POPPER’S GEWÖHNUNGSTHEORIE ASSEMBLED
AND FACED WITH OTHER THEORIES OF LEARNING

Arne Friemuth Petersen
Former Professor of General Psychology
University of Copenhagen

Abstract. With the publication of Popper’s Frühe Schriften (2006), renewed possibilities for inquiring into the nature and scope of what may be termed simply ‘Popperian Psychology’ have arisen. For although Popper would never have claimed to develop such psychology there is, however, from his earliest to his last works, a wealth of recommendations as to how to come to grips with problems of the psyche without falling victim of inductivist and subjectivist psychology.

The fact that most theories of learning, both traditional and modern, have remained inductivist, and therefore logically invalid, places Popper’s hypothetico-deductive approach to learning and the acquisition of knowledge among the most important conjectures in that entire domain, akin to Edelman’s biological theory of consciousness. Central to Popper’s approach and his final rejection of all inductive procedures is his early attempt at a theory of habit-formation, Gewöhnungstheorie (in ‘Gewohnheit’ und ‘Gesetzerlebnis’ in der Erziehung, 1927) – a theory not fully developed at the time but nevertheless of decisive importance for his view on education and later works on epistemology, being ‘of lasting importance for my life’ (2006, p. 501).

Working from some of the original descriptions and examples in ‘Gewohnheit’ und ‘Gesetzerlebnis’, updated by correspondence and discussions with Popper this paper presents a tentative reconstruction of his Gewöhnungstheorie, supplemented with examples from present-day behavioural research on, for example, ritualisation of animal and human behaviour and communication (Lorenz), and briefly confronted with competing theories, notably those by representatives of the behaviourist tradition of research on learning (Pavlov and Kandel).

1. INTRODUCTION

In a monograph entitled ‘Gewohnheit’ und ‘Gesetzerlebnis’ in der Erziehung, which Popper submitted in 1927 as his final examination project to The Institute of...
Education in Vienna, he presents the results of his studies of what he calls 'dogmatic' and 'critical thinking' illustrated by many descriptions and examples, his own and others, of how 'habit-formation' (Gewöhnung) and 'law-driven experience' (Gesetzerlebnis) may operate in child and adult cognition and reasoning about the world.

The phenomenological part, representing about 2/3 of this work of 90 printed pages, was to be followed by a theoretical part containing a theory of 'law-driven experience' (Theorie des Gesetzerlebnisses) and a theory of habit-formation (Theorie der Gewöhnung), but not a single page of this theoretical part was found in Popper's Nachlass. It may even be questioned whether he ever put it down on paper or preferred to leave his monograph unfinished as stated in his List of Publications (Popper, 1974b, p. 1202).

However, we know from a letter written by Popper to Dr. Albert Krassnigg in 1970, now reprinted in Frühe Schriften, that he considered it of lasting importance for his life (‘für mein späteres Leben grundlegend’, p. 501) when in 1925 or 26 he discovered that the generally accepted 'Theory of habit-formation through repetition' was false and needed to be replaced by a better one. In the same letter he briefly states his theory like this: 'repetition can never make anything conscious, on the contrary, repetition turns conscious events into unconscious ones: all conscious learning is always learning by trial and error' (ibid.). About 15 years before this letter he had referred to his Hausarbeit as: ‘My logical criticism of Hume’s psychological theory, and the considerations connected with it (most of which I elaborated in 1926–27, in a thesis entitled “On Habit and Belief in Laws”),…suggests that the critical attitude is not so much opposed to the dogmatic attitude as super-imposed upon it… A critical attitude needs for its raw material, as it were, theories or beliefs which are held more or less dogmatically. (Popper, 1963, p. 50.)

So, clearly, this early theory must be considered central to Popper’s whole philosophy – as he also admits to Dr. Krassnigg when he writes: ‘This psychological insight later carried me to a logical critique of the theory of induction – which became the very beginning of my independent philosophy of the natural sciences.’ (2006, p. 502.)

For students of Popper’s philosophy it would be of obvious interest to know more about his theory of habit formation and its connection with his later rejection of induction and inductive procedures. Judging from the scattered remarks, quoted above, it would also be of interest for theorists of learning and education, since Popper’s theory runs counter to traditional theories and teachings within these domains, although adherents of the latter might grant Popper that habit-formation and law-driven experience seem incompatible with educational goals like free self-determination, independence and autonomy, since notions like ‘habit’ and ‘habit-formation’ carry with them a decisively passive element and a strong moment of automatism, leaving aside active thinking, reflection and insight (op. cit., p. 91).
Thus motivated, we may be ready to embark on a reconstruction of Popper’s enigmatic theory of habit-formation, which will take us through a central part of his epistemology.

2. TENTATIVE DELINEATION OF DOGMATIC AND CRITICAL THINKING

In a short introductory section to the 1927 Monograph, called ‘Psychological Objective’, Popper characterises (2006, p. 94) dogmatic thinking as a kind of thinking which ‘blindly’ accepts (hinnehmt) basic principles or rules as true without even considering that they may be false or proved wrong by argument or experience; they are maintained (festgehalten) and stubbornly employed whenever something in a situation invites their carrier to do so.

Critical thinking is tentatively characterised (op. cit., p. 95) by a non-ending questioning of currently accepted and maintained basic principles or rules in order to adopt and employ them after they have been substantiated (bewäiren), at best by experience.

Typical of dogmatic thinking are thus the attributes and examples noted in Table I.

The following phenomenological part brings more illustrations to the identified phenomena under the headings ‘Typical attitudes towards the unknown’, ‘Coming to grips with things’, ‘Maintaining an approach through “Conservatism”, “Pedantry” or “Ritual”’ – with which the whole work comes to an abrupt ending. In this collection of examples there are two references to the theoretical part, originally planned to conclude the thesis: one (op. cit., p. 123) on the diminution of excitability upon repeated stimulation (the phenomenon of habituation), and a second one (op. cit., p. 148, note 2) on the apparent conflict between learning by mechanical association (the existence of which Popper denies in later works) and learning by insight.

