristic ‘good’. After all, a work of art can be a failed painting, a bad building or a terrible piece of music.

A good work of art lifts you out of your earthly existence

A good work of art must immediately move your heart, touch your soul, reach into your bones. When you see and feel a good work of art, you forget for an instant that you are there. It’s a sort of religious experience. When you’re back with both feet on the ground again – something which might take a few hundredths of a second – you discover that something very complex is going on. You then try to analyse this complexity: Who created the work of art? How was it created? From which period does it date? In which context can it be placed? And so on.

The maker of a work of art is a specialist. Of course, he has his vision; that is a given. In my view, the starting point must always be the transformation of reality. A work of art portrays something of the here and now. It does something with what we see around us; it transforms it, captures it. It must also have something innovative, otherwise it’s just a repetition of that reality, not a transformation. The specialist has a number of means at his disposal to achieve that transformation. We can expect him to use those skills in an expert fashion - after all, being a specialist, he has trained for this. Of course, there are always people who are born geniuses, who simply have ‘the gift’, like Mozart. And there are also self-taught artists who have a particular desire to create something. By constantly working in and with the subject matter or the material, they gradually become more and more comfortable with it.

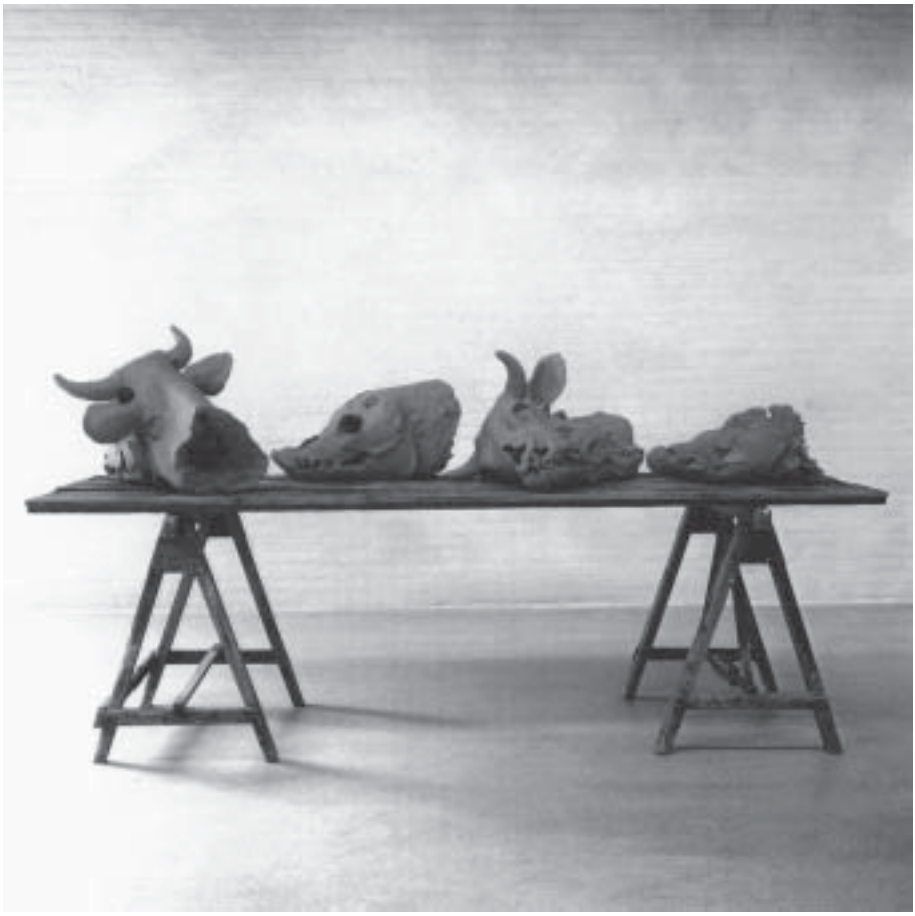
If the specialist has employed the factors I have just mentioned correctly, and has also incorporated into his work the clichés of life and death, old and new,

beautiful and ugly, black and white, light and dark, and so forth, then we might expect that the puzzle which is the creation of a good work of art is complete. However, a good work of art always has an additional, indefinable factor – call it mystery, emotion, intuition, chance - which plays a key role in the creation process.

