MECHANISMS OF LEARNING AND MOTIVATION:

A Memorial Volume to Jerzy Konorski

Edited by
Anthony Dickinson

Robert A. Boakes

Psychology Press

MECHANISMS OF LEARNING AND MOTIVATION:
A Memorial Volume to Jerzy Konorski

List of Contributors

M. Sebastian Halliday
University of Sussex
Brighton, England

Eliot Hearst
Indiana University
Bloomington, Indiana

Allan R. Wagner
Yale University
New Haven, Connecticut

Robert A. Rescorla
Yale University
New Haven, Connecticut

John W. Moore
University of Massachusetts
Amherst, Massachusetts

Nicholas J. Mackintosh
University of Sussex
Brighton, England

Michael J. Morgan
University of Durham
Durham, England

Anthony Dickinson
University of Cambridge
Cambridge, England

Michael F. Dearing
University of Sussex
Brighton, England

Robert A. Boakes
University of Sussex
Brighton, England

Kazimierz Zielinski
Nencki Institute
Warsaw, Poland

Jeffrey A. Gray
University of Oxford
Oxford, England

J. N. P. Rawlins
University of Oxford
Oxford, England

J. Feldon
University of Oxford
Oxford, England

Jadwiga Dabrowska
Nencki Institute
Warsaw, Poland

Ronald G. Weisman
Queen's University
Kingston, Canada

Peter W. D. Dodd
Queen's University
Kingston, Canada

Vincent M. LoLordo
Dalhousie University
Halifax, Canada

Sara J. Shettleworth
University of Toronto
Toronto, Canada

William K. Estes,
The Rockefeller University
New York, New York

MECHANISMS OF LEARNING AND MOTIVATION:

A Memorial Volume to Jerzy Konorski

Edited by
Anthony Dickinson
University of Cambridge
Robert A. Boakes
University of Sussex

Psychology Press
Taylor & Francis Group
NEW YORK AND LONDON

First Published 1979 by Lawrence Erlbaum Associates, Inc.

This edition published 2014 by Psychology Press
711 Third Avenue, New York, NY 10017

and by Psychology Press
27 Church Road, Hove, East Sussex, BN3 2FA

Psychology Press is an imprint of the Taylor & Francis Group, an informa business

Library of Congress Cataloging in Publication Data

Main entry under title:

Mechanisms of learning and motivation.
Includes bibliographical references and indexes.
1. Learning in animals. 2. Learning, Psychology of. 3. Motivation (Psychology) 4. Konorski, Jerzy. I. Konorski, Jerzy. II. Dickinson, Anthony. III. Boakes, R. A.
QL785.M5 156'.3'15 78-11370

ISBN 978-0-898-59460-7 (hbk)

Publisher's Note
The publisher has gone to great lengths to ensure the quality of this reprint but points out that some imperfections in the original may be apparent.

Contents

Preface

This volume consists of a series of chapters honoring a Polish psychologist and neurophysiologist who died in 1973. Although his name was familiar to all of the contributors, many had had no personal contact with him and had gained acquaintance with his ideas only through his publications. This unusual venture deserves some explanation.

In the autumn of 1975 the editors had become increasingly interested in Konorski's research and theories, and noted that his name was being cited more frequently in current research journals. We discussed the idea of organizing a third in the series of occasional conferences on animal psychology that have taken place at the University of Sussex, with the aim of providing a forum for discussion of Konorski's work. A list of possible topics, drawn from the range represented in Konorski's *The Integrative Activity of the Brain* (1967), was tentatively sketched. Reactions to this proposal, and to the prospect of delivering a paper that would analyze some aspect of Konorski's work and relate this to other theoretical approaches or to recent research developments, were sought from a number of psychologists. Somewhat to our surprise the reactions were uniformly favorable, frequently very enthusiastic, and of great help in determining the program for the conference. With this evidence of widespread interest, plans went ahead and the event took place between June 30 and July 2, 1977.

The chapters in this volume are revised versions of the papers presented at this conference. The theme ensured a degree of commonality at such events, and a major purpose of many of the revisions was to take account of the exchange of ideas and information that had occurred. Although Konorski's last book (1967) covered a wide spectrum of issues in both

animal and human psychology, the overwhelming emphasis in the present volume is on learning and motivation in animals. This reflects both the editors' own interests and the fact that Konorski's major *empirical* contribution was in this field. This choice should not be taken as a judgment that his theoretical ideas do not have wider importance.

In the past decade, major changes have occurred in the study of animal behavior within Western psychology. New developments have taken a variety of forms and, although workers in different research areas have shared the adoption of many basic assumptions and the rejection of others, there has been no widely accepted framework providing a guide to these changes. As Halliday discusses in the first chapter of this book, there appears to have developed a rare situation whereby a number of people have independently found that their present ideas and interests are more closely related to those pursued in Warsaw over the past half-century than to the work of their own intellectual predecessors in the West. It is as if groups from different starting points had entered an apparently new region, started to explore, and then discovered — often unexpectedly — the existence of a sketch, uncertain in parts, but of the region as a whole. The points of departure represented here include concern with the nature of classical and of instrumental conditioning; the analysis of inhibitory effects; attempts to understand the physiological basis of learning; research on autoshaping and its relevance to problems studied in the context of operant conditioning; and problems arising when various kinds of stimuli, responses, reinforcers, and species are compared. Within each area of special interest there naturally exist different points of view and various opinions of the value of specific ideas suggested by Konorski; but overall there seems agreement that Konorski's map provides a better general guide than any other.

There are strong pressures on researchers to take a somewhat parochial view of their field. These, we feel, have contributed to the neglect in the West of Konorski's work. One purpose of the conference was to encourage contributors to present an account of research on some specific topic in a more general context than is usual. Our hope is that this will be considered valuable and that the present book will help to place Konorski's work in the proper perspective.

Permission to reproduce quotations from Konorski's publications was given by the Cambridge University Press, the University of Chicago Press, and Prentice-Hall, Inc. The bibliography of Konorski's work is based on one published in *Acta Neurobiologiae Experimentalis,* with the permission of its editor, and supplemented by additional references kindly supplied by Professor Zielinski.

Our thanks go to the contributors for their willing and prompt cooperation and to our colleagues from the Laboratory of Experimental Psychology

at the University of Sussex, especially to Sebastian Halliday and Nick Mackintosh for their valuable advice. The conference was made possible by grants from the Experimental Psychology Society and the European Training Program in Brain and Behavior, and by the help of many undergraduate, graduate, faculty, and staff members of the Laboratory of Experimental Psychology. We are grateful for the additional support toward travel expenses provided by the Society for the Experimental Analysis of Behavior, the Royal Society, and the Polish Academy of Sciences. Finally, we would like to thank our wives for their patient help during the conference and the preparation of this book.

A. DICKINSON
R. A. BOAKES

Page Intentionally Left Blank

1 Jerzy Konorski and Western Psychology

M.S. Halliday
University of Sussex

The boundaries between the disciplines of physiology, physiological psychology, and psychology are ill-defined and constantly shifting; Konorski always described himself as a physiologist, but he made major contributions in all these areas during his long and productive life. The focus of this volume is on conditioning, and so this chapter concentrates on this aspect of his work, while recognizing that he himself would certainly have deplored any arbitrary divisions between his various interests. In particular, I would like to consider why 20 years ago Konorski was largely ignored in the West, whereas today he is seen as a figure of great importance.

The mid-1950s were a significant turning point both in Konorski's career and in the development of American learning theory. In an autobiographical sketch Konorski (1974) wrote:

> There are at least two reasons why I can divide my postwar life into two distinct periods, with the demarcation point being 1955. First, 1955 was the year in which, two years after Stalin's death, the "thaw" began, when Kruschev came to power in the USSR and dissociated himself sharply from the Stalinist period. This was immediately reflected in all fields of cultural life in the USSR and even more so in Poland. In my own field the pseudo-Pavlovian indoctrination vanished completely, and I stopped being a revisionist and a servant of capitalism [p. 207]. © 1974. Adapted by permission of Prentice-Hall, Inc., Englewood Cliffs, N.J.

He is here referring to a period, beginning about 1949, during which Pavlovian theory was assimilated to Stalinist orthodoxy, and in which he, as a critic of Pavlov, came under heavy attack; there was even a time during which he was given "leave of absence," though remaining as head of the

Department of Neurophysiology at the Nencki Institute (Dr. W. Wyrwicka, cited in Mowrer, 1976). Konorski refused to give way under this pressure, and his integrity was ultimately rewarded by full recognition in his own country. 1955 was also significant because in that year the Nencki Institute moved from its temporary location in Lodz to a new and specially constructed building in Warsaw, replacing the original Institute which had been destroyed by the Nazis. Konorski had played a major part in planning this building and remarked (1974), "I can boast that the Department of Neurophysiology, with all its virtues and defects, is my own child (p. 208)." Thus for the first time since before the war, he had satisfactory facilities for carrying out his research.

Around 1955 American learning theory was also at a transitional point. Up to that time the subject had been dominated by the comprehensive learning theories developed in the 1930s, which attempted to explain all learned behavior using a limited number of relatively simple principles and largely without regard to physiological evidence. By the end of the decade this approach seemed inadequate. Attention was being concentrated on much more limited areas of behavior and on smaller-scale theories; at the same time physiological psychology was making major advances, and Skinnerians were advocating a nontheoretical approach. A representative statement of the old viewpoint, together with some disquiet about its adequacy, can be found in *Modern Learning Theory: A Critical Analysis* (Estes et al., 1954). In this scholarly work the authors attempted a critical survey of the major theories of learning, in the belief that (1954) "the major contributions of the outstanding theorists of our time have been made" and that "their contributions have set the framework for the larger part of the current theoretical literature and have placed their stamp on the vigorous output of our experimental laboratories (p. xi)." Their intention was not merely critical but also constructive. Their interest in the future was clearly expressed in the introduction, where they state that their "primary concern is with learning theory as it may function in the long-range development of a science of behavior (p. xv)." At some time during the meeting at which their book was planned, some of the participants must surely have looked to the future and wondered who would be seen as the significant learning theorists in a generation's time. If that group of psychologists had been asked to draw up a list of likely candidates from those born between about 1900 and 1915, it is almost certain that, however long the list, the name of Konorski would not have appeared. It is not surprising that there is no mention of him in *Modern Learning Theory*, which is primarily a study of five American theorists; but in ignoring him, the authors were confirming the judgment of American psychologists in general. In 1954 there were no references to Konorski in the *Journal of Comparative and Physiological Psychology,* the *Journal of Experimental Psychology,* the *Psychological Bulletin,* or the *Psychological Review.* He was not mentioned in the two

major textbooks that had recently appeared, Osgood's *Method and Theory in Experimental Psychology* (1953) or Woodworth and Schlosberg's *Experimental Psychology* (1954), and both editions of Hilgard's *Theories of Learning* (1948 and 1956) are equally silent. In 1954, therefore, there appeared to be universal agreement about the unimportance of Konorski's contribution to psychology.