3. HABIT-FORMATION CONSIDERED AS A RESULT
   OF PROGRESSIVE SIMPLIFICATION OF INITIAL ACTS
   AND A SOURCE OF UNCONSCIOUS EXPECTATIONS

With these concepts and hints in mind we now proceed to Popper’s opus magnum, Die beiden Grundprobleme der Erkenntnistheorie, written 1930–33 (published only in 1979), in which we find ideas of behavioural selection used as Popper argues in favour of a Deductivist Psychology of Knowledge (Section 4) – a Kantian inspired psychology. To grasp some of the basic elements for our runaway theory and its biological significance, the main lines of thought may be sketched as follows.
Table I

Poper’s characterization of ‘dogmatic thinking’

(1) *Hinnehmen* = uncritical acceptance of general principles, rules, etc.

(2) *Festhalten* = hard-nosed maintenance of principles and rules in the face of facts and arguments;

with the core-characteristics named:

(3) *Gesetz erlebnis* = law-generated experience which, expressed in words, functions according to the doctrine:

'So ist es, so muss es sein' = ‘So is it, so must it be’.

[A case of G.E. Moore’s ‘Naturalistic fallacy’!]

When faced with the unknown a child, or an adult, may go through 3 stages in establishing such a law-generated experience:

(i) *Einstellung* = attitude towards the unknown;

(ii) *Setzung* = adoption of some measure (principle, rule, action) or non-action;

(iii) *Festhalten* = same as (2), above.

Examples:

ad 1: A 3.5 year old child tells to his mother about the red-green railway lights: ‘Do you see that the red light always becomes green when the sun shines.’

ad 2: When somebody pointed out to the philosopher Hegel that Nature did not agree with his Natural Philosophy, the great man replied: ‘So much worse for Nature’.

ad 3: A 3 year old girl who had seen her aunt wearing a necklace once insisted: ‘Aunti-Elise must always wear a necklace’.

ad i-ii: Popper describes (pp. 112-113) how a new apprentice in the workshop of a cabinet-maker is (i) somewhat scared by the new setting, and (ii) tries to find out how to behave correctly towards his Master and his older peers, (iii) maintaining any minute rule of thumb he finds to navigate by professionally and socially.

As his starting point for showing that a deductivist and non-sensualist psychology is possible, Popper starts out from an idea of Ernst Mach who, surprisingly enough, in his *Principien der Wärmelehre* explains the origin of mental constructs, and of thought, on a biological basis – more precisely, as a result of the interactive relationship between reaction and perception. Mach states: ‘What is reacted to in the same way falls within the same concept. As many types of reactions, as many types of concept.’ (Popper, 1979, p. 24.)
This idea of Mach’s is supported by the neurophysiologically founded argument that the nervous system can be divided into an afferent and an efferent side, corresponding to the distinction between reception and reaction. Popper now argues that physiological and psychological reactions may well be triggered by stimuli from outside, but with respect to the specific type of reaction – the way in which the reaction takes place (Reaktionsablauf) – the stimuli are heavily dependent on the reacting apparatus itself. ‘The releasing, objective stimulus can thus be regarded as the material condition for the reaction, since it is necessary in order for the reaction to take place; the reacting apparatus, however, contains the formal conditions for releasing the reaction. Such reactions can thus be considered “subjectively pre-formed”… and therefore not “resulting from experience’.” (Ibidem)

Therefore, when a human being or an animal shows a consistent reaction in similar situations, it is because a tentative assignment of the reaction to the stimulus has been established. Such kind of assignment process always precedes the test situation – that is, situations in which it may show its biological value – so that, until then, the potential reaction functions as an anticipation, an ‘unfounded prejudice’ or dogma, which is a synonym for expectation. Popper says: ‘Whether or not this provisional assignment is abandoned depends on its biological value. The test method is a selective one: if the anticipations are of no value, they are weeded out.’ (Op. cit., p. 26.)

Being an admirer of Darwin for his exploits and use of hypothetico-deductive methodology, Popper hoped from the outset that selective principles could account not only for the evolution of species and their bodily ‘hardware’, as Darwin had shown, but also for their ‘software’, such as learning ability and individual adaptation. In other words, the problem was to explain how the organism’s pre-formed reactions, which logically cannot have developed through experience, can serve as a basis for the acquisition of experience about the surrounding world. In short: how to explain that organisms learn only through deductive procedures.

Due to the lack of ethological observations, which had only just begun to be known to non-specialists when Popper worked on his book, he turned to Herbert Spencer Jennings’s studies of lower organisms (1906) to answer this question. Jennings had developed a theory of ‘exploratory movements’ to account for various problem-solving behaviour in protozoa like Paramecium. Jennings had observed that if this animal is exposed to a noxious stimulus, then it would run though its whole repertoire of pre-programmed behaviour until one of the preformed behaviour patterns turned out to be functional in ridding the animal of the stimulus. If the noxious stimulus was repeated, then the entire behavioural repertoire was run through again, and even after several repetitions no changes occurred within the behavioural sequence itself – only a decrease in the time taken to browse through the repertoire (Ablaufsverkürzung) was seen, resulting in the ‘adequate’ response occurring more promptly (see Figure 1). The decrease in time, when the behavioural sequence was repeated, was obtained by gradual simplification and better integration of the ciliary movements.
Fig. 1. The selection-effect of repeated trials in problem-solving behaviour of Paramecium; theoretical curve based on Jennings’s 1906 observations.

Traditional learning theory, such as that of Russian and American behaviourism, had explained similar reactions in other animals differently, since it takes all reactions to be the result of repeated stimulation that gradually ‘builds up the reactions’ so to speak. Popper’s deductivism, however, holds it for impossible that repetition can bring about anything new, either in learning or in cognition. Repetition can only make things disappear: ‘Habit and practice only remove the meanderings in the course of the reaction – straighten it out as it were. Nothing new is thus created by repetition. The increasing promptness of a reaction should not be regarded as its gradual creation (natura facit saltus).’ (Op. cit., p. 28.)