If you come across a good work of art in the right place at the right time, it can evoke a response which can perhaps be described as the religious experience I spoke of earlier. Such moments are so intense and blissful that we suddenly feel as if there is more to this earth than, say, our day-to-day worries. That's what I mean by lifting us out of our earthly existence.

The inner need to create images

There have been a few moments when a good work of art has evoked a sensation like that in me – and without alcohol or drugs, I should add; a completely pure sensation, in other words. And when that happens, you obviously want to know how it happens and what precisely is happening. If it's a piece of music, for example, you might try to discover what was going on in the composer's mind. You can read about this, or you can try to discover what happens by making music yourself. I think the same thing happens with visual art. At an academy of art you learn by carrying out assignments how the process of creating visual images operates. At a certain point, you have enough insight and a need arises to strike



out on your own, to create your own images. You don't need any more assignments. You are then able to give form to your inner need to create an image. The question then, of course, is whether your own image is also able to evoke that intense experience. For me, the kick is using material to make something which gives me a feeling of intense happiness. But it's a very short-lived feeling. That's where the 'drive' comes from, that makes you go on producing new images.

Because that moment of bliss is so short-lived, I also try to experience 'quality moments' in my daily life. For me, the setting is important and also provides a frame of reference which enables me to hold my own. I expect the same commitment from the specialists around me, each in their various disciplines. That's why I can get so angry when fellow-specialists – artists, architects, design-ers, and so on - don't do their work well. Without asking to, I am then forced to encounter their pollution. This is anything but uplifting art.

The visual arts as my chosen field

Sadly, I'm no *homo universalis*, and as I can't make a contribution to society in all domains of art, I've chosen to try to excel in one artistic discipline. That discipline is the visual arts; sculpture is my field. Since 1985 I've used clay in my work because that's the material in which I can give form to my thoughts and feelings in the most striking way. Precisely because it's so unformed and meaningless on its own, clay creates possibilities for imparting form and meaning. The idea of creating an image from clay which is a result of my own actions, right down to the tiniest detail, is something that fascinates me enormously. In 1994 I also began working in bronze and later still in aluminium.

The Dr H.P. Heineken Prize for Biochemistry and Biophysics

The Dr H.P. Heineken Prize for Biochemistry and Biophysics 2000 was awarded to Dr James E. Rothman for clarifying the mechanism of intracellular membrane fusion.

The cell contains very small membrane-enveloped vesicles that carry a large variety of proteins between different compartments in the cytoplasm. This delivery process, which involves vesicle flow and membrane fusion, is vital for the growth and division of every cell. How this process comes about was a great mystery, and one of the great unsolved questions of biochemistry and cell biology.

James Rothman discovered the molecular principles of intracellular membrane fusion and demonstrated that the specificity of fusion was dictated by the pairing of snare proteins between membranes. This historic discovery provided a single unified principle for understanding important physiological processes, including the release of insulin into the blood, communication between nerve cells in the brain and the entry of viruses like hiv (the aids virus) to infect cells. Defects in the control of these pathways are important in diabetes and most likely also in certain cancers. Currently a major effort is under way to develop a new generation of drugs to control aids by blocking the membrane fusion process.

James Rothman was born in Haverhill, Massachusetts, usa, in 1950 and is an American citizen. He has a Ph.D. in Biological Chemistry (Harvard Medical School), and has worked at the Sloan-Kettering Institute in New York. Dr Rothman has also been vice-chairman of the Sloan Kettering Institute since 1994. Amongst his other honours, Dr Rothman has received the Gairdner Foundation International Award (1996), the King Faisal International Prize in Science (1996) and the Lounsbery Award of the National Academy of Sciences (1997).