In the second half of the 1950's and in the 1960s, sporadic references began to appear in the literature, and their frequency has increased up to the present day. The Science Citation Index gives six references to Konorski in 1961; by 1970 this had risen to 35, while in 1976 there are no less than 130 citations. Over the past 20 years Konorski has moved from a position of total obscurity, at least in the West, to that of a major figure in the study of learning. It is interesting to consider why there has been this remarkable reversal in the perceived importance of his work.

There are, of course, straightforward reasons for this change. Scientific contacts with Poland in the 1930s were slight; they were totally disrupted by the war and in the postwar years. Much of Konorski's work had been published in Polish or Russian and was unavailable in England or America. Furthermore, apart from a visit to London in 1946, it was not until the late 1950s that Konorski was free to visit Western countries. As he remarked in his autobiography (1974):

> This first visit of mine to America [in 1957] had very great significance for the further development of our scientific work and its relation to the work going on in the United States. Owing to my connections with American scientists, almost every member of our laboratory was able to visit the United States for one or two years, which gave him (or her) the opportunity to become acquainted directly with the American investigations; this significantly enlarged their scientific horizons. Consequently, our work became much better known by American scientists than it was before [p. 209].

Together with the publication of *Integrative Activity of the Brain* by the University of Chicago Press in 1967, this does much to explain the increasing prominence of Konorski in the 1960s and 1970s. However, it does not entirely account for the neglect he had suffered in previous years. To understand this fully, we need to consider more carefully both the nature of Konorski's contribution and the changing climate in American psychology over the past 20 years.

It is important to realize that much of the work for which Konorski is now celebrated had been published well before 1955, a considerable part of it in the 1930s. This is now seen as pioneering work which went largely unrecognized at the time, but which is directly relevant to modern thinking about animal psychology. But if the work was unrecognized, it should not have been unknown. Konorski and Miller had published two theoretical papers in the *Journal of General Psychology* (1937a, 1937b) in which they described their own work and criticized Skinner's distinction between two

types of conditioned reflex. These papers were widely cited for a time and are discussed in both Skinner's *The Behavior of Organisms* (1938) and in Hilgard and Marquis's *Conditioning and Learning* (1940), two books which would have been on the shelves of any animal psychologist in the 1940s and 1950s. Apparently the topics discussed in these papers simply did not seem important at that time. This is borne out by the treatment of the operant–respondent distinction in the chapter on Skinner in Estes et al., *Modern Learning Theory*; only a paragraph is devoted to this topic, and this is dismissive in tone. For example (1954):

> No attempt is made to reduce this differentiation [operant/respondent] experimentally. On the contrary there is an insistence on maintaining it. Although many experiments have shown that behavior which Skinner terms "operant" obeys the same laws as the static laws which define "respondents" when the appropriate experimental conditions are introduced, these are rejected because measures deemed inappropriate to operant behavior are introduced [p. 289].

If the discinction between operant and classical conditioning could be treated in this cavalier fashion, it is hardly surprising that the work of Konorski, whose main contribution up to that time had been to illuminate the nature of that distinction, passed without notice.

A later source for Konorski's ideas which should also have been familiar is *Conditioned Reflexes and Neuron Organization,* (1948). In England this book received favorable reviews in *Mind,* the *British Medical Journal,* and the *Times Literary Supplement,* but no mention in *Nature* or any of the psychological journals. Nevertheless, it was widely read by psychologists and was recommended, for example, to Cambridge undergraduates around 1960. In America, however, the book received little or no attention: Only 399 copies were sold in the United States (Mowrer, 1976) and it was not reviewed by any major psychological journals, although it received some attention in medical publications. This can hardly have been accidental, for it was published by the Cambridge University Press and consequently must have come to the attention of the review editors of the psychological journals; they presumably decided that a review would not elicit much interest. Konorski himself commented (1974):

> My explanation of this fact is that at that time experimental psychology was strongly Skinnerish or Hullian, and physiological explanations of the mechanisms of conditioned reflexes were utterly unpopular. I suspect that either people did not read the book at all, not being attracted by its title... or else if they had it in their hand, they rejected it [p. 204].

Despite this lack of interest in his book, Konorski was gratified to discover on his first visit to the United States in 1957 that he was by no means unknown. The overall picture is therefore one of neglect rather than of

ignorance, and the apparent blindness of Western psychologists to the value of Konorski's work was less a consequence of sensory deprivation than of unwillingness to see.

What, then, were the characteristics of his work that made it so unattractive to Western behaviorists at that time? A number of somewhat interrelated factors have combined to produce this effect. First, Konorski was deeply concerned with conditioned reflexes, even describing instrumental learning as a conditioned reflex of Type II. Western behaviorists from the 1920s onward had taken over the primitive ideas of classical conditioning, but had gone little further than this. Konorski himself made this point very clearly (1948):

> As is well known, conditioned reflexes are of particular importance in the American behaviourist school of psychology. But if the manner in which behaviourism exploits the science of conditioned reflexes is analysed more closely, it will be found that it consists chiefly in utilizing the basic and most elementary concepts of this field, i.e. the Pavlovian *nomenclature,* for the denomination of particular phenomena.... How little the achievements of the Pavlov school have been taken into account, even by those investigators who have been most directly concerned with the study of conditioned reflexes, may be judged, for example, in the chapter on the conditioned reflex written by Dr. H. S. Liddell of Cornell University in Fulton's famous monograph, *Physiology of the Nervous System.* Stating that... Pavlov's theory... "is at present of historical interest only," Liddell simply disregards the huge body of experimental evidence which forms the bases of the theory and which must find explanation in some way or another.... Liddell takes the inadmissible course of drawing a veil of silence over most of the facts and deals only with the most elementary phenomena in the field. One regrets to have to say that many American authors behave in exactly the same way [p. 2].

It was precisely this state of affairs that Konorski hoped to remedy by the publication of his book, but the method he chose took little account of the nature of behaviorism. He believed that the main source of resistance to Pavlovian work arose from dissatisfaction with Pavlovian physiological explanations—a dissatisfaction which he fully shared. If, therefore, Pavlovian findings could be brought into line with the neurophysiological tradition of Ramon y Cajal and Sherrington, the way would be cleared for a rapproachement between the American and Russian approaches to the benefit of both. Konorski had first been introduced to Sherringtonian neurophysiology in 1933 on his return from Pavlov's laboratory in Leningrad, by Dr. Liliana Lubinska, who later became his wife and scientific collaborator. Pavlov's view of the activity of the cerebral cortex involved the interplay between waves of cortical excitation and inhibition, and was totally alien to Western neurophysiology, with which Konorski now became familiar. In writing about this period of his life he noted that (1948) "as the years passed I became more and more convinced that Pavlov's theory was not correct, as it could not be reconciled with the evidence of

general physiology of the central nervous system (p. xix)." As he wrote later (1974), "One of the two theories should be rejected in toto, and the facts so far explained by the rejected theory should be reinterpreted in the framework of the other theory (p. 198)." His earlier book, which was the fruit of many years thought and study, was a remarkably effective attempt (1974) "to explain the whole bulk of the experimental work collected by Pavlov's school by the Sherringtonian principles of functioning of the central nervous system (p. 198)." Unfortunately the method that he chose might have been deliberately calculated to repel the Western psychologists to whom it was addressed. At that time learning theorists were profoundly unimpressed by physiology and were concerned with explanations of behavior in terms of intervening variables and hypothetical constructs independent of possible physiological mechanisms. This is, I believe, a second reason that Konorski's work was largely ignored, as he himself later recognized. There is a certain irony in the fact that, while he was rejected in the West because of his interest in classical conditioning and physiology, he was under ideological attack in the USSR because of his interest in conditioned reflexes of the second type (instrumental learning) and his rejection of Pavlovian physiology.

Another feature of Konorski's work, which probably contributed to its unpopularity in America, was its concern with processes of inhibition. This was a natural consequence of taking Pavlov's experimental findings seriously; the concept of inhibition is integral to the analysis of classical conditioning, but was anathema to most behaviorists. The objections were, first, that the use of the concept smacked of the worst kind of neurologizing and, secondly, that it was in any case redundant. Skinner spoke for many who would have disagreed with him in other ways when he said (1938): "We do not need the term [inhibition] because we do not need its opposite. Excitation and inhibition refer to what is here seen to be a continuum of degrees of reflex strength and we have no need to designate its two extremes (p. 18)." The general distaste for the concept of inhibition needs no further documentation, but it must have done much to reduce enthusiasm for an author who used the term as freely as a Hullian would talk of reinforcement. Of course, reactive and conditioned inhibition were familiar terms in the context of Hullian theory, but they had always aroused disquiet and were tolerated largely because of Hull's immense prestige; by the early 1950s these concepts were under heavy attack as the weakest part of the Hullian edifice.

Another feature of Konorski's work, which he had borrowed from Pavlov, was his use of single animals in experiments and the total absence of statistics. This, too, was very foreign to the Western approach, which at that time depended almost entirely on group experiments and statistical analysis.

The final aspect that produced withdrawal rather than approach is in some ways the most interesting, because it illustrates most clearly the changes in Western psychology that have occurred in the last 20 years. Throughout his career, Konorski was primarily concerned with the mechanism of association—how particular stimuli, or stimuli and responses, become associated with one another. His early enthusiasm for Pavlov's work and his training in Pavlov's laboratory between 1931 and 1933 ensured that he never made the mistake, so common among Western learning theorists, of assuming that the process of association itself was a relatively simple one. Familiar as he was with Pavlovian concepts of irradiation and concentration of excitation and inhibition, positive and negative induction, and the rest, Konorski was well aware that the study of the mechanism of association in conditioned reflexes of Type I or Type II was no easy matter. In time he became dissatisfied with the Pavlovian account, largely because it did not take the problems of association seriously enough. As a part of a powerful criticism of Pavlovian theory, Konorski wrote (1948):

> The fundamental feature of Pavlov's theory, which sets its mark on it and determines all its statements, is the assumption that not only the process of excitation evoked by the application of an active conditioned stimulus, but also the process of inhibition... is localised in the cortical centre of this stimulus, in the point of the cerebral cortex to which it is "addressed." ... This assumption predetermines that all the fundamental processes which occur... when positive and inhibitory conditioned reflexes are in action... take place at the very beginning of the cortical part of the corresponding reflex arcs, and that one or another conditioned reaction, or the absence of it, is only to be regarded as an indicator which gives notice of the "sign" and intensity of the process. Thus the reflex arc as a whole disappears completely from sight, and attention is concentrated on unspecified states of excitation and inhibition... unspecified because it does not matter in the least which executive neurons they are linked to. In our view this assumption is the "original sin" of all Pavlov's theory [p. 38].