What divides the traditional interpretation and Popper’s interpretation of such behavioural phenomena are obviously their different views on the function of repetition. This difference comes out more clearly in the next example – Konrad Lorenz’s description of his famous goose, Martina – to which Popper refers in his Intellectual Autobiography (Popper, 1974a, p. 160, Note 44).

Lorenz (1966, p. 57–58) describes how an original escape behaviour performed one day by Martina in a sudden frightening situation, was seen to shorten progressively over the course of a year, finally ending up as a ritualised behaviour pattern which, despite its radical simplification, still calmed the goose (see Figure 2 and text). In contrast to Lorenz’s use of behaviourist learning theory in his interpretation of Martina’s development of a behavioural habit, Popper admits no inductive procedures in his interpretation of how repetition transforms Martina’s behavioural detour towards the window into a ritualised behaviour pattern before mounting the stairs in Lorenz’s home, and he takes this example as a
case in point of how animals and humans develop new adaptations. Each time
Martina performs her ritualised behaviour it changes a little as it becomes shorter
in space and time; thus ‘there is no genuine “repetition” but rather a change
through error elimination…which helps to make certain actions or reactions
automatic, thereby allowing them to sink to a merely physiological level, and to be
performed without attention.’ (Popper, op. cit., p. 38.)

Fig. 2. Schematic representation of progressive simplification,
or ritualisation, of a fear reaction in the greylag goose, sketched on location
and in accordance with the following passage in Konrad Lorenz (1966, p. 57–58):

‘In our house in Altenberg the bottom part of the staircase, viewed from the front door,
stands out into the middle of the right-hand side of the hall. It ascends by a right-angled
turn to the left, leading up to the gallery on the first floor. Opposite the front door is a
very large window. As Martina, following obediently at my heels, walked into the hall,
the unaccustomed situation suddenly filled her with terror and she strove, as frightened
birds always do, towards the light. She ran from the door straight towards the window,
passing me where I now stood on the bottom stair. At the window, she waited a few
moments to calm down, then obedient once more, she came to me on the step and
followed me up to my bedroom. This procedure was repeated in the same way next
evening, except that this time her detour to the window was a little shorter and she did
not remain there so long. In the following days there were further developments: her
pause at the window was discontinued and she no longer gave the impression of being
frightened. The detour acquired more and more the character of a habit, and it was
funny to see how she ran resolutely to the window and, having arrived there, turned
without pausing and ran just as resolutely back to the stairs which she then mounted. The habitual detour to the window became shorter and shorter, the 180° turn became an acute angle, and after a year there remained of the whole path habit only a rightangled turn where the goose, instead of mounting the bottom stair at its right-hand end, nearest the door, ran along the stair to its left and mounted it at rightangles."

Popper and Lorenz discussed and corresponded about these and related issues over several years and when I visited Lorenz in Altenberg as late as 1976 he vividly showed me the location where he had followed Martina's behavioural development, repeating his explanation of the phenomena, so no agreement on the matter had apparently been reached between the two. This was perhaps only to be expected as we can see from a letter Popper wrote in 1975 (Popper, 1975b), in which he gives his point of view and comments interestingly on the position of Lorenz:

'Quite recently I received from Lorenz his Nobel Lecture plus his Nobel Autobiography. I replied by sending him a letter in which I criticized his theory of conditioning – a theory which he upholds even though he has now given up induction by repetition. (Tinbergen, as far as I know, still upholds inductivism.)

The situation is highly interesting. According to Konrad Lorenz's autobiography (in Les Prix Nobel en 1973) and his Die Rückseite des Spiegels... it was Erich von Holst who converted Lorenz from the reflex theory to a theory which is certainly much better, but still not quite good enough: that the stimulus does not stimulate the response, but rather inhibits an inhibiting neurone which normally suppresses the (potentially always ready) response. The inhibition of the inhibition allows the (normally suppressed) excitatory neurone to fire, and the response takes place.

I believe that this is a much better model than that of an input-output reflex arch; but still vastly oversimplified (even on Lorenz's own arguments). But worse is to come: Lorenz, though he rejects reflexes, believes in conditioning and reinforcement as if he were still a reflexologist.

My point is that learning is not "by association"; it is not a process consisting of elementary processes such as reflexes, or Holst-models, or associations. All these models are completely misleading. The central point is that all learning is adaptive modification – sometimes complication, sometimes simplification – of previously existing highly complex adaptive skills. Thus learning is never associative, but selective and modifying. It does not put together simple elements, but modifies existing complex structures. These complexes may, in their turn, consist of elements such as neurones and synapses; but the idea that an association corresponds to a thickened synapse is totally mistaken; a slight modification of a skill probably corresponds (assuming the synapses are the modifiable physical elements) to the modification of a thousand synapses, of a whole network. Even lower animals which are not yet capable of learning (of modification) have highly complex behaviour; and it is such complex behaviour which is the antecedent of all learning processes.

Nothing is therefore more misleading than the idea that the association of meaningless syllables is a model of an elementary learning process, and that more "natural" learning processes are complexes of such elements. Learning begins with inbuilt fears and expectations; and what is learnt first is to distinguish between what is
biologically wholesome and what is dangerous. In other words, learning starts with meaning – with what is biologically meaningful; which makes it quite impossible to accept any of the associationist and atomist models. A Gestalt-switch is needed, as Kuhn would say…’

That habit-formation also implies the formation of expectations on the part of the active organism comes out from another surprising observation on Martina made by Lorenz (1966, p. 58):

‘One evening I forgot to let Martina in at the right time, and when I finally remembered her it was already dusk. I ran to the front door and as I opened it she thrust herself hurriedly and anxiously through, ran between my legs into the hall and, contrary to her usual custom, in front of me to the stairs. Then she did something even more unusual: she deviated from her habitual path and chose the shortest way, skipping her usual right-angle turn and mounting the stairs on the right-hand side, “cutting” the turn of the stairs and starting to climb up. Upon this, something shattering happened: arrived at the fifth step, she suddenly stopped, made a long neck, in geese a sign of fear, and spread her wings as for flight. Now she hesitated a moment, turned round, ran hurriedly down the five steps and set forth resolutely like someone on a very important mission, on her original path to the window and back. This time she mounted the steps according to her former custom from the left side. On the fifth step she stopped again, looked round shook herself and greeted, behaviour mechanisms regularly seen in greylags when anxious tension has given place to relief. I hardly believed my eyes. To me there is no doubt about the interpretation of this occurrence: the habit had become a custom which the goose could not break without being stricken by fear.’

