
Coding and representation: Searching for a home in the brain

Endel Tulving

The idea that mental experience may leave residue in the soul, or mind, that allows later remembering of the experience is as old as recorded history, and probably older. The idea that this residue is physical, somewhere in the brain, is more recent, having been first proposed by Robert Hooke (1627–1703) who thought that memory is ‘as much an Organ as the Eye, Ear, or the Nose’, and that it has ‘its situation somewhere near the Place where Nerves from the other Senses concur and meet’ (Young 1965, p. 287).

The existence of this ‘residue’ with a remarkable staying power is now taken for granted, but much about it has remained baffling. How is it formed? What is its nature? What kind of thing, or entity or stuff is it? What is the relationship between the experience and its residue? What is the relationship between the residue and remembering that it enables? Where does the residue reside? Does every experience leave a residue? If not, then what determines which ones do and which ones do not? If yes, what kind of a place is it that can ‘hold’ an individual’s untold experiences? Does the residue last forever? (Not many scientists believe this, but I think that some do.) Does it last at least as long as the individual is alive? These and related questions have been raised and debated, sometimes hotly debated, throughout the human intellectual history. At the present time, no one knows what the answers are, although we have undoubtedly made progress in getting a better grip on the questions.

The terms that have been used to refer to the memorially relevant components of the after-effects of experience have varied with the fashions of the times, the accumulated pertinent knowledge already available, and even the languages and dialects spoken by those who have thought deeply about the matter. A frequently used term is ‘representation’, another is ‘coding’—as in the title of this section of the book. Other well-known terms are ‘engram’, ‘memory image’ and ‘memory trace’. Each has its own connotations that vary from context to context and even from writer to writer, although the concept lying behind all these terms has been and continues to be relatively unambiguous.

Memory trace

A memory trace is the neural change that accompanies a mental experience at one time (time 1) whose retention, modified or otherwise, allows the individual later (at time 2) to have mental experiences of the kind that would not have been possible in the absence of the trace.

The critical ingredients of this definition are: (1) mental experience at time 1; (2) neural change; (3) retention of aspects of the change; (4) mental experience at time 2; and (5) the relationship between experiences at time 1 and time 2. The concept of memory trace ties together these features in an organized whole. Every single component is critical in the sense that its absence would be tantamount to the absence of the whole.

Some features of the definition may be worth emphasis, in order to minimize misunderstanding. First, the definition applies to cognitive memory, the kind of memory that has to do with mental experience. It has nothing to say about learning and memory in which mental experience of the kind that the definition refers to is missing. Thus, much of what has been written in the traditional literature on skill learning, conditioning, priming and simple forms of associative learning seems to have done perfectly well without invoking representation-like concepts of the mental type to which I refer here. All kinds of mental concepts, of course, were anathema to many psychologists during the heyday of behaviorism.

Secondly, the definition implies that not all but only some of the physiological/neural after-effects of an experience constitute the memorially significant 'residue', i.e. the memory trace. The question of which ones, and the whole issue of how to separate the wheat of the memory trace from the chaff of all sorts of neural activity that has nothing to do with memory, remains a challenging problem for the neurobiological side of the science of memory. The definition also assumes that the memory trace is a dynamic, changing, malleable entity (Dudai 2002; Nader 2003) rather than a 'fixed, lifeless' sort of thing that many cognitive psychologists came to look down upon, thanks to Sir Frederick Bartlett's well-known disdainful phrase.

Thirdly, the definition explicitly distinguishes between, and relates to each other, two kinds of entities: physical (neural change) and mental (experience), and thereby brushes on one of the central issues of our science: how does the mental arise out of the neural? Talking about mental experiences as separable from neuronal processes may be questioned by those who think that because mental experience depends on neuronal processes it must also be in some sense *reducible* to neuronal processes. Aside from the problem created by many meanings of the concept of reduction, the logic of this type of argument has always escaped me. At least for practical purposes—to get on with the

'business'—I find it more congenial to operate in a conceptual framework in which the mind depends on the brain but also has properties and capabilities that are different from those of the brain.

Fourthly, the definition reminds us that the memory trace is not just mere residue, or after-effect of a past experience, not just an incomplete record of what was. It is also a recipe, or a prescription, for the future. However, as it is usually only an impoverished record, it is also only an unreliable guide to what will happen in the future. What actually happens—what kind of a future experience it enables—depends not only on its properties at the time of attempted retrieval but also on the conditions prevailing at the time of retrieval. This was one of the deep insights of Richard Semon, the early and unappreciated prophet of memory, an insight that the subsequent experimental work by others has more than vindicated (Schacter 2001a).