Much of his work before 1955 was devoted to providing a more satisfactory model, couched in physiological terms, of the basic conditions required for establishing excitatory and inhibitory associations in conditioned reflexes of Type I and Type II. In this respect Konorski was representative of the Russian psychophysiological tradition, descending, through Sechenov, from nineteenth-century German physiology. Within this tradition the reflex arc, including "reflexes of the brain," was the fundamental subject matter of physiology and psychology. Indeed, insofar as Pavlov can be seen as having a psychological rather than a physiological theory, Konorski was firmly within the Pavlovian tradition; for example, at no point in this book did he question the essentially Pavlovian idea of classical conditioning as stimulus substitution. By contrast, behaviorism was ultimately rooted in Darwinian theory, with its emphasis on adaptation to the environment. Learning was seen as the major mechanism of adaptation, and learning theorists were

concerned with the way in which trial-and-error learning adapted an animal to a changing environment. Within this context the important issues appeared to be such things as the order of elimination of errors in a maze, place versus response learning, and, above all, the role of reinforcement in learning. Most Western psychologists either were, like Skinner, largely uninterested in the mechanisms of association, or believed that it was a primitive mechanical process, adequately described by writing a big *H* between a little *s* and *r*—thus leaving them free to get on with the important business of deciding, for example, whether drive reduction was necessary for reinforcement. Thus, Konorski was addressing questions which seemed largely irrelevant to American behaviorists.

It is notoriously difficult, as any reader of Konorski must know, to convert an "aversive" stimulus into an "attractive" one; but over the last 20 years this is exactly what has happened to Konorski's ideas vis-a-vis Western psychology. The change has come about as a result of profound alterations in our approach to animal learning; one could almost say that the appetitive and aversive systems have changed roles. It is common for one generation to react against the preconceptions of the preceding one, but, with respect to the assumptions of 25 years ago, the reversal in thinking in the West has been remarkably complete. Classical conditioning is now of central theoretical interest, the change being conveniently marked by the publication of *Classical Conditioning* (Prokasy, 1965). At the same time, following the appearance of Rescorla and Solomon's classic paper (Rescorla & Solomon, 1967), the interactions between classical and instrumental conditioning provide one of the most vigorous areas of research in psychology. Inhibition has also become a respectable concept in the study of both classical and instrumental learning (e.g., Boakes & Halliday, 1972; Rescorla, Chapter 4; Zielinski, Chapter 10 in this volume). Changes in attitude toward physiology are perhaps less marked, and some prejudices linger on; but there are clear signs of a useful dialogue deveoping between physiological psychologists and students of learning—as is shown, for example, in the physiological contributions to this volume (Moore, Chapter 5; Gray, Chapter 11). Finally, as a result of Kamin's work on blocking (e.g., Kamin, 1968) and Wagner and Rescorla's theoretical and experimental work (Wagner & Rescorla, 1972; Rescorla & Wagner, 1972), the study of the formation of associations has moved to the center of the stage as possibly the most important problem in the study of animal learning. I have mentioned these changes only briefly; they are recent and familiar enough to require little comment. What is remarkable is that one looks in vain to American behaviorism of 25 years ago for the intellectual antecedents of contemporary views; whereas Konorski, who comes from a tradition largely independent of Western psychology, speaks clearly and convincingly, over an interval of more than a generation, on the very issues that are now seen as most important.

Other chapters discuss in detail the significance of Konorski's work for modern approaches to a variety of aspects of learning and motivation. Nevertheless it seems worthwhile to enumerate here some of the ways in which his experimental and theoretical work anticipated contemporary developments, in some cases by as much as 30 years.

In 1928 Konorski and Miller published their first experimental papers, one of which was entitled 'Sur une forme particulière des reflexes conditionnels.'' This paper described experiments carried out with a single dog and makes the first clear distinction between classical and instrumental conditioning (or conditioned reflexes of the first and second type, as they call them). They found that if, during the presentation of a conditioned stimulus, the dog's leg was passively flexed on some of the trials and if food was delivered only on those trials, then the movement would begin to appear spontaneously in response to the conditioned stimulus. If, on the other hand, the reflex was reinforced with an aversive stimulus—such as blowing air into the dog's ear—on those trials when its leg was passively flexed, the dog would increasingly extend its leg until it became so rigid that it was almost possible to pick the dog up by that leg. In discussing these findings, they wrote (1928):

> The phenomena that we have just described have the same general properties as conditioned reflexes: they originate without doubt in the cortex, and they are not innate but are formed and disappear during the life of the individual. It is for this reason that we regard them as conditioned reflexes, but their mechanism is different from the conditioned reflexes of Pavlov.... Hence, unable to reduce the reflexes of the second type to the conditioned reflexes of Pavlov, we should consider them as the second fundamental mechanism of the function of the cerebral cortex.... Differences between conditioned reflexes of the first and second type are considerable. (1) In the ordinary conditioned reflex, the conditioned stimulus always elicits the same reaction as the reinforcing stimulus, while in conditioned reflexes of the second type these reactions are different.... (2) The role of Pavlov's conditioned reflex is limited solely to signalisation while the role of conditioned reflexes of the second type is completely different [according to their relationship to the positive or negative reinforcing stimulus]. (3) In conditioned reflexes of the first type, the reaction is effected by organs innervated through the central or autonomic nervous system, while, in conditioned reflexes of the second type, the effector can probably be only a striate muscle [p. 188; 1969, translation by Skinner].

In their next papers they showed that similar results were obtained when the flexion was induced by passive movement, or by mild electric shock, or when the experimenter merely waited for the response to occur spontaneously. In the 1928 paper, without reporting any findings, they also outlined the procedures for avoidance learning and omission training and illustrated their relationship to instrumental reward and punishment training. This is an impressive achievement for a first publication, and much confusion could have been avoided, both in Russia and the West, if the paper had been widely known and discussed.

It is interesting to know something of the circumstances under which this early experimental work was done. Konorski and Miller were medical students who had become disenchanted with the pedestrian character of their studies and had, quite accidentally, come across a copy of Pavlov's works. As Konorski reported (1974):

> From these works we learned for the first time about conditioned reflexes and we immediately realized that this was exactly the field of science we were looking for. The extent of our excitement brought about by this discovery is difficult to describe. We became entirely involved in studying Pavlov, and only by some miracle, not quite clear to me yet, did we succeed in being graduated in medicine [p. 186].

At that time there was little psychology in Poland, and certainly no work on conditioned reflexes; however, they eventually persuaded a professor at the Free Polish University to let them use a small third-floor room in an apartment block, in which part of his department was housed. Konorski recalled (1974), "After this the first thing we had to do was buy a dog. For this purpose we went to the market place, found the area where people sold dogs, and after long deliberation, chose a young and nice bulldog, which cost us ten zlotys (about one dollar). We called him Bobek. He immediately became friendly with us and we brought him to our 'laboratory.' The housekeeper agreed to let him stay in her apartment." The laboratory arrangements were somewhat primitive: "putting together two square stools we made a 'Pavlovian stand' and used cardboard for a screen. . . . Pieces of food were thrown from the small aperture in the screen by the experimenter." For kymograph recording they used toilet paper, "which was both cheap and convenient, provided that it was relatively smooth and did not have transversal perforations. You can imagine the comical picture presented by two serious young men going to a paper store and asking to be shown all the possible varieties of toilet paper, scrutinizing them thoroughly, and choosing the one which fulfilled both conditions (pp. 187–188)." This period of his life could be seen as an almost perfect illustration of the rather hackneyed point that good research arises from good ideas, not from good equipment.

Konorski and Miller were certainly the first to publish experiments using the omission procedure. These experiments follow essentially the same pattern as those already described, except that passive flexion of the leg signaled nonoccurrence of food. Strong extension reflexes developed which were sufficient, on most trials, to keep the dog's paw on the floor against the calibrated pull of the apparatus. In these experiments they were recording salivation as well, and found that it was elicited by the stimulus if the dog was successful in keeping its foot on the floor, but was immediately inhibited if the leg was flexed, signaling the absence of food. In order to demonstrate the importance of the instrumental contingency, they also showed that there was no extension of the leg to a Type I classically

conditioned stimulus that signaled food regardless of the dog's behavior. These early papers provided almost the only example of omission training for some 30 years, and it was not until the publication of Sheffield's paper on omission training in salivary conditioning (1965) that the potentialities of this technique were realized in the West. (For more recent applications, see Hearst, Chapter 2; Mackintosh and Dickinson, Chapter 6; Zielinski, Chapter 10.)

The analysis of avoidance conditioning has, from about 1940, provided both a major problem for learning theorists and an important spur to theoretical development. Understanding of avoidance learning was slow to develop, largely because of confusion about the distinction between instrumental and classical conditioning. Yet in his 1948 book Konorski provided a very clear analysis of avoidance which is in many ways similar to Mowrer's (1960) theory and which is based in experiments carried out in the 1930s or before. The first experiments were, in fact, carried out in 1928 on the admirable Bobek. Mowrer has recently written, "If I had known of Konorski's *Neuron Organization and the Conditioned Reflexes,* I would have had special reasons for citing it (which I did not do) in *Learning Theory and Personality Dynamics*" (Mowrer, 1976). Konorski's analysis of avoidance is based on the following type of experimental procedure (1948):

> The external stimulus (say the metronome) is reinforced by a negative unconditioned stimulus, e.g. the introduction of acid into the dog's mouth. The conditioned reaction consists of the secretion of saliva.... When this reflex is formed, from time to time we provoke a passive flexion of the limb in association with the metronome, and on these trials the acid is not introduced. After a comparatively short time a reflex [of the second type] is formed in which the dog itself raises the limb to the metronome—a movement saving it from the introduction of acid. This reflex is soon established so firmly that for weeks and months the dog goes on raising its paw to each application of the metronome, even though the reflex is never reinforced [p. 228].