If they ever discussed these surprising observations in detail, Lorenz and Popper could, no doubt, have agreed on the importance of expectations in the life of animals and humans, even though their theories and vocabulary were far apart, as we have seen. The situation was perhaps that Lorenz could not give in to the logical argument from a philosopher, not even his old friend Karl Popper who was also a naturalist of sorts. This was, however, not the case with the zoologist Julian Huxley whose concept of ‘ritualization’ Lorenz adopted (1971, p. 146 and note 32) sacrificing his own term ‘formalisation’, which he had introduced to denote ‘a behaviour that has become a pure ceremony’ (op. cit., p. 25). We can only regret that the otherwise excellent Sir Julian did not invite Popper to The-Royal- Society discussion on the ritualisation of behaviour in animals and man (1966); it would then not have been too late to realise that Popper’s contemporary theory of problem-solving and learning could have contributed to a resolution of the question of the origin and phylogeny of ritualisation, especially of animal and human communication.

Whatever the case, we can now bring together in Table II the various developments in Popper’s reasoning about the reactions of living beings towards something new, which reveal the dynamics of problem-solving and the role of repetition in the selection of behaviour by error-elimination. In doing so it is easily seen that Popper’s schema for problem-solving emerges by itself on the left-hand side of the table.
Table II

The Dynamics of Problem-Solving and the Role of Repetition in the Selection of Behaviour by Error-Elimination

<table>
<thead>
<tr>
<th>P1</th>
<th>P2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dogmatic stage</strong></td>
<td><strong>Critical stage</strong></td>
</tr>
<tr>
<td>D1 Pre-programmed assignment (Setzung)</td>
<td>C1 Unconscious selection of behaviour patterns (error-elimination)</td>
</tr>
<tr>
<td>D2 Assignment by analogy (re-use of assignment from similar situations)</td>
<td>C2 Conscious search for errors (self-criticism)</td>
</tr>
<tr>
<td>D3 Tentative assignment (by imitation or exploration)</td>
<td>Habit-formation (unconscious dispositions)</td>
</tr>
</tbody>
</table>

In 1937, when in New Zealand, Popper had used the idea of problem-solving in a paper, ‘What is dialectic?’, read at a philosophical seminar at Canterbury University College, to account for the development of human thought, especially in philosophy – no doubt in order to place Hegelian thinkers among other erring animals. Without drawing up any schema, and with no reference to Lloyd Morgan, who had introduced the term ‘trial and error’ in 1894 to describe animal learning, Popper declared that the usual method by which solutions to problems in life may be found is, indeed, ‘the method of trial and error’, and he described its generality as follows: ‘This, fundamentally, is… the method used by living organisms in the process of adaptation. It is clear that the success of this method depends very largely on the number and variety of the trials: the more we try, the more likely it is that one of our attempts will be successful.’ (Popper, 1940; 1963, p. 312.)

There is also an updated remark about the importance of ‘dogmatic thinking’ for this process: ‘We may describe the method employed… as a particular variant of the trial and error method. Men seem inclined to react to a problem either by putting forward some theory and clinging to it as long as they can (if it is erroneous they may even perish with it rather than give it up$^1$), or by fighting against such a theory, once they have seen its weakness. This struggle of ideologies, which is obviously explicable in terms of the method of trial and error, seems to be
characteristic of anything that may be called a development of human thought.’
(Ibidem) The indicated footnote marks an answer to a frequent criticism of
Popper’s insistence on the important function of ‘dogmatic thinking’ made by
students and colleagues who did not know his original, unpublished work: ‘The
dogmatic attitude of sticking to a theory as long as possible is of considerable
significance. Without it we could never find out what is in a theory – we should
give the theory up before we had a real opportunity of finding out its strength; and
in consequence no theory would ever be able to play its role of bringing order into
the world, of preparing us for future events, of drawing our attention to events we
should otherwise never observe.’

In this remark the impact of theories on perception is very well explained.
The negative side of this psychological effect should, however, not be forgotten,
\textit{i.e.} the equal impact of \textit{false} theories on perception which makes their carriers
blind and deaf to other ways of viewing the world. Cases in point are the prevailing
‘mirror-theories’ of consciousness and the behaviourist theories of learning which
make their adherents live in not only a world of \textit{perceptual} illusions but also in a
world of \textit{cognitive} illusions.

4. \textsc{From Habit-Formation to Learning by Action and Selection}

As we have seen Popper’s schema for problem-solving not only took long to
emerge but, due to circumstances of war and emigration, many years went by
before he had occasion of making it known to the public. It happened in 1965 at
Washington State University where he was invited to give ‘The Arthur Holly
Compton Memorial Lecture’ (Popper, 1966).

Knowing some of the preceding steps in its long incubation process, and
especially how it was to be used later, it is interesting to note that the now famous
tetradic schema was presented for the first time as ‘an evolutionary sequence of
events’ (\textit{op. cit.}, p. 24) – as a process that had to do more with the evolution of
animal species by natural selection, or \textit{indirect problem-solving}, than with the
individual animal adapting itself by learning to cope in given contexts, that is, by
\textit{direct problem-solving} (Petersen, 1985, pp. 235f.). Indeed, the procedure
represented by the schema shown in \textit{Figure 3} is described as a restatement of
evolutionary theory which is characterised as an ‘evolution of new means for
problem-solving, by new kinds of trials, and by new methods or error-elimination;
that is to say, new methods for \textit{controlling} the trials.’ (Popper, \textit{op. cit.}, p. 21.) The
individual organism is considered a hierarchical system of \textit{plastic controls}, where
the regulated subsystems make trial-and error movements which are constantly
checked by the higher-level systems of control resulting in some trials being
entirely suppressed while other trials are kept and perhaps modified by ‘local’
error-elimination.
Popper’s theory thus consists of ‘a certain view of evolution as a growing hierarchical system of plastic controls, and of a certain view of organisms as incorporating – or in the case of man, evolving exosomatically – this growing hierarchical system of plastic controls. The Neo-Darwinist theory of evolution is assumed; but it is restated by pointing out that its “mutations” may be interpreted as more or less accidental trial-and-error gambits, and “natural selection” as one way of controlling them by error-elimination.’ (Popper, op. cit., p. 23.)