The Dr A.H. Heineken Prize for Medicine

Professor Eric R. Kandel of the Howard Hughes Medical Institute, Columbia University, New York, has been awarded this year's Dr A.H. Heineken Prize for Medicine. The prize has been awarded to Professor Kandel for his pioneering research on the molecular mechanisms underlying learning processes and memory. Using the marine snail *Aplysia californica* as a model, Eric Kandel and colleagues

have managed to bridge the enormous gap between the physiology of behaviour and classical psychology. The simple structure of the nervous system in this primitive invertebrate is especially well suited to investigating learning and memory formation at the cellular and molecular level. In an impressive series of neurophysiological experiments, now used as standard examples in most neuro-scientific textbooks, the group led by Eric Kandel has explained the fundamental neuronal mechanisms underlying learning processes at the cellular level. This work and recent studies by Kandel and colleagues involving genetically modified mice have led to the discovery of neuronal mechanisms responsible for non-associative and associative learning processes (for example classical conditioning) and for the development and functioning of short- and long-term memory in lower and higher animal species. His discoveries open up entirely new ways of understanding human memory and its disorders.

Eric Richard Kandel was born in Vienna, Austria, in 1929 and received his medical degree from New York University School of Medicine in 1956. He is a University Professor of Physiology and Psychiatry at the Center for Neurobiology and Behavior, Columbia University College of Physicians and Surgeons. He is also a Senior Investigator at the Howard Hughes Medical Institute. Prof. Kandel has received an impressive list of honorary degrees, awards and other marks of distinction in the course of his long career.

The Dr A.H. Heineken Prize for Environmental Sciences

Professor Poul Harremoës (1934) of the Technical University of Denmark at Lyngby is the winner of this year's Dr A.H. Heineken Prize for Environmental Sciences. The prize has been awarded to Professor Harremoës in recognition for his contributions to the theory of biofilm kinetics in relation to biological waste water treatment and for his successful organisation of the international scientific community in water pollution research and control.

Poul Harremoës belongs to the pioneers who tried to track down pollution with radioactive tracers. As head of the Department of Environmental Science and Engineering at the Technical University of Denmark, he has contributed to making this department one of the biggest and broadest university departments world-wide and an acknowledged world leader in the field. Dominant fields of research have been oxygen depletion in rivers, nitrification-denitrification and biofilm kinetics applied to waste water treatment.

Professor Harremoës promotes scientific co-operation in water pollution research and control on an international scale in the International Association for Water Pollution Research and Control, iawprc (now International Water Association, iwa).

Poul Harremoës was born in Denmark in 1934. He graduated from the Technical University of Denmark in 1957 and from the Massachusetts Institute of Technology in 1959. He became a full professor at the Technical University of Denmark in 1972 and the head of the Department of Environmental Engineering. Professor Harremoës has lectured all over the world. He received the Stockholm Water Prize on behalf of the Department in 1992. Professor Harremoës has initiated and organised numerous international conferences and has been a member of the scientific committees for many such conferences. He joined the Scientific Committee of the European Environmental Agency in 1994.

The Dr A.H. Heineken Prize for History

Professor Jan de Vries of the Department of History at the University of California in Berkeley has been chosen for this year's Dr A.H. Heineken Prize for History. The prize has been awarded to Professor de Vries for his pioneering research into the development of the European economy between 1500 and 1800.

Professor Jan de Vries has conducted pioneering research into the early modern history of the European economy, specifically in the Dutch Republic (the Northern Netherlands). Professor De Vries has used economic theories and concepts to organise a huge quantity of wide-ranging historical data in a most original and transparent manner, allowing him to reveal astonishing viewpoints and patterns. He has succeeded in tracing the origins of the modern market economy and has shown the transition from the early modern economy to industrial society. He has gone further than any other historian thus far in analysing the role that urbanisation played during this period.

Professor De Vries's studies have focused mainly on the way in which different economic parties (households, farmers, artists, municipal authorities or groups of labourers) responded to market trade and, in turn, contributed to its development and expansion. He has succeeded in showing the links between macroeconomic developments and history at local level, so that his work often has a very practical grounding. For example, he has analysed the role of the barge in the modernisation of Holland, and showed how the flourishing art market depended on economic variables. His work not only makes a highly significant contribution to the study of economic and demographic history, but also provides signposts for art historical research.

Jan de Vries was born in the Netherlands in 1943. He moved to the United States as a boy and is an American citizen. He obtained his Ph.D. in 1970 (Yale Univer-sity). Since 1977 he has been Professor of History and, since 1982, Professor of History and Economics at the University of California, Berkeley, usa.