He also observed that, once established, the avoidance response would be made spontaneously to other conditioned stimuli that had independently been associated with acid in the mouth, or indeed with other unconditioned stimuli, such as a puff of air in the ear. This type of experiment is in many ways superior to the shuttle box, so familiar to Western psychologists: There is a clear separation of the classically conditioned response and the instrumental avoidance response; the leg flexion is a "pure" avoidance response, uncontaminated by any escape contingency; control of the stimulus situation is far better, since the animal cannot move around in the apparatus; and the development of the classically conditioned response can be monitored during all stages of avoidance learning. Thus, for example, Konorski was able to observe that the first spontaneous avoidance response occurred on the trial immediately following the first trial, during which passive movement of the leg had produced a marked reduction in salivation.

Only when the feedback from leg flexion had become a conditioned inhibitor did the avoidance response emerge; this was in accord with his theoretical position — which was that the reinforcement of the avoidance response was provided by the conditioned inhibition generated by kinaesthetic feedback. He also suggested that the termination of the warning signal might itself act as a conditioned inhibitor of the same sort. He was, however, puzzled at the persistence of avoidance behavior and commented (1948):

> So why is it that the character of this [conditioned] stimulus as a conditioned defensive one is somehow "preserved" when it is accompanied by a movement constituting a conditioned inhibitor? One has the impression that this very movement somehow protects the stimulus against extinction.... But we are unable to explain the causes of this phenomenon [p. 231].

We should recognize that Konorski's 1948 explanation of avoidance was not entirely satisfactory; in particular, as he noted later (1967), "it gave precise rules determining whether and when a given exteroceptive stimulus would elicit a given instrumental response, but it did not answer the question why it would do so (p. 393)." But it is a formulation which avoided misunderstandings and confusions and clearly identified the important questions.

This brief description of Konorski's early work on avoidance illustrates another area in which he anticipated modern experiments: the concurrent measurement of classical and instrumental responses. Such measurements were, of course, a natural consequence of his standard experimental procedures. For example, he showed that when passive movements of the leg were followed by presentation of food, spontaneous leg movements appeared soon after the dog began to salivate to the passive flexion, and that in extinction salivation and leg flexion declined together and disappeared at much the same time. However, he also concluded that it was not possible to make any general rules about the temporal relationship between the classical and instrumental components; sometimes salivation preceded the instrumental response, sometimes it followed it, and on occasion it was absent altogether. He therefore turned to the more powerful superimposition technique, which has become so popular in the West in the last 10 years. A typical example of one of these experiments, which was originally reported in 1936, appears in *Integrative Activity of the Brain* (1967):

> A classical alimentary CR was established in a dog to a bell and a metronome set at 120 beats per minute. A metronome set at 60 beats per minute was differentiated by not reinforcing it by food.... Then... the dog was trained to lift his hindleg in the experimental situation, each movement being reinforced by food. [In the test stage] ... for a few seconds food was not presented and consequently the dog performed the trained movement with maximum frequency. Then either a positive type I CS (bell or M120) or the inhibitory one (M60) was presented.... It was found that in response to

> the positive CS the dog immediately stopped performing the trained movement, looked at the feeder and salivated copiously. On the contrary, in response to the differentiated stimulus, salivation was much reduced but the movements of the leg continued to appear with maximal frequency [pp. 369–371].

This is but one of a large number of similar experiments and is intended merely to illustrate the way Konorski was making good use of the superimposition technique in the early 1930s and the interesting results he was obtaining. (For further discussion of these procedures, see Boakes, Chapter 9; Morgan, Chapter 7.)

Konorski was also the first to understand the special character of conditioned inhibition. Pavlov had regarded conditioned inhibition as but one among the varieties of internal inhibition and certainly did not accord it any special status; he claimed that (1927) "all the experimental evidence ... establishes conclusively that the nervous processes on which conditioned inhibition depends are identical in character with those of extinctive inhibition (p. 86)." The only difference that Pavlov acknowledged was in the procedure for producing inhibition in the two cases. Konorski (1948) was dissatisfied with this argument and, after subjecting the Pavlovian account of the mechanism of conditioned inhibition to a devastating criticism, he proposed his own theory. This is couched in physiological terms, but can be interpreted in purely behavioral ones. In outline he assumes that the process of extinction involves the formation of inhibitory connections between the center of the CS and that of the US, but not the destruction of the existing excitatory connections. Thus after extinction the CS has both excitatory and inhibitory properties, and these are roughly in balance. In the case of conditioned inhibition, however, because there are no excitatory connections between the center for the conditioned inhibitor and the center for the US, all the connections which are formed are inhibitory (1948).

> So in the case of conditioned inhibition we deal with an interesting situation, and one very suitable for analysis, in which the excitatory and inhibitory connexions have their starting point in two different cortical centres: the centre [for the conditioned stimulus] is the starting point for almost exclusively excitatory connections ... while the centre [for the conditioned inhibitor] is almost a pure source of inhibitory connections [p. 155].

Not merely did he identify conditioned inhibition as being in a class of its own because it has almost purely inhibitory effects, but he also recognized that this property makes it a uniquely valuable analytical tool. (See Rescorla, Chapter 4.)

It would be possible to extend this catalogue of experimental and theoretical contributions considerably, but to do so might prove an exercise in overkill; it should already be clear why Konorski's work seems

particularly relevant to contemporary animal psychologists. I would like, however, to draw attention to one other finding that has an especially modern ring. Konorski pointed out that, in what would now be called a free operant situation, the instrumental responses are under the control of ill-specified background stimuli; under these circumstances the background stimuli alone do not predict food, whereas these stimuli plus the movement of the leg do. Therefore, by analogy, if the passive training procedure is to be used with a discriminative stimulus which is to have the same relationship to the response as the background stimuli had in the free operant situation, it is essential that two sorts of trial are given: those where the CS occurs, the leg is flexed and food is presented, and others where the CS occurs without flexion and no reinforcement is given; these latter trials being equivalent to the periods in the free operant situation when the background stimuli were present but the dog was not lifting its leg and was not getting food. This was the training procedure which Miller and Konorski used in their original experiment and which resulted in rapid acquisition of the instrumental response. When they went to work in Leningrad in 1931, Pavlov suggested that this was not a necessary condition for the development of Type II conditioned reflexes and that all that was needed was mere contiguity between the external stimulus and the response. Miller and Konorski therefore did a further experiment in which the metronome was paired with mechanical flexion of the leg and reinforced with food on every trial. Konorski (1948) noted:

> The result we get then is as follows: first, an alimentary conditioned reflex of the first type is formed to the metronome, the dog displaying the secretory and the motor alimentary reaction (turning the head, ... licking the lips, etc.). Secondly, the dog begins to raise the leg actively with greater or lesser frequency during the intervals between the conditioned stimuli.... But the active movements are not performed to the application of the metronome; in fact, rather the reverse is observed.... There is not even a hint of any "connection" between the metronome and the raising of the leg. ... From the foregoing data it follows that in order to form a conditioned reflex of the second type—metronome ⟶ raising the leg—it is not sufficient to have a simple concurrence of these two phenomena and to reinforce them by food. In order to form such a reflex, a situation has to be created in which the movement is an *indispensable* condition of the dog's obtaining food [pp. 219–220].

This analysis, with its recognition of the importance of background stimuli and the distinction between continuity and contingency dates back to the early 1930s; yet it is entirely contemporary in conception. (See Mackintosh & Dickinson, Chapter 6.)

One is naturally led to ask how Konorski saw so clearly things to which most other psychologists of the time appeared blind. Such questions can never be satisfactorily answered, but it is possible to suggest some reasons. Much must be attributable to Konorski's exceptional personal qualities; on

this topic I shall mention only one or two points derived from memoirs by his students and collaborators. His dedication to his work was total; once having found the subject of his choice, after early disappointment with mathematics and medicine, his commitment was lifelong and complete. More than one of his students saw in him the embodiment of Pavlov's maxim that "science requires one's whole life, and even if one had another life to offer it would still not be enough." Two anecdotes from Dr. Fonberg's memoir of him may serve to illustrate his enthusiasm and energy. The first dates from 1958 when an international physiological convention was being held in Poland. A Sunday morning had been set aside for boating on a nearby lake. It was a beautiful day and everyone was setting off in different directions when the word was passed around that Konorski was holding a seminar. When they turned back to investigate, the boatmen discovered that he was holding a floating seminar from his boat in a reedy bay; and so the morning, although spent on the water, was devoted to physiology rather than recreation. The other story concerns Konorski's return from his first trip to America; as Dr. Fonberg recalled the occasion (1974):

> He landed in Warsaw on a Saturday afternoon; most of the staff at the Institute was already preparing to leave for home. Konorski, however, went straight from the airport to his laboratory; he had a cup of tea in his study, washed his hands and then proceeded to organize a meeting. He wanted to tell everyone about the interesting research being carried out in the United States and went on lecturing until 9 o'clock in the evening. After the strain of a trans-Atlantic flight, the change in climate and the time lag, Konorski was still radiating with energy, so fervently did he want to share with us his scientific "adventures" in the USA [pp. 662–663].

In adition to his phenomenal energy, he possessed more than his fair share of determination and self-confidence. This is perhaps seen most clearly in his relationship with Pavlov. At the time he was in Leningrad, Pavlov was the absolute ruler in his laboratory. Konorski (1974) wrote:

> He assigned the problems to be worked out to each of his co-workers and he controlled all stages of their research. Only in exceptional circumstances did a student follow his own lines of research.... In the absence of Pavlov we usually talked about him, quoted what he said, how he behaved, etc. People boasted when Pavlov spoke to them at some length, and... the attitude of Pavlov toward an individual was the main factor determining the hierarchy within the group.... The atmosphere reigning in the group reminded one of that usually encountered in royal courts, with Pavlov being the indisputable king [p. 193].

Add this to Pavlov's notorious aggressiveness in argument, and it can easily be appreciated that it would be a brave man who would oppose him. According to Konorski, Pavlov strongly disagreed with the view that there

were two kinds of conditioning and failed to see any difference between them. However, throughout his two years in Leningrad Konorski carried on his own line of research and braved Pavlov's disapproval, which must have required considerable strength of will for a young and inexperienced scientist. There is, in fact, reason to believe that Pavlov respected Konorski for his willingness to oppose him in argument—even if he disapproved of his scientific views so strongly that when Konorski and Miller published their work in his journal they did not dare use their own terminology of Type I and Type II conditioned reflexes. Konorski's integrity and strength of purpose were once again put to the test in 1949 when he was denounced as a physiological deviationist, and once again he refused to give way. However, his firmness seems never to have degenerated into dogmatism; to the end of his life he was prepared to entertain new ideas, and indeed there was a very marked change in his theoretical outlook between 1948 and 1967. In the closing passage of his autobiography, he wrote that "the full cycle of renewal of my scientific ideas requires about two decades" and went on to regret that there would not be time for another cycle.