Unlike most evolutionary thinkers who consider natural selection processes and learning processes to be of different nature, Popper (in his second Herbert Spencer Lecture, 1975a, p. 73f.) takes these processes to be basically alike; furthermore, in considering progress in science as a means used by Homo sapiens to adapt himself to the environment, Popper compares three levels of adaptation: (i) genetic adaptation, (ii) adaptive behavioural learning, and (iii) scientific discovery – and he explains the fundamental similarity between the levels by advancing the hypothesis that the mechanism of adaptation is the same on all three levels: ‘Adaptation starts from an inherited structure which is basic for all three levels: the gene structure of the organism. To it corresponds, on the behavioural level, the innate repertoire of the types of behaviour which are available to the organism; and on the scientific level, the dominant scientific conjecture or theories. These structures are always transmitted by instruction on all three levels: by the replication of the coded genetic instruction on the genetic and the behavioural levels; and by social tradition and imitation on the behavioural and the scientific levels. On all three levels, the instruction comes from within the structure. If mutations or variations or errors occur, then these are new instructions, which also arise from within the structure, rather than from without, from the environment.’ (Op. cit., p. 74.)

Inspired by Niels Kaj Jerne’s natural-selection theory of antibody formation, especially Jerne’s 1967 article, Popper concludes his argument by saying: ‘scientific discovery depends on instruction and selection – on a conservative or
traditional or historical element, and on a revolutionary use of trial and the elimination of error by criticism, which includes severe empirical examinations or tests’ (Popper, op. cit., p. 78). Thus contrary to first impressions, instruction is the conservative power, while selection is the revolutionary power. And, with this, we are brought right back to our point of departure: Popper’s early work on dogmatic and critical thinking, and how he later used these ideas in his understanding of animal and human learning.

In retrospect, it is remarkable that around 1930, before the first major results of ethology and related disciplines had become known, Popper could draw up the main lines of a deductivist theory of learning relying only upon examples from the behaviour of infusoria, described by Jennings, the psychology of George Bernard Shaw and a few others. For had Lorenz’s work on imprinting (1935), Spitz’s (1945) and Harlow’s (1959) discoveries of the attachment phenomenon in young primates and human infants, and the description by Carey and Bartlet (quoted in Rice, 1990) of a so-called ‘fast-mapping’ in the young child’s learning of new words, just by hearing them once – been available to Popper at the time, these and other cases of pre-programmed behavioural dispositions could have illustrated his non-inductivist theory of learning in a more systematic and convincing way.

Nevertheless, logically, the theory is highly attractive as it is resumed, for example, in Section 4 of Die beiden Grundprobleme (Popper, 1979, pp. 23f.): a necessary condition for organisms to be able to solve problems or learn from experience is that a tentative assignment is established between an already existing behavioural disposition and some stimuli pertaining to the problem situation, as in the case of Martina, discussed before. Living beings do indeed make such jumps to conclusion thereby venturing the first step towards a possible solution to a given problem. The assigned relationship thus established will, however, only be maintained insofar as the anticipatory behaviour is biologically functional in solving the problem. If the problem situation changes, then the organism will have ‘to jump to another conclusion’ by applying another behavioural disposition more to the point of the changed situation. It is easily seen that such a continual selective elimination of pre-formed behaviour patterns, of expectations or hypotheses in real life situations has much in common with evolution by natural selection.

The assignment process, which seems akin to a semi-irreversible form of imprinting, may be considered a dynamic expression of the affinity of organisms for imposing regularities on the world. Such process may be ascribed to the existence, as much in the control system of unicellular organisms as in the nervous system of multicellular animals, of the mentioned differentiation between an afferent side and an efferent side (i.e. between reception and reaction), and especially because the efferent side is normally more autonomous than the afferent side. It is thus characteristic, as pointed out above, that the effect of incoming stimuli depends largely on the reacting apparatus itself: the stimulus which triggers the reaction is therefore only its material condition (since it determines the moment
of the appearance of the reaction and certain quantitative aspects of it), while the efferent apparatus constitutes its formal condition (since this will determine the whole character of the reaction).

It is comforting to note that some recent developments in neurobiological theory, such as Karl Pribram’s model of ‘test-operate-test-exit’ (TOTE), were aimed at explaining precisely this type of central control of receptive mechanisms: ‘…perception is in essence a “motor” phenomenon… perception per se is more a reflection of the response patterns instigated in the brain by an input than it is a resultant of the input patterns’. (Pribram, 1971, p. 91.)

Thus, for logical, reasons, no direct assignment can occur between the sensations (derived from ‘repeated perception’), as has been assumed by traditional learning theory. The sensations must first be assigned within the register of pre-existing reactions, before – via this pre-formed ‘carrier system’ – indirect assignments can be established between the sensations themselves (those fictive sensorial links which, ever since Aristotle, traditional learning theory has taken to be the basis of all experience – as if experience only consisted of accumulated sensorial associations).

When set theory appeared at the end of the 20th Century, it was thought, also in psychology – as often is the case on the occasion of some such new development – that the new tool could solve a lot of old problems in various fields. Regarding our present context it may be that the generalisation made by Schlick (1925, p. 35) of Dedekin’s notion of assignment (Zuordnung, which Popper also employed in his doctoral dissertation, 1928; 2006, p. 239f.) – ‘In thinking, there is basically no other function than that of assignment’ – when used properly could help to bring learning theory out of its prevailing inductivist impasse by contributing to our understanding of the logical foundation of problem-solving.