Brain and mind

Fifthly, the definition of memory trace is given in terms of something happening in a physical object, the brain, and in that sense it is tempting to think of the trace also as a physical object, or physical entity. It is not. A definition of memory trace in terms of physical changes in the brain does not make the memory trace a physical entity itself. Memory trace is a change, and change is not an entity. It is a relationship (difference) between two things that are physical objects, the brain at time 1 ('immediately before' the experience) and time 2 ('immediately after' the experience), but the relationship itself is not a physical object. To illustrate: think of drawing a straight line. After you have drawn it, the line exists physically with all its properties. Then you grab the pencil again and make the same line a bit longer. After you have done it, the 'second' line exists physically with all its properties. The difference between the two does not exist anywhere other than in your mind. You can of course arrange for a comparison between them, by making a copy of the first line while it exists, and you can note their differences in length. However, the difference you note exists only in your mind, in physical reality there exist only two lines.

Sixthly, because there is no object or entity in the brain that can be said to *be* a particular memory trace (i.e. to represent a particular experience), it is in principle not possible to observe it as such, to identify it as such or to determine its properties as such. Memory trace, as defined, is something that makes something else possible, if and only if some other conditions are fulfilled at time 2. Since time 2 has not yet arrived, the relationship between the memory trace and the experience it yields is indeterminate. The situation is not unlike that of the relationship between an elementary particle's location and its momentum as described by Heisenberg in the 1920s.

Seventhly, the definition relates ‘mental experience’ to a physical happening (neural change). It is important to note that both experience and neural activity are part of the reality with which cognitive neuroscience deals. However, there is also a difference between them, one that many people get excited about. The difference is that the experience is real (only) from the first-person’s point of view, whereas the neural events are real, at least in principle and increasingly so in actual scientific practice, from the third-person point of view. The ‘subjectivity’ of the first-person experience has traditionally been held as an obstacle to the study of such experience. That is not a problem, however. The only ‘problem’ is that the first-person experiences are not *directly* observable. Yet, like countless other things in the universe that scientists study, they are *indirectly* observable. In psychology, indirect observation has been successfully applied since day one. The most respectable and oldest part of experimental psychology, psychophysics, is all about indirect observation of ‘subjective’ experiences, and so is much about cognitive study of memory. The important criteria are not ‘objectivity’ and ‘subjectivity’. The important criteria are the possibility of empirical validation and rejection of hypotheses, reliability and replicability of the empirical findings, and the coherence of the story (theory) that relates facts to one another in a way that would not be possible without science.

Eighthly, the definition reflects the basic assumption that memory trace, like any other concept, can be fully understood only in relation to other concepts. Here, the other concepts are experiencing something at time 1 (dealt with under the concept of *Encoding* in Section 6 of this volume), and experiencing something at time 2 (dealt with under the concepts of *Retrieval* and *Remembering* in Sections 10 and 11, respectively, of this volume).

Epilogue

In *Elements of Episodic Memory* (Tulving 1983) I talked about memory traces synonymously with engrams. I defined engram as the state of the memory system before and after the encoding of the event, as well by its position and relationship to other hypothetical concepts in GAPS, the general abstract processing system, which I proposed as the conceptual framework for studying episodic memory. The conceptualization of memory trace offered here is not greatly different from that of almost a quarter century ago. The main difference has to do with the ‘locality’ of engrams. Then they resided in the ‘memory system’, now they have found a home in the brain. I think this is progress.

Science of Memory: Concepts

Edited by

Henry L. Roediger III

James S. McDonnell Distinguished University Professor
at Washington University in St. Louis

Yadin Dudai

Sela Professor of Neurobiology at the Weizmann Institute
of Science, Rehovot, Israel

and

Susan M. Fitzpatrick

Vice President of the James S. McDonnell Foundation

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Oxford University Press, Inc., publishes works that further
Oxford University's objective of excellence
in research, scholarship, and education.

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Copyright (©) 2007 by Henry L. Roediger III,
Yadin Dudai, and Susan M. Fitzpatrick.

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016
www.oup.com

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise,
without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data

Science of memory: concepts / {edited by} Henry L. Roediger III, Yadin Dudai, and Susan M. Fitzpatrick.
p. cm. -- (Science of memory)

Includes bibliographical references and index.

ISBN-13: 978-0-19-531044-3 (pbk.)

I. Memory. I. Roediger, Henry L. II. Dudai, Yadin. III. Fitzpatrick, Susan M., 1956-
BF371.S45 2007
153.1'2--dc22
2006029621

Typeset in Minion by
Cepha Imaging Private Ltd., Bangalore, India

Printed in Great Britain
on acid-free paper by Ashford Colour Press Ltd., Gosport, Hampshire

10 9 8 7 6 5 4 3 2 1