Konorski must also have been an exceptionally gifted experimenter. His original experiments with Bobek in the ramshackle laboratory in the Warsaw apartment block have stood the test of time far better than hundreds of experiments carried out in well-equipped laboratories elsewhere. He clearly had an exceptional flair for the right technique and a sensitivity to the relevant experimental details.

Finally, Konorski had the considerable advantage of not being directly involved in either the Pavlovian or the behaviorist tradition. Had he been a normal student of Pavlov's, he might never have been allowed to start on his unconventional line of research; or had there been a behaviorist tradition in Poland, he might have ignored the merits of Pavlov's work. As it was, he was able to develop his experimental techniques and interpret the results untrammelled by conventions and preconceptions; and this led him in directions which seem, at present, to have been remarkably farsighted.

This discussion of Konorski's work and its reception in the West has dealt almost entirely with the period before 1955; this is because, paradoxically, his earlier work is closer to the concerns of modern Western animal psychology than his more recent thinking. *Integrative Activity of the Brain* may be seen as a development and revision of his earlier views, but it is much more loosely structured and ambitious than his earlier book. It is an attempt to provide a model—couched in terms of a "conceptual nervous system"—not merely for learning but also for perception and motivation. In his book Konorski introduced a range of new concepts; gnostic units, preparatory and consummatory conditioned reflexes, and the idea of mutually inhibitory preservative and protective (appetitive and aversive) motivational systems. Unlike his previous book, *Integrative Activity of the*

Brain was widely noticed in the psychological journals; and the reviews, although varying in enthusiasm, were all respectful and regarded the book as an important contribution. However, the general tenor of these reviews was that the book contained intriguing insights and interesting experimental results, rather than a satisfactory overall framework for understanding the brain and behavior. In general it seems that the reviewers' assessment has been confirmed over the last 10 years, but of course it was some 20 years before the importance of Konorski's earlier work was fully recognized. Konorski's (1974) own reaction is both interesting and appropriate:

> Contrary to my expectations the reaction to the book was rather poor. I had a feeling that many of my friends and colleagues simply disliked it, or had not read it....
>
> I am very curious to know what will be the final fate of the book; will it eventually win general recognition, which I think it deserves in spite of its shortcomings, or will it have no important impact on the further development of behavioural science? [p. 214].

In the light of his previous book, it would be a rash man who would firmly predict the latter outcome.

ACKNOWLEDGMENTS

I am grateful to R.A. Boakes, A. Dickinson, N.J. Mackintosh, and S. Soltysik who read the draft of this paper and made a number of very helpful suggestions.

REFERENCES

Boakes, R.A., & Halliday, M. S. (Eds.), *Inhibition and Learning.* London: Academic Press, 1972.

Estes, W. K., Koch, S., MacCorquodale, K., Meehl, P. E., Mueller, C. G., Jr., Schoenfeld, W. N., & Verplanck, W. S. *Modern learning theory: A critical analysis.* New York: Appleton-Century-Crofts, 1954.

Fonberg, E. Professor Jerzy Konorski. *Acta Neurobiologiae Experimentalis,* 1974, *34,* 655–664.

Hilgard, E. R. *Theories of learning.* New York: Appleton-Century-Crofts, 1948.

Hilgard, E. R. *Theories of learning* (2nd ed.) New York: Appleton-Century-Crofts, 1956.

Hilgard, E. R., & Marquis, D. G. *Conditioning and learning.* New York: Appleton-Century-Crofts, 1940.

Kamin, L. J. "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation.* Miami: University of Miami Press, 1968.

Konorski, J. *Conditioned reflexes and neuron organization.* Cambridge: Cambridge University Press, 1948.

Konorski, J. *Integrative activity of the brain: An interdisciplinary approach.* Chicago: University of Chicago Press, 1967.

Konorski, J. Jerzy Konorski. In G. Lindzey (Ed.), *A history of psychology in autobiography* (Vol. 6). New Jersey: Prentice-Hall, 1974.

Konorski J., & Miller, S. On two types of conditioned reflexes. *Journal of General Psychology,* 1937, *16,* 264–272. (a)

Konorski, J., & Miller, S. Further remarks on two types of conditioned reflexes. *Journal of General Psychology,* 1937, *17,* 405–407. (b)

Miller, S., & Konorski, J. Sur une forme particuliere des reflexes conditionnels. *Les Comptes Rendus des Seances de la Société de Biologie,* Société Polonaise de Biologie 1928, *99,* 1155–1157.

Miller, S., & Konorski, J. On a particular form of the conditioned reflex. (Originally published 1928. English translation by B. F. Skinner.) *Journal of the Experimental Analysis of Behavior,* 1969, *12,* 189–189.

Mowrer, O. H. *Learning theory and behavior.* New York: Wiley, 1960.

Mowrer, O. H. How does the mind work? *American Psychologist,* 1976, *29,* 843–857.

Osgood, C. E. *Method and theory in experimental psychology.* New York: Oxford University Press, 1953.

Pavlov, I. P. *Conditioned reflexes.* Oxford: Oxford University Press, 1927.

Prokasy, W. F. (Ed.), *Classical conditioning: A symposium.* New York: Appleton-Century-Crofts, 1965.

Rescorla, R. A., & Solomon, R. L. Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological Review,* 1967, *74,* 151–182.

Rescorla, R. A. & Wagner, A. R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory.* New York: Appleton-Century-Crofts, 1972.

Sheffield, F. D. Relation between classical and instrumental learning. In W. F. Prokasy (Ed.), *Classical conditioning: A symposium.* New York: Appleton-Century-Crofts, 1965.

Skinner, B. F. *The behavior of organisms.* New York: Appleton-Century-Crofts, 1938.

Wagner, A. R., & Rescorla, R. A. Inhibition in Pavlovian conditioning: Application of a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning.* London: Academic Press, 1972.

Woodworth, R. S., & Schlosberg, H. *Experimental psychology.* New York: Holt, 1954.

2 Classical Conditioning as the Formation of Interstimulus Associations: Stimulus Substitution, Parasitic Reinforcement, and Autoshaping

Eliot Hearst
Indiana University

I. INTRODUCTION

Jerzy Konorski is best known to Western experimental psychologists for his early articles proposing and analyzing the distinction between classical (Type I) and instrumental (Type II) conditioning (Miller & Konorski, 1928; Konorski & Miller, 1937a, 1937b). His contributions to this topic influenced the expression of Skinner's specific views on respondent and operant conditioning (Skinner, 1937, 1938), and in the long run probably also played a significant role in the development of the more intricate "two-process learning theories" that were subsequently advanced and became extremely popular. (See Mowrer, 1976, and Rescorla & Solomon, 1967, for summaries and comments.) However, Konorski's general theoretical system, the ways in which he modified it over 20 years, and much of the research that he performed in conjunction with his colleagues at the Nencki Institute received relatively little attention in the West during his lifetime. This neglect may be partially due to the comparative inaccessibility of most of the journals in which he presented his ideas and findings, but is more likely attributable to the neurophysiological emphasis and speculative nature of his writings—features that would inevitably have prejudiced a large number of behaviorists against careful examination of his work. Halliday (Chapter 1) develops some of these themes in attempting to explain why Konorski's contributions were largely disregarded for many years.

Konorski died at a time when many Western psychologists were becoming disenchanted with extremely behavioristic, response-centered approaches to the analysis of learning. He probably would have been pleased

with the various recent attempts to discuss learning and memory in terms that go well beyond simple relationships among observables. Nowadays workers in the field of animal and human learning seem less reluctant to speculate about potential intervening processes or stages that may help to explain how an objective, environmental stimulus produces some observable response. Expectancies, spatial maps, reinforcer representations, short- and long-term memories, internal clocks, retrieval processes, channel capacities, and perceptions of causal relations have in the last few years been seriously proposed as mechanisms underlying overt behavior, even in species as unpretentious as the rat and pigeon (e.g., Hulse, Fowler, & Honig, 1978). Many workers find that the postulation of such processes or mechanisms has considerable heuristic value, although they recognize the dangers of any move in a mentalistic direction. Therefore, the Zeitgeist is now more receptive to approaches like Konorski's, and increased attention to various aspects of his work (e.g., the conceptualization and measurement of inhibition: see Boakes & Halliday, 1972; Gray, Chapter 11; Rescorla, Chapter 4; appetitive–aversive interactions: see Dickinson & Dearing, Chapter 8; constraints on learning: see LoLordo, Chapter 14; Shettleworth, Chapter 15) was already discernible in the West at the time of his death.

In this paper I will try to briefly summarize Konorski's views on classical conditioning, compare them with other significant alternatives, and evaluate them in the context of some contemporary research on the topic. His views on mechanisms of reinforcement and the basic connections formed during classical conditioning are stressed here, as well as possible systematic and empirical implications of his distinction between preparatory and consummatory reflexes. Consideration of these issues leads us to reexamine Pavlov's concept of stimulus substitution and Konorski's notion of "parasitic" reinforcement, particularly as applied to recent research concerning the classical conditioning of directed movements (Hearst & Jenkins, 1974; Schwartz & Gamzu, 1977). The findings challenge the clarity and validity of the distinction between classical and instrumental conditioning, just as some of Konorski's own final writings (e.g., 1969, 1973) seriously questioned the sharpness of this distinction, which had formed the basis of his first major scientific contribution.

II. KONORSKI ON CLASSICAL CONDITIONING

Like Hebb (1949), Konorski distinguished between two general kinds of connections that can be formed within the nervous system. One type involves the integration of sensory input and underlies what other writers have called *perceptual learning;* in Konorski's framework such learning entails actualization of potential linkages between units of lower-level (receptive) analyzers and units within the higher level (perceptive or gnostic) areas of

the brain. Gnostic units developed in this manner represent stimulus objects with which the subject has become acquainted; for example, as a new acquaintance's face becomes familiar, it is recognized as a unitary percept. The second type of connection provides the basis for *associative learning* and involves the development of linkages between different gnostic units as a result of their synchronous activation. Thus the sight of a familiar person's face may often evoke an image of that person's voice, because the appropriate visual and auditory gnostic units have been simultaneously activated in the past.

According to Konorski, classical or Type I conditioning provides an especially useful arrangement for studying the formation and maintenance of this second type of connection—that is, an association between perceptive units. The laws of classical conditioning do not differ basically from those governing the acquisition of other interperceptive connections, but the Type I procedure possesses great experimental and analytical advantages because the US reliably evokes a definite, observable response. As a result of contiguity between transmittent stimuli (CSs) and the recipient stimulus (the US), the former eventually come to elicit an overt response that (barring certain practical complications, to be discussed below with respect to "stimulus substitution") is the same as the response consistently elicited by the latter, biologically important stimulus. Measurement of this response to the CS permits us to study the acquisition and persistence of the S–S association in an objective way; the CR acts as a "tracer" enabling detection of the status of the connection.