5. SPEAKING AGAINST LIPSIUS: VERSUS THE INDUCTIVIST TENETS OF TRADITIONAL LEARNING THEORY

As should have become clear from the foregoing, one of the astonishing lesions to be drawn from the Lorenz-Popper debate over habit-formation is that Lorenz’s firm adherence to the theory of learning by conditioning which, like so many others brought up in a tradition of behaviourist physiology and psychology, he had no doubt adopted as a student could not even be refuted by the precise descriptions he himself later contributed.

Popper, on his side, continued to develop his deductivist theory of habit-formation into a general theory of problem-solving which, in his Autobiography (1974a, p. 109 and p. 143), he took as the main characteristics of life in answering Schrödinger’s question ‘What is life?’ Some time later, in his contribution to The Self and Its Brain (written in collaboration with John Eccles) Popper distinguishes 8 different stages in the process of actively learning by trial and error, or by problem solving, or by action and selection:
The active exploration, guided by inborn and acquired “knowledge how”, and by
(background) “knowledge that”.
(2) The production of a new conjecture, a new theory.
(3) The criticism and testing of the new conjecture or theory.
(4) The rejection of the conjecture, and the recording of the fact that it does
not work.
(“Not this way.”)
(5) Repetition of this process (2) to (4) with modifications of the original
conjecture or with new conjectures.
(6) The discovery that a new conjecture seems to work.
(7) The application of the new conjecture involving additional tests.
(8) The practical and standardized use of the new conjecture (its adoption).
I conjecture that only in (8) does the process take on, in stages, the character

From this logical stronghold Popper bluntly declares Pavlov’s unconditioned
and conditioned reflexes for non-existing, and he goes on to criticise the
behaviourist theory of learning by the following points (op. cit., p. 136f.).

(i) The behaviourist considers the organism a passive device waiting to be
stimulated; Popper, on the other hand, attributes to the organism ‘an active interest
in its environment’, a mainly ‘unconscious exploratory instinct’ leading to
exploratory behaviour which is not just a complex of ‘reflexes’ in Pavlov’s sense
but the manifestation of an inborn aim-structure that generates a general curiosity
and species-typical activity towards the environment, especially towards other
members of the species.

(ii) Even the so-called ‘unconditioned’ and ‘conditioned responses’ are not
just reflexes; from Popper’s point of view ‘Pavlov’s dog rather developed the
obvious theory or expectation: that the food would arrive when the bell rang. This
expectation made its saliva flow – exactly as the expectation raised by the visual
perception or the smell of food.’

(iii) The behaviourist assumes that all biologically important regularities to
which an organism can adapt itself consist in coincidences, like that of the bell and
the arrival of food; but the structure of the environment to which organisms have to
adapt themselves, Popper argues, has ‘no similarity with Hume’s perpetual
conjoined impressions’ or, we may add, with the artificial constructs of contiguity
and contingency theories of learning which are unwittingly construed to account
for animal behaviour in laboratory settings.

(iv) Against the inductivist use of repetitive procedures especially in human
learning

Popper argues (op. cit., p. 134): ‘Repetition does play a role in behavioural
adaptation, but it does not contribute to discoveries. Rather, it helps, after the
discovery is made, to make of it an unproblematic routine and therefore to make it
unconscious. (This is so with... skills like walking, or bicycling, or piano playing.) Repetition, or practising, is no way of acquiring new adaptations: it is a way of turning new adaptations into old ones, into unproblematic background knowledge; into unconscious dispositions.'

(v) Regarding the above mentioned 8 stages required for learning by problem-solving or by action and selection, Popper (op. cit., p. 142) points out, that there is no reason to assume that the implied procedures are of the same nature, like certain known reflexes; neurologically, learning could well consist of a ‘hierarchical organisation of structures of structures’, something like a dynamic hologram.

One of the examples chosen by Pavlov (1927; 1960, p. 22f.) to show that the principle of conditioning has a bearing upon the reactions of dogs outside the laboratory, relates to an experiment carried out by a colleague: ‘Dr. Zitovich took several young puppies away from their mother and fed them only milk for a considerable time. When the puppies were a few months old he established fistulae of their salivary ducts, and was thus able to measure accurately the secretory activity of the glands. He now showed these puppies some solid food – bread or meat – but no secretion of saliva was evoked... Only after the puppies have been allowed to eat bread and meat on several occasions does the sight or smell of these foodstuffs evoke the secretion.’

At first glance Pavlov seems to have proved his principle also for this demonstration. On closer inspection, however, and viewed in the light of ethological theory, a deductive interpretation of this experiment, along the lines suggested by Popper, appears to be more plausible. What happened in the puppies when they adapted themselves to liquid food may be described as an assignment of a foodstuff to a pre-programmed, alimentary behaviour repertoire (and digestion) – followed by a progressive simplification of the species-typical pattern of feeding behaviour relative to the nature of the assigned food (milk) and the different ways it was delivered, first by the mother animal and then by the experimenter. Ethologists will say that the puppies establish a search model in which only milk is represented at first; when solid food is introduced later, the explorative, alimentary behaviour pattern is needed again to establish a new assignment, followed by a progressive simplification of the behaviour needed for coping with the new foodstuff. Solid food is in this way incorporated into the young animal’s search model regarded nutrition – and only then doe the saliva flow.