Professor de Vries and Ad van der Woude co-authored the standard work *The First Modern Economy. Success, Failure and Perseverance of the Dutch Economy from 1500 to 1815*, which was awarded the Gyorgy Ranki Prize for the best book on the economic history of Europe. Professor De Vries was elected as a foreign member of the Royal Netherlands Academy of Arts and Sciences in 1989. He is also a member of the British Academy, the Society for Dutch Literature (Maatschappij der Nederlandse Letterkunde), and various international scholarly organisations. From 1991 to 1993 he held the post of president of the Economic History Association, and he is currently an editor of the *Journal of Economic History*.

The Dr A.H. Heineken Prize for Art

The Dr A.H. Heineken Prize for Art 2000 was awarded to Guido Geelen for the unorthodox and innovative way in which he uses the traditional material of clay. Guido Geelen, who prefers to be called a creator of sculptures, creates images in which ceramics form an important element. These images are often constructed from ceramic reproductions of a wide variety of everyday objects. Guido Geelen has, for example, created a ceramic sculpture composed of dogs, car tyres, vacuum cleaners, televisions and computers, all made of red clay compacted together. By organising these items in a particular way and then stacking them on top of one another, Guido Geelen has created a ceramic wall which fascinates viewers despite – or perhaps because of – its seemingly morbid character.

Guido Geelen was born in Thorn, the Netherlands, in 1961. He attended the Institute for Draughtsmanship, Craftsmanship and Textiles (tehatex) in Tilburg and went on to study at the Academy of Fine Arts for two years. Guido Geelen was awarded the incentive prize for applied art by the Amsterdam Fund for the Arts in 1988, the Charlotte Köhler Prize for Sculpture in 1989, and the incentive prize for fine art by the Amsterdam Fund for the Arts in 1990.

Guido Geelen's work can be viewed at the Stedelijk Museum of Modern Art in Amsterdam, the ptt Art and Design Collection in The Hague, the Kruithuis Museum for Contemporary Art in 's-Hertogenbosch, the North Brabant Museum in 's-Hertogenbosch, the Kröller-Müller Museum in Otterlo, the Boijmans van Beuningen Museum in Rotterdam and the De Pont Museum in Tilburg.

The 2000 Dr A.H. Heineken Prize for Art (Dfl 100,000) is being awarded to an artist who works in ceramics. The Netherlands has a rich tradition in ceramic art, ranging from the centuries-old Delftware industry to the innovative Dutch pottery from around **1900** (when independent potters first began to flourish) and right up to the revolutionary ceramic work being produced in our own day. It is

precisely this aspect that the jury has decided to emphasise by awarding the Dr A.H. Heineken Prize for Art to artist Guido Geelen for his entire oeuvre.

The jury that has awarded this prize has acted independently of the Royal Netherlands Academy of Arts and Sciences. However, a number of Academy members took part in the jury in a private capacity.

Dr H.P. Heineken Prize for Biochemistry and Biophysics

1964 Erwin Chargaff
1967 Jean L.A. Brachet
1970 Britton Chance
1973 Christian de Duve
1976 Laurens L.M. van Deenen
1979 Aaron Klug
1982 Charles Weissmann
1985 Bela Julesz/Werner E. Reichardt
1988 Thomas R. Cech
1990 Philip Leder
1992 Piet Borst
1994 Michael J. Berridge
1996 Paul M. Nurse
1998 Tony J. Pawson
2000 James E. Rothman

Dr A.H. Heineken Prize for Medicine

1989 Paul C. Lauterbur
1990 Johannes J. van Rood
1992 Salvador Moncada
1994 Luc Montagnier
1996 David de Wied
1998 Barry J. Marshall
2000 Eric R. Kandel

Dr A.H. Heineken Prize for Environmental Sciences

1990 James E. Lovelock
1992 Marko Branica
1994 BirdLife International (Colin J. Bibby)
1996 Herman E. Daly
1998 Paul R. Ehrlich
2000 Poul Harremoës

Dr A.H. Heineken Prize for History

1990 Peter Gay
1992 Herman van der Wee
1994 Peter R.L. Brown
1996 Heiko A. Oberman
1998 Mona Ozouf
2000 Jan de Vries

Dr A.H. Heineken Prize for Art

1988 Toon Verhoef
1990 Marrie Bot
1992 Carel Visser
1994 Matthijs Röling
1996 Karel Martens
1998 Jan van de Pavert
2000 Guido Geelen