Drive and reinforcement are terms used frequently by Konorski, but he relied much more on drive as an explanatory concept. The facilitating influence of a certain drive or combination of drives is necessary for associative learning to proceed, but, as we will see, Konorski noted differences between the actions and effects of appetitive and aversive drives that have significant implications for his theory. Concerning the concept of reinforcement, Konorski stated that he would use the term only in an operational, nonphysiological sense, to refer to a stimulus consequence that establishes and sustains a CR to the conditioned stimulus.

In 1948 Konorski discussed the behavioral effects of classical conditioning in terms of the acquisition and maintenance of what he later referred to as consummatory responses. The final version of his complete theory (1967) treats classical conditioning as involving the formation of and interaction between two types of CRs, preparatory and consummatory, which are not necessarily conditioned to the same CSs. Preparatory reflexes tend to direct subjects toward attractive stimuli and away from aversive stimuli, and presumably depend on central drive states like hunger or fear. External or humoral stimuli may activate such drives, whose centers are said to be located in the hypothalamus and limbic system; such activation produces a relatively nonspecific excitation of neural units. The behavioral effects of a

drive include the facilitation of attention to impinging stimuli and the excitation of general motor activity.

Consider a hungry dog fed for the first time in an experimental room. Because of the simultaneous activation of the gnostic units corresponding to the experimental context and the gnostic units corresponding to the drive state, the room itself rapidly comes to evoke what Konorski called the hunger CR, which reflects a conditioned motivational state and is indexed by motor restlessness and searching behavior, as well as by heightened sensitivity to a variety of stimuli, particularly those appropriate to the prevailing drive — gustatory or olfactory stimuli in this case. Drive (preparatory) CRs comprise relatively diffuse behaviors, often not easily measurable; these responses are typically tonic or prolonged and are most likely to be observed in some general context or in the presence of comparatively long-lasting CSs. In human beings hunger CRs are mainly determined by the amount of time since the last meal, and not by very specific stimuli like the sight of food on the dinner table.

Through its arousal of afferent and efferent mechanisms, the conditioning of a drive CR to the experimental setting serves as a necessary precursor to the conditioning of specific consummatory CRs to CS. These latter CRs are the ones that experimental psychologists have normally focused upon and measured during an experiment in classical conditioning: salivation, leg flexions, eye blinks. Such responses are relatively discrete and phasic; and, besides being originally evoked by quite specific stimuli (USs), they are themselves most easily attached to intermittent, brief stimuli. Thus human beings begin to salivate as their favorite food is placed in front of them, but do not constantly salivate while they are on their way to the restaurant where it is to be served. According to Konorski, consummatory CRs take longer to condition than preparatory CRs and are less resistant to extinction. The classical conditioning of consummatory responses is presumed to be mediated by centers in the thalamocortical system, and involves connections between gnostic units representing the explicit CS and the effective (proximal) stimuli constituting the US — for example, gustatory for food, somatic pin-prick for shock. In contrast to the restlessness and activity that characterize the hunger CR and initially occur also during the CS, the motoric effects produced by a well-established brief CS for food are said by Konorski to involve cessation of movement and directed attention toward the US delivery site.

Interactions between drive and consummatory reflexes, both conditioned and unconditioned, are important for Konorski's general analysis of classical conditioning. In this regard there are basic differences between appetitive and aversive procedures: Although appetitive- and aversive-drive CRs facilitate their related consummatory reflexes (e.g., hunger excites salivary CRs, and fear generally augments leg-flexion responses to shock),

in the case of appetitive conditioning the occurrence of the consummatory reflex temporarily inhibits the relevant drive, whereas in the case of aversive conditioning the delivery of the US has no such inhibitory effect. Thus food in the mouth inhibits hunger, as evidenced by the temporary cessation of stomach contractions and by decreased general motor activity, whereas the delivery of shock does not decrease fear.

According to Konorski, the drive inhibition that occurs during consumption of an appetitive US eventually becomes conditioned to CSs immediately preceding US. This process accounts for the reduced activity and "calmness" that are often observed during CSs for food, and also plays a major role in the gradual impairment of alimentary CR performance that Pavlov (1927) frequently noticed, referred to as "extinction or inhibition with reinforcement (pp. 234ff.)" and regarded as dependent on some inevitable functional exhaustion of active cortical elements, which reinforcement delayed but did not prevent. Attribution of these effects to conditioned drive inhibition helped Konorski handle the fact that introduction of a new set of CSs, or interspersed presentation of CS−s that are similar to CS+, often decreases or prevents such performance decrements during CS+. The less certain the subject is with respect to exactly when or whether the US will be delivered, the more strongly is the relevant drive maintained; unexpected nonreinforcement is presumed to evoke the relevant drive and facilitate performance on later trials (cf. Wagner, Chapter 3). Konorski seemed to predict that in many instances partial reinforcement should maintain preparatory and consummatory CRs better than continuous reinforcement, particularly in appetitive situations—a prediction not completely borne out by experimental findings (see Mackintosh, 1974).

Thus Konorski viewed classical conditioning in a connectionistic manner, as a process requiring the synchronous action of two stimuli and the presence of an appropriate drive. An important corollary is that associations are formed not only between the US and contextual or intermittent external stimuli, as stressed earlier and in standard analyses of classical conditioning, but also between the contextual and intermittent stimuli themselves and between drives and the external stimuli. Of course it is the central representations of these events (gnostic units) that are actually connected and mediate the development of CRs; Konorski (1967, p.267) presented several arguments against the possibility that classical conditioning involves establishment of direct connections between CS representations and efferent centers. For him, the CS produces either an image or hallucination of the US, which then brings about the CR. Sensory preconditioning is also treated as a form of S–S learning, governed by the same basic mechanisms as classical conditioning; and Konorski speculated that one drive underlying the formation of such associations between neutral stimuli may involve arousal produced by the stimuli themselves ("curiosity drive").

Although Konorski assumes a definite S–S position concerning classical conditioning, he goes beyond most earlier approaches of this kind by offering rather extensive and yet flexible predictions about the types of behavior that should appear and interact in the classical-conditioning situation. (He cannot easily be accused of leaving his animals "buried in thought," a well-known taunt directed at one S–S theorist of the past.) Konorski (1967) described specific preparatory and consummatory reflexes evoked by various drives and USs, and listed several factors (e.g., CS duration, intertrial interval, US magnitude) that should favor domination of one or the other general type of reflex in the presence of situational or intermittently presented cues. For example, because long CS–US intervals are used, Konorski believed that the CER paradigm (perhaps the most common procedure employed by Westerners who study Pavlovian conditioning) involves mainly preparatory (fear) CRs manifested by freezing, crouching, defecation, and so on, rather than consummatory (shock) CRs like limb flexion. It is certainly true that during a CER stimulus lasting one or two minutes, subjects do not display much flinching or jumping; rather, they ordinarily show the more diffuse or tonic forms of behavior that Konorski characterized as preparatory. Along these lines, Konorski obviously expected the optimal interstimulus interval (ISI) for establishment of Type I CRs to differ depending upon the response measure—an expectation that is solidly confirmed by experimental findings (Mackintosh, 1974).

Because, by Konorski's definition, classical conditioning involves the transfer to CS of only those behavioral effects originally elicited by US, it is clear that he conceived classical conditioning in a way very similar to the view implied by Pavlov's concept of stimulus substitution. However, Konorski (1967, pp. 268–270) suggested several reasons why CRs may occasionally not match URs. One type of exception to this rule is explained by the idea of "parasitic" instrumental reinforcement. For example, food that is delivered immediately after the subject has looked at or moved toward the CS could serve to instrumentally condition such movements; these responses may become quite strong because of their accidental coincidence with subsequent food, even though they are not originally responses to food and are, in fact, antagonistic to the motor behavior directed toward the location of food. Konorski's concept of parasitic reinforcement is, for all practical purposes, identical with the notion of "superstitious" operant conditioning employed by Skinnerians, and later in this chapter I will attempt to discuss the relevance of such concepts for our understanding of some recent findings on the Pavlovian conditioning of directed movements.

The variables that Konorski stressed in his analysis of classical conditioning are not very different from the factors that most other theorists would emphasize: CS intensity, US magnitude, drive level, spatial contiguity of CS

and US, and CS–US asynchrony. Besides these standard parameters of classical conditioning, he also discussed the facilitating effects on CR performance of CS intermittency and "roughness". Nonmonotonous CSs (e.g., of irregularly changing intensity) produce stronger CRs. This is an interesting type of manipulation, which has been neglected in Western work on Pavlovian and instrumental conditioning (but see Pavlov, 1927, Chapter 14). Konorski related such effects to physiological adaptation; their possible relevance to the effects of stereotypy of stimulation on reflex habituation and sensitization, and to the operant–respondent distinction, has been stressed by others (see Razran, 1971).

Konorski presented an extensive discussion of theory and data concerning the transformation of excitatory CRs into "inhibitory" CRs and of appetitive CRs into aversive CRs. In such cases, the old connections are presumed to remain intact and to influence the formation, maintenance, and memory of the new connections. Therefore, if a CS+ is converted into a CS−, the old excitatory linkage between the CS and US centers survives and exerts an antagonistic influence on learning and performance appropriate to the new (excitatory) linkage between CS and no-US centers. Transformations of appetitive CS+s into aversive CS+s, and vice-versa, generally prove very difficult to achieve because of the strong mutual antagonism between protective and preservative drives. However, conversion of an appetitive CS− into an aversive CS+, or an aversive CS− into an appetitive CS+, is relatively easy because antidrives in one category are not antagonistic to drives in the other category. (See Dickinson & Dearing, Chapter 8, for an extensive discussion of appetitive–aversive interactions.)

To sum up, Konorski offered an interpretation of classical conditioning that relied heavily on S–S contiguity and drive. His approach was not really response-centered, but because he included both preparatory and consummatory reflexes within his framework, it was possible for him to make relatively flexible predictions about the forms of overt behavior that should initially appear and subsequently persist or decrease during various kinds of Type I conditioning. In his writings he anticipated many issues that are of great contemporary interest for students of classical conditioning: the control exerted by contextual cues, the measurement and significance of inhibitory learning, the effects of certainty and of unexpected events, the complications created by possible parasitic or "superstitious" reinforcement of responses, the rebound motivational effects that may follow stimulus offsets, the antagonism existing between certain drive states, the experimental operations that may alter already-established central representations of external stimuli, and the value of treating the "CR" as an interacting set of functionally related behaviors rather than as a single reflex. These are topics that many contributors to the present volume address.