To be sure, Pavlov did not receive the Nobel Prize in 1904 for his work on conditioning but for his earlier work on the principal digestive glands. This may, however, be said about Professor Eric R. Kandel who won the Prize in 2000 for his work on the cellular and molecular processes of memory employing methods of conditioning. Although the emphasis is on memory Kandel’s work is clearly a continuation of the behaviourist tradition, and many will say that it has enlivened dwindling hopes.
Kandel set out like a comparative zoophysiologist studying neural mechanisms of behaviour and learning in molluscs, the results of which were communicated in a number of papers and in *Cellular Basis of Behavior* (1976) and *Behavioral Biology of Aplysia* (1979). Apart from their contributions to evolutionary biology, with descriptions of neuronal mechanisms for homologies, divergence and speciation, these studies were undertaken in order to replace traditional intervening variables, such as *arousal*, *motivation*, *drive*, *associative learning*, etc., as explanations of behaviour by *cellular mechanisms* – for example, whether classical conditioning could be an associative extension of *cellular sensitisation*. During this work *Aplysia* turned out to be an ideal experimental animal, having large synaptic cells suitable for implantation of micro-electrodes and a modicum of responsiveness. So, after having established that habituation, sensitisation, and classical conditioning could be obtained from individual cells of the synapse known as R2 in this animal, Kandel went on to demonstrate that the synapse did, indeed, change (was strengthened or grew) as a result of these processes being provoked for a sufficiently long time. These effects, as with Pavlov, were taken right away as proof that simple forms of learning had taken place (Kandel, 2006, p. 159f. and p. 165f.).

At no other moment in the present account has the conflict between a deductivist and an inductivist interpretation of learning appeared more acutely. Does learning, as a matter of fact, equal a thickening of a synapse or not? Unfortunately, Popper does not seem to have heard of Kandel’s work, at least there are no written traces of that. How would he have reacted to those experimental results? – The following is an attempt to answer that question in drawing upon the foregoing account of his contributions to a deductivist theory of problem-solving and learning.

With his wonderful investigations into the biology of molluscs and their behaviour, Dr. Kandel may claim to have shown that at the synaptic level a *priming*, or *re-tuning*, of the cell occurs as a function of the stimulation received, in other words: a change in the electro-chemical state of the synapse occurs, enabling the cell either ‘to anticipate’ some important (‘alarming’ or ‘favourable’) stimuli as in the case of sensitisation and conditioning, or to return to a baseline of neural activity (‘ignoring unimportant’ stimuli) as in the case of habituation. However, it seems far too easy to jump to the conclusion, as Kandel does, that what the synapse is manifesting in these cases are *simple forms of learning* – i.e. ‘learning at the cellular level’. It would have been not a lesser achievement to claim to have demonstrated how behavioural *anticipation* is present already on the cellular level – even though this control can only be part of the story (if, indeed, the learned society can agree that this is really what happens, as readily as they accepted the less convincing, but more popular, hypothesis about learning).
Re-phrasing a comment by Paul Weiss (1968, p. 24), ‘the story of “molecular control of cellular activities” is bound to remain fragmentary and incomplete unless it is matched by knowledge of what makes a cell the unit it is, namely, the “cellular control of molecular activities” – it may be claimed that the story of synaptic control of *behavioural anticipation* will remain incomplete unless it is matched by knowledge of how the nervous system exerts an influence on the individual synapse, or network of synapses, and how the entire hierarchical aim-structure of the organism guides its activities and interchanges with the environment. A similar argument may hold for learning and memory, and biologists and psychologists inspired by Popper talk about this guiding effect as ‘*downward causation*’ (e.g. Campbell, 1974, p. 179f.), meaning that a macro structure may, *qua* whole, act upon its constituents, molecules and atoms, as, say, the cell acts upon molecules inside its membrane.

‘*Downward causation*’ is thus at work when a higher-level control system influences a lower-level one. The effect is clearly a *selective* one, whereas the effect of ‘*upward causation*’ is *instructive* (Petersen, 1983, p. 72). The direction of causation inherent to Pavlov’s and Kandel’s behaviour models is clearly *upward*, rendering their principles and tenets *inductive* – *i.e.* learning by repetition, or induction – and thereby logically invalid (Popper, 1977, p. 134f.).

Anticipation, or *feedforward*, is constantly at work in the adaptation of living beings to their environment – even plants are capable of anticipation – and there are examples of it on all levels of the animal kingdom. One such example is *animal play* which consists of both pre-programmed, species-typical behaviour patterns that are adjusted and calibrated during play, long before the young animal needs to use them in serious adult contexts, and of acquired behaviour which, depending on the constraints on learning of the given species, may solve local problems of adaptation for one or more individuals. Much learning takes place in play and exploration, and it is particularly conspicuous when young animals and children learn motor skills. Indeed, the ease with which tentative movements are endlessly repeated prompts the conjecture that play is Nature’s ‘invention’ of a reward system for executing (preferably in a non-serious, playful context) behavioural acts, until they are perfected by gradually eliminating innervations of superfluous muscular movements (Petersen, 1988, p. 38f.).

The much disputed fact that only little learning occurs when conditioning is at work can be seen in the following example from real life, which may amuse admirers of Dvorak’s music. Earlier this year a music lover told the Danish Music Radio that she had been unable to listen to Dvorak’s ‘Symphony from The New World’ ever since she had seen Anatol Litvak’s horror movie, *The Snake Pit* (1948), in which Dvorak’s music accompanies disturbing scenes from a psychiatric ward in America - apparently in order to obtain a sarcastic effect. So strong were her later aversions towards this beautiful music that she always had to leave the place where it was played, the effect on her being almost corporal. – Now, what
has our poor music lover retained from Dvorak’s Symphony? Surely nothing
compared to what the musicians participating had to learn and master, perhaps not
even what an average ‘passive’ listener might remember with respect to melody,
harmony, rhythm, etc. of this music. For our ‘conditioned’ music lover, just a tiny
part of the Symphony had become an alarm signal so intensive that it drowned
everything else. Obviously, to get to know and learn, say, a piece of music,
something else than conditioning has to be called upon, and it is this gap between
theory and reality that one typically encounters when conditioning is claimed to be
the basis of all learning.