III. ALTERNATIVE THEORETICAL VIEWS

Konorski's treatment of classical conditioning differs from several influential Western approaches to the topic. One general difference involves his view that such conditioning is merely a special case of the learning of interperceptive (S–S) associations and that the overt behaviors that develop during classical conditioning reflect the associative process rather than constitute basic elements of it. This perception-centered interpretation contrasts with response-centered treatments following the Thorndikian tradition, in which associations between the CS and particular *responses* are said to develop during classical conditioning, either because of contiguity alone or as a result of contiguity plus drive reduction (reinforcement). In the 1930s and 1940s theorists took much more definite or extreme positions than today concerning the relative importance of S–S versus S–R associations in conditioning. Nevertheless, similar issues, sometimes rephrased or disguised, are still with us in the 1970s as a source of debate and experimentation. (See Hearst, 1978; Rescorla, 1978.)

In addition to his belief that the basic elements of Type I associative learning involve only stimulus representations, Konorski also differs from other influential writers in terms of the conditions required for establishment of a Type I CR. For him, not only contiguity between CS and US is essential, but also some background drive. However, other workers (e.g., Rescorla, 1967) would conceive "contiguity" as acting in a somewhat more complex manner than Konorski explicitly described; and still other authors would either scrupulously avoid a separate drive concept or would argue that classical conditioning can proceed in a motivationally neutral setting. With respect to the concept of "reinforcement," Konorski (1967) asserted that he employed the term in a nontheoretical sense; but of course various other theorists would insist that some strengthening event, corresponding to Thorndike's "satisfying state of affairs," must follow a CR in order for it to be acquired and maintained. Certain of these latter theorists believe that mere conjunctions of the CR and reinforcer are sufficient for learning, whereas other writers presume that the CR gains strength because it actually affects the value of the reinforcer. And along different lines, which echo one aspect of the classic exchange between Guthrie and Pavlov in the 1930s (Guthrie, 1930, 1934; Pavlov, 1932), many Western learning theorists would strongly object to Konorski's emphasis on physiological speculation and reductionism, regardless of the behavioral research and psychological theory he proposed. A persistent East–West difference in the treatment of learning and behavior concerns the view, much more prevalent in the East, that behavioral research must go hand in hand with neurophysiological research if work on conditioned reflexes is to be truly valuable.

This section of the present chapter briefly examines several differences between Konorski's views and those of some other writers, expecially with

respect to basic associative elements, contiguity, and reinforcement. More complete summaries and empirical evaluations of related issues can be found in Mackintosh (1974).

A. Interperceptive Versus Stimulus-Response Associations

Konorski followed the Pavlovian tradition in his stress on classical conditioning as a process that basically involves the formation of connections between neural centers representing the CS and US. This view is not very different in spirit from the interpretation offered by so-called S–S theorists like Tolman (1932), Bolles (1972), and Bindra (1972), nor from the cognitive interpretation of Pavlovian conditioning recently outlined by Rescorla (1978). All these writers make a distinction between learning and performance and would presumably view the specific CRs that appear during classical conditioning as indices of the establishment of connections between stimulus events — for example, learning the "causal texture of the environment" (Tolman & Brunswik, 1935) — rather than as fundamental or primary units in the associative process. In other words, we are fortunate that certain stimuli happen to evoke various unconditioned responses, for this outcome permits us to study CS–US associations in an objective and reliable way; in classical conditioning it is not the responses themselves that are being learned, for they are wired into the organism from the very outset.

Alternative points of view take issue with such a deemphasis of the response and are generally much more peripheralistic. For example, Guthrie (1935) argued that classical conditioning occurs merely because the procedure ensures that the UR is elicited shortly after the CS is presented. Such close contiguity of a stimulus and a response results in the formation of an association between them so that future presentations of the CS will evoke a response closely resembling the UR — provided that the CS is reproduced fairly precisely on subsequent trials and that no other strongly interfering behaviors are evoked during the CS or US. Although Hull (1943)agreed with Guthrie that the basic associations in classical conditioning are S–R in nature and that contiguity of CS and UR was necessary for conditioning, he believed that the US must, in addition, involve reduction of some drive in order for successful conditioning to occur. Skinner (1938, 1953) cannot be categorized as an S–R theorist in the traditional sense of the label, but his discussions of classical conditioning are certainly response-centered. He accepts the validity of the concept of stimulus substitution, but he maintains that reflex responses constitute only a small part of the total behavior of an organism. Thus Pavlovian conditioning, according to Skinner (1953), can add "new controlling stimuli, but not new responses (p. 56)." This opinion provides one of the reasons for the overwhelming emphasis on operant conditioning in his writings and experimental research; he has been particularly

interested in using reinforcement to "shape" complex behavior patterns from available but simpler responses.

A variety of studies have attempted to pit S–S and S–R explanations of Pavlovian conditioning against each other, but without definite resolution of the issue. Blockage of the UR during CS–US pairings by means of surgical procedures or the injection of drugs does not prevent conditioning; when the blocked response is again physically possible, it is evoked by the CS without further training. Sensory-preconditioning experiments are often successful, even though they involve the pairing of two stimuli neither of which evokes a definite overt response. And an appeal to the conditioning of specific peripheral responses cannot explain the effects of Type I CSs on instrumental responding (see Rescorla & Solomon, 1967), because for example, a shock signal will reduce VI food-rewarded behavior and generally increase Sidman avoidance behavior.

Even though these results seem to support an S–S approach over conventional S–R interpretations, which are very peripheralistic, the findings might be handled by less extreme forms of S–R theory. The definition of a "response" could be broadened to include central nervous system activity or motivational states, or it might be argued that subtle responses were conditioned that the experimenter failed to record or was technically unable to measure. Consequently, many workers feel that the S–S versus S–R controversy boils down to a semantic one related to the definition of a response. Although this conclusion has some force, it provides no excuse for the experimental neglect of such phenomena as latent learning, sensory preconditioning, place learning, and so on, which were closely linked historically with the S–S versus S–R battle. And in this volume Rescorla (Chapter 4) describes recent research that attempts to analyze aspects of associative learning in terms of response-centered versus "representational" explanations; his findings are important whether or not they are cast in the framework of S–S versus S–R interpretations.

B. Contiguity Versus Informativeness

According to Konorski, associative learning occurs when two gnostic units linked by a potential connection are synchronously activated; optimal conditioning requires that a US closely follow the onset of CS. There is virtually no one, regardless of his theoretical persuasion, who would disagree with the general notion that temporal contiguity promotes classical conditioning, so long as the CS precedes the US. However, recent empirical and theoretical analyses have suggested that conceptions of classical conditioning that consider only pairings of CS and US are misleading or incomplete. In addition to being paired with the US, the CS apparently must also be informative with regard to deliveries of the US; pairings of CS and US are necessary but not sufficient for excitatory learning. The use of concepts like

predictability (Kamin, 1969), *contingency* (Rescorla, 1967), and *validity* (Wagner, 1969) represents the attempts of different workers to encompass the fact that an analysis in terms of mere conjunctions of CS and US is insufficient to explain a variety of recent findings in Type I conditioning.

Experiments on "blocking" (e.g., Kamin, 1969) provide one illustration of the need for a theory of classical conditioning that goes beyond simple contiguity. For example, after a noise (CS) has been paired several times with shock (US), a light may be presented simultaneously with the noise while the noise–shock pairings are continued. Subsequent tests with only the light usually show that it has developed little, if any, power as a CS even though it had been paired with shock often enough for powerful CER conditioning to have appeared if the prior noise-alone trials had not been given. One popular interpretation of this finding is that the light, being redundant, supplied no new information to the subject and therefore conditioning did not occur.

Rescorla (e.g., 1972) performed a series of studies that also illustrate the deficiency of theoretical accounts of classical conditioning that rely excessively on contiguity of CS and US. Even though CS and US were frequently paired in Rescorla's research, the delivery of ("extra") USs in the absence of CS, at the same rate as during CS, resulted in little evidence of conditioning to the CS. These experiments demonstrated that the relative frequencies of the US in the presence versus absence of CS, rather than the sheer number of CS–US pairings, determine the strength of conditioning to a CS. Once again, if a formerly neutral stimulus supplies no information about occurrences of the US, the stimulus will apparently acquire little excitatory power even though it occurs frequently in conjunction with the US.

In his initial treatment of such results, Rescorla (1967) offered a molar interpretation based on the concept of "correlation" or "contingency"; subjects were presumed to possess the sophistication needed to calculate over-all US probability and compare it with US probability during CS. A few years later he and Wagner (Rescorla & Wagner, 1972) proposed a theory that handled a wide variety of these "informational" effects, but on the basis of a more molecular, trial-by-trial model that clearly asserted the importance of temporal contiguity. Their theory is so well known that there is no point in describing it further in this chapter. However, with respect to Konorski's approach, the Rescorla–Wagner model is not so much inconsistent with Konorski's views as it is a much more detailed and precise exposition of how contiguity acts to determine the amount of associative strength accruing to different stimulus elements — including contextual or background cues — during classical conditioning. Actually, when talking about situational cues, unexpected events, and excitatory–inhibitory interactions in conditioning, Konorski anticipated in an informal way some of the important features of the new model; moreover, Halliday (Chapter 1) comments on Konorski's apparent realization of the distinction between con-

tiguity and contingency. Incidentally, Konorski's earlier (1948) approach to the analysis of classical excitatory and inhibitory conditioning seems to come closer to making specific predictions of the kind offered by Rescorla and Wagner than does the less formal, more qualitative approach proposed by Konorski in 1967. (Rescorla, Chapter 4, discusses related issues in more detail than is possible here.)

C. Mechanisms of Reinforcement

Konorski (e.g., 1967, p. 266) considered the term "reinforcement" imprecise and misleading, partly because it was used differently by Pavlov and Hull. However, for purposes of convenience and readability, he decided to employ it when the presentation of a stimulus, in either the classical or instrumental procedure, leads to the establishment of a conditioned response to the CS and sustains this response after it has been established. He stated that he would use the word reinforcement in a purely operational sense, implying nothing about the physiological mechanisms of associative learning.

One can debate whether Konorski adhered to a strictly operational use of the term throughout his 1967 book. However, a more important question involves his views concerning the function of the US in classical and instrumental conditioning and concerning the relationship between the US and the particular responses that emerge from these procedures. For instrumental conditioning he viewed the US as supplying the drive-reduction mechanism that accounts for learning and performance of the successful instrumental movement, which closely resembles a Law of Effect interpretation (although his concomitant use of the principle of retroactive inhibition has Guthrian implications; see, e.g., his summary on pp. 513–514). In classical conditioning, on the other hand, his approach seems to follow Pavlov's ideas on stimulus substitution quite closely. According to Pavlov's interpretation, pairings of CS and US establish the CS as a substitute for US; evidence for this interpretation is taken to be the CS's acquired ability to evoke the behavioral effects elicited by the US. This approach does not imply that the CR is actually strengthened by its consequences; the US does not act to "reinforce" prior responses, but to produce a set of responses (URs) that can then be transferred to its substitute, the CS.

I will return to the topic of stimulus substitution in my discussion of some contemporary research on autoshaping in the next section of this chapter. However, to foreshadow these later comments, it seems worth mentioning here the basic aspects of alternative conceptions of classical conditioning that view the US as serving the function of a response strengthener. These approaches (e.g., Hull, 1943) have some difficulty handling the close resemblance between CR and UR that is said to characterize classical conditioning, but we will see shortly that identity of

CR and UR may not be so very common in Type I conditioning anyway. In one version of this general response-strengthening interpretation—a version that I will call the "response–reinforcer contiguity" explanation — any response occurring during the CS and closely followed by a drive-reducing event like food delivery or pain termination will be strengthened. Because it evokes specific URs, the US increases the likelihood that responses identical with or resembling URs will occur in the general situation and will become the behaviors strengthened in this manner; however, in theory at least, virtually any response that happens to appear during CS would also increase in probability due to its mere conjunction with US. This interpretation seems very similar to Konorski's description of the establishment of parasitic instrumental responses and to Skinner's (1948) analysis of superstitious reinforcement, for in both cases responses are strengthened even though there is no contingency of any kind between behavior and US delivery; the behaviors are not presumed to actually affect the delivery or value of the US in any way.

In the other version of this response-strengthening view of US function, which I will call the "response–reinforcer contingency" explanation, the CR is definitely assumed to modify the effectiveness of the US (e.g., Perkins, 1968; see Gormezano & Kehoe, 1975, for a critical discussion). Salivation may increase the palatability of food, making it more attractive; and freezing or flinching may reduce the pain produced by a shock, making it less aversive. Although in many cases the exact nature of the contingency between the CR and subsequent changes in US value is not easily specified by a theorist or assessed by an experimenter, the CR–US relationship presumably operates as it would during standard instrumental conditioning: Those responses are strengthened that bring about a more favorable consequence than would occur in their absence. So far as I can tell, Konorski did not specifically assess this "instrumental" explanation of classical conditioning, but the possibility has arisen frequently in discussions of the merit of the classical–instrumental distinction (see Hearst, 1975a).

Of course it is quite conceivable and very likely that these various mechanisms — stimulus substitution, response-reinforcer contiguity effects, and response-reinforcer contingency effects—are all operating in a given example of so-called Type I conditioning. Furthermore, one of the factors may be relatively more important in the initial stages of training (e.g., in producing the first few CRs in the presence of the CS), whereas the other factors may serve mostly to maintain and strengthen the CR once it has been established (cf. Champion, 1969).

We turn now to some recent findings in the field of classical conditioning that bear on these and related issues. The arrangements that I stress involve the measurement of directed actions of the whole organism (autoshaping and sign tracking). A brief review of the most important findings concern-

ing this topic precedes a discussion of their relevance to Konorski's theory and to other interpretations of classical conditioning.

IV. DIRECTED MOVEMENTS AND S-S ASSOCIATIONS

A. The Phenomena of Autoshaping and Sign Tracking

The discovery of autoshaping by Brown and Jenkins in 1968 served as the stimulus for a large number of North American experiments investigating the effects of Pavlovian conditioning procedures on directed actions of the whole organism. (See reviews and evaluations of this work by Hearst & Jenkins, 1974, and Schwartz & Gamzu, 1977.) The general topic has attracted considerable attention for several reasons. First of all, evidence demonstrating the classical conditioning of complex skeletal responses—for example, orientation movements, approach-withdrawal behavior, and signal-contact responses—surprised many psychologists who viewed classical conditioning as limited to autonomic responses or to relatively isolated, undirected "local" movements like eye blinks, leg flexions, or knee jerks. Therefore, the new findings had clear implications for the general validity of the classical-instrumental distinction, based as it often has been on presumed differences in the conditionability of autonomic and skeletal responses. Furthermore, results of research on autoshaping are relevant for such current issues as (1) the "arbitrariness" of conventional operant responses like the pigeon's key peck and the rat's lever press, because these responses also appear during exposure to appropriate Type I procedures, and (2) the relative merits of response-centered versus perception-centered theories of simple learning, because approach tendencies toward a CS for food develop and persist after mere "observation" of CS-US contingencies and despite actual loss of reinforcement for approach behavior toward CS. The following abbreviated summary of research on autoshaping should provide a background for discussion of its various theoretical implications, particularly with regard to Konorski's views.

In contrast to conventional (instrumental) procedures that also use pigeons, response keys, and grain reinforcement, Brown and Jenkins (1968) made 4-sec grain delivery (US) contingent only on a prior stimulus (CS = 8-sec key illumination). Such a CS-US contingency was instituted after brief preliminary training in which pigeons learned to eat from a grain dispenser that was illuminated only during the 4-sec periods of US access. As pairings of CS and US continued, birds first exhibited increases in activity during CS and oriented toward the lighted key; then they began to approach it; finally they pecked at it. Brown and Jenkins named this phenomenon autoshaping and intended the term to cover the approach and contact behavior that develops toward a localized signal for an appetitive US. The procedure was *auto*matic and the pigeon could be said to have

shaped it*self* to perform a response that in previous research had required the experimenter's assistance—that is, the method of successive approximations or "manual shaping."

Subsequent experiments have demonstrated that the acquisition and maintenance of key pecking in this type of situation depend mainly on a positive contingency between key illumination and grain delivery; birds rarely peck keys that are illuminated randomly with regard to grain presentations (e.g., Wasserman, Franklin, & Hearst, 1974).[1] Furthermore, autoshaping has been observed in experiments involving a variety of different species, responses, reinforcers, and general situations. For example, rats approach, contact, and often depress a lever whose insertion into the chamber signals food or electrical stimulation of the lateral hypothalamus (Peterson, Ackil, Frommer, & Hearst, 1972); fish approach and touch a lighted target that precedes delivery of food (Woodard & Bitterman, 1974); and baby chicks approach and peck a key that signals the onset of a period of heat in a cold chamber (Wasserman, 1973; Wasserman, Hunter, Gutowski, & Bader, 1975). Finally, autoshaping is usually long-lasting; most subjects that are tested for thousands of trials continue to approach and contact the CS.

There is also a negative counterpart to these standard findings. Not only do pigeons approach and contact a localized visual stimulus that signals an appetitive reinforcer, but they also position themselves relatively far from the same stimulus when it signals that the reinforcer is *not* going to appear (Hearst & Franklin, 1977; Wasserman et al., 1974). The greater the frequency of food in the absence of the stimulus, the stronger is such withdrawal behavior. In these studies the spatial position of the pigeons was monitored on every trial by means of switches beneath the floor of the chamber.

---

[1]Several recent papers (e.g., Davol, Steinhauer, & Lee, 1977) have argued that the establishment and maintenance of autoshaped behavior may be explained in terms of (1) pecking generalized from the lighted magazine to the lighted key, plus (2) the peck-food conjunctions that occur after the first (generalized) key peck. Although magazine-key generalization effects certainly may promote autoshaping in various situations, this overall explanation of the phenomenon is inadequate. As Hearst and Jenkins (1974) pointed out, the first key peck is controlled by the degree of correlation between CS and US; birds are much less likely to peck the key, or they fail to peck it at all, if the CS and US are randomly or negatively related rather than positively related. Furthermore, various experiments (e.g., Peterson et al., 1972; Wasserman, 1973; Woodruff & Williams, 1976) have demonstrated autoshaping in situations in which no magazine was used and no directed US behavior was involved (the US entailed brain stimulation, or heating of a chamber, or injection of water directly into the pigeon's oral cavity). The proponent of the generalization explanation also has difficulty handling the fact that the first autoshaped key peck often takes 40 to 100 trials to appear and depends dramatically on the duration of the intertrial interval. Moreover, the "omission" and "observation" results to be described shortly also cannot easily be incorporated into an explanation that stresses key peck–food conjunctions in the acquisition and maintenance of autoshaped behavior.

A few additional facts about autoshaping, reviewed in more detail by Hearst and Jenkins (1974), should be mentioned here. First of all, autoshaping does not emerge or is relatively weak when other stimuli in the situation (auditory or visual) predict the arrival of a reinforcer as well as does the CS. Thus, autoshaping develops to the extent that the CS is a nonredundant, "informative" predictor of the reinforcer. Second, the type of behavior directed toward the CS depends on the kind of US with which it is correlated. A pigeon's key pecks at a food-predictive signal are evenly spaced, brief, and forceful, whereas its pecks at a water-predictive signal are irregularly spaced, sustained, and relatively weak; birds seem to be "eating" the former signal and "sipping" or "drinking" the latter. Analogously, rats gnaw or lick a CS (insertion of a lighted lever into the chamber) that predicts food, but sniff or explore the same CS when it signals electrical stimulation of the lateral hypothalamus, a US that itself produces sniffing and exploring. Thus the type of behavior directed toward the CS often resembles or is somehow appropriate to the type of US that it predicts—a point to which I return shortly.

The term *sign tracking* was coined by Hearst and Jenkins (1974) to cover these and several other effects in animal and human learning. Sign tracking refers to behavior (e.g., eye movements, bodily orientation, approach, signal contact) that is directed toward or away from a feature of the environment (a sign) as a result of the relation between that environmental feature and another (the reinforcer, in a typical experiment). Studies of sign tracking provide useful arrangements for assessing the relative behavioral control exerted by S–S ("classical") versus R–S ("instrumental") contingencies in simple learning.

B. Autoshaping and Type I Conditioning

On the surface, at least, autoshaping as a procedure and a phenomenon seems to conform to Konorski's definition of classical conditioning. It involves the paired presentation of two stimuli, one of which gives rise to a definite overt response from the outset of the experiment. As a result of the pairing procedure, the initially neutral stimulus—a lighted key, in the case of the pigeon—acquires the capacity to elicit the same response (approach and pecking) as does the biologically important stimulus (grain presentation). An association between a CS and a US is presumably formed.

However, this view of autoshaping is perhaps too simplistic. Several important and troublesome questions come to mind:

1. Is the response to the CS really the "same" as the response to the US? One could argue that the response to the US involves approaching and pecking in the grain dispenser, whereas these behaviors are directed toward