6. ‘WHAT IS TRUE IN LOGIC IS TRUE IN PSYCHOLOGY’

Around 1929 when Popper abandoned psychology to devote himself to an
objectivist epistemology, to the logic of discovery, he did so after having wrestled
with the problem of induction – especially with ‘the bankruptcy of eighteenth-
century reasonableness’ (Russell, 1946; 1979, p. 645) left over from Hume’s
failure to solve the psychological part of the problem – and after reading Kant and
speculating a great deal about the rise of polyphonic music, which Popper
considers ‘the most unprecedented, original, indeed miraculous achievement of our
Western civilization, not excluding science (Popper, 1974a, p. 43).

These speculations led him to the Kantian conclusion that knowledge cannot
be a mere copy or impression of reality, and that knowledge is genetically or
psychologically a priori though not necessarily a priori valid: ‘Our theories are our
inventions; but they may be merely ill-reasoned guesses, bold conjectures, or
hypotheses. Out of these we create a world: not the real world, but our own nets in
which we try to catch the real world. This being so, then what I originally regarded
as the psychology of discovery had a basis in logic: there was no other way into the
unknown, for logical reasons.’ (Op. cit., p. 46.)

Some forty years later, in Chapter 1 of Objective Knowledge – An
Evolutionary Approach (1972), when he returned to the problem of induction and
announced his solution to it, in particular Hume’s psychological problem of
induction, Popper advanced at the same time a principle of transference, ‘What is
true in logic is true in psychology’, in order to avoid induction altogether and to
bring some logical order into such fields as psychology, scientific methodology,
history of science, and others.

What kind of principle is this?! John Watkins (1972) exclaimed a few days
after Objective Knowledge had appeared.

A possible answer to this question is that the principle has come down in
disturbed versions from Kant’s teaching via two of Popper’s teachers: Carl Stumpf
(1892, p. 481f.) who unlike Kant, and apparently under the influence of the
growing positivist Zeitgeist, attached a refereeing role to psychology when he
stated: ‘Something cannot be epistemologically true and [at the same time] psychologically false’ – and Karl Bühler (1922, p. 212) who safeguarded himself with the statement: ‘I take side with Stumpf: what is true in logic cannot be false in psychology, and vice versa.’ (I am indebted to my friend Troels Eggers Hansen for the references to these quotations.)

The fate of Popper’s ‘principle of transference’ seems so far to have been as undecided as its origin may have been. Popper appears in line with Kant’s ‘metaphysical deliberations’ when he underlines that it is a principle of transference, something like the regulative idea of truth content: truth content of statements being transferred from the domain of logic to domains of science. Bühler’s statement, ‘what is true in logic cannot be false in psychology’, attractive as it may sound to a Popperian like Professor Tom Settle (1981), is nevertheless not a principle of transference.

As has been the case with Popper’s ‘principle of transference’, his ideas about problem-solving and learning by action and selection have not only taken long to emerge but have apparently not been applied enough to cases where they could have made a difference by clearing things up. However, it can be gauged from the selectionist theory of brain functioning developed by Professor Gerald M. Edelman in works like Neural Darwinism – The Theory of Neural Group Selection (1987) and The Remembered Present – A Biological Theory of Consciousness (1989) that the significance of selection in learning and cognition may, nevertheless, have started to dawn upon scientists in these domains. In the present historical context it is of particular interest to reflect upon the possibility that Edelman’s important contribution to brain science and learning has been inspired by discussions Edelman might have had with Popper on the occasion of the Conference on ‘Problems of Reduction in Biology’ at Villa Serbelloni, Bellagio, Italy, in September 1972. If we are to believe a remark made by Dr. Oliver Sacks (1993, p. 43) when he reviewed Edelman’s Bright Air. Brilliant Fire – On the Matter of the Mind (1992) then Edelman, after the Conference, drafted the first version of his brain theory ‘in a single sitting, during a thirteen-hour wait for a plane in the Milan airport’. This version appeared as part two of a joint publication with Vernon B. Mountcastle entitled The Mindful Brain: Cortical Organization and the Group Selective Theory of Higher Brain Function (1978).

Of course, Edelman had been aware of the importance of selection for biological processes, especially those connected with phylogenetic changes. But despite the fact that he had heard, in 1966, the lecture by the Danish immunologist Niels Kaj Jerne (1967), which had enthused Popper, on the analogy between the immune system and the nervous system with respect to a selectionist theory of learning, it was apparently only after the Serbelloni-conference that he urgently embarked on the development of a theory of brain functions and consciousness in which selection plays an important role. In a paper given at the Conference, Edelman (1974, p. 54) had made a step in that direction, realising that selection
could work on short-term phenomena, such as ‘clonal selection’: ‘Clonal selection is a fixed system with variations on a theme; in natural selection completely new themes are possible (although relatively rare) and the system is open-ended. Above all, at least for the basic protein structure of antibodies, natural selection is obviously required before clonal selection can operate.’ From there it was only a short step from the assumption that selection might also work on short-lived nerve cells as he then launched in his theory on neural group selection and re-entrant cortical integration, inspired in various ways by Jerne’s and Popper’s ideas.

7. REPLACING PARADIGMS BY THE SEARCH FOR TRUTH

All this holds great promise for our understanding of learning and consciousness. Yet there may be a long way to go before these new ideas will have replaced the old paradigms in these domains, since in many laboratories around the world rats and mice are still seen to run according to behaviouristic principles learning by repetition and induction! The fact that not a single experiment from Thorndike onwards, within North American laboratories, has ever examined the disappearance of a previously learnt response (Boakes, 1984, p. 223) seems to have branded any approach to behavioural selection by error-elimination as alien to this whole tradition, conveniently ignoring observations and innovations which research in Europe have contributed since Gestalt-psychology and ethology entered the scene.

Somehow the situation resembles that of cybernetics and information technology, which operate with confused ideas about information impossible to replace by truly better ones, as it also seems to be pretty late to hope for a replacement of the untenable behaviourist tenets of learning by principles that can stand up to logical scrutiny and survive it. Nevertheless, in the name of progress, ‘paradigmatically correct’ research should be replaced by a critical search for truth, even in the face of the impossible, and here Karl Popper’s lifetime work can serve as a helpful searchlight.

BIBLIOGRAPHY


Watkins, J. (1972), *Personal communication*.