Classic Studies in Psychology

by Dr. Stephen M. Williams
(Author Note at End of Book)

This book is dedicated to the pioneers whose work is described.

Table of Contents

Preface

This is a collection of notes on classic studies in Psychology. There are three studies for each of seven areas of Psychology:

- A- *About Psychology*
- B- *Biological Psychology*
- C- *Cognitive Psychology*
- D- *Research Issues*
- E- *Age (Developmental Psychology)*
- F- *Differences between people (Abnormal Psychology)*
- G- *Social Psychology*

Each of the seven sections has exercise questions at the end of the section.

I was an examiner 1990 – 2020 for Boards such as AQA and OCR (GCE A-Level) and International Baccalaureate (Diploma Level). These notes would give you a head start in such examinations.

The selection of a set of classic studies gives scope for choice. What strikes me now about my selection is that though I spent the whole of my career in Britain, all but one of these studies is by Americans. This reflects the great predominance within World Psychology so far of American Psychology.

There is always something to be gained by consulting the original study, and a full bibliographic reference is given for each—one factor governing the choice of studies was their ready availability through Inter-Library Loans. Some can also be found online through *Google Scholar*.

A1: The emerging field of the sociology of psychological knowledge

American Psychologist 30, pp. 988-1002

Allan.R. BUSS (1975)

Psychology is a field of study that calls on us to look at ourselves, the students. It is a further extension of this idea to call on us to look at ourselves *as* students, as Psychologists, to look at the social rules governing fully-fledged Psychology researchers, in a project of self-analysis. Buss presumed a bit by calling this an emerging field for it has still, in 2014, to emerge fully.

This field is located within the wider field of sociology of science, which is located within the yet wider field of sociology of knowledge generally. Since the work of Karl Marx, Max Weber and Karl Mannheim, the idea has been around that human thought is linked with the social context of the thinker.

Marx's class analysis of society maintained that the ideas of the ruling class become dominant, and, naturally, these ideas legitimate a stratified (class-layered) society. Weber is best known for showing how Christian Protestantism moulded Western society by way of promoting capitalism. Mannheim believed that *all* knowledge is shaped socially, so the objective picture involves considering all the perspectives of the diverse groups within society, which inevitably are parts and not the whole of the truth. Mannheim's views have been held to undermine respect for truth and so promote cynicism. He himself believed his outlook had been made feasible by the historical trends towards greater democracy and disrupting the intellectual monopoly of the Church.

Buss's article distinguishes between the macro-sociology of society in general and the micro-sociology of professional academics. He wants his new field to embrace both. He looks to the existing speciality of Intellectual History, which already teaches that fact and value are not independent. So, for example, research that is mechanistic rather than humanistic, deterministic rather than taking account of free will and atomistic rather than holistic will lead to mechanistic, deterministic and atomistic views of man. In other words, it will see her as a machine, a bundle of mental "faculties".

Buss wants to study professional societies, such as the British Psychological Society and the American Psychological Association. Both these have seen breakaway movements which emphasize an image of Psychology as science. In Britain, this is the Experimental Psychology Society, and in America, the Psychonomic Society. Another topic for sociology of Psychology would be the funding of research. It is sometimes bequests by wealthy people that energize branches of Psychology, as when Arthur Koestler endowed a chair in Parapsychology at Edinburgh University.

Ray Over is a psychologist who has done some work that could be called sociology of Psychology. He has looked at manpower planning in the academic job market, which includes the question of the need for new Master's programmes, especially during times of economic recession. He has also considered the relationship of research productivity to the age and gender of the researcher, as well as the whole issue of co-authorship.

Buss lists some more topics to be covered in his emerging field:

- peer-reviewing for journals - the existence of "invisible colleges"
- minorities, such as ethnic minorities, in academe

- hiring procedures - the existence of the old boy/girl network
- the original emergence and institutionalization of Psychology itself
- feedback to academics from applied Psychologists
- the influence of Psychology upon politicians
- the intelligibility and impact of Psychology for lay people.
-

Buss then goes on to discuss the specific branch of Psychology known as Individual Differences (see Section F of this book). He holds this study to be politically right-wing. He says it originated in the growth of capitalism with the new emphasis upon the division of labour, that is, specialization. Associated with nineteenth-century capitalism was Liberalism, a philosophy that claimed that society now gave people freedom. So their different class positions within society must reflect inborn differences. Belief in inborn differences that could be measured by psychological tests legitimized American immigration rules of a racist character.

The view that a belief in genetic difference is right-wing was challenged by Hans Eysenck. He asserted the existence of differences in ability is accepted even by Communism, with its slogan “From each according to his ability”. Though Lysenko had great success, in Stalin’s Russia, in outlawing genetic research, Eysenck claimed genetic studies of twins’ intelligence still went on behind the Iron Curtain. But is Russia a left-wing political structure? Even under Communism, “organization men” gained huge power.

Buss discusses further specific branches of Psychology from a sociological perspective. He relates the rise of Humanistic Psychology to other features of the 1960s - civil rights movements, anti-war protests and demands for university reform. Within Developmental Psychology quantitative measures of learning potential such as IQ, that foster competition and pragmatism are related to capitalism. He also sees fact-centred Behaviourism as elitist, conservative and totalitarian.

Buss concludes with a few questions about Psychology for sociological treatment. Are applied psychologists such as clinical psychologists time-servers keeping the *status quo* on the road? Or are they rather natural progressives because of their constant contact with the adverse consequences of the *status quo*? Should a statement be viewed as true because it promotes unselfishness rather than because it corresponds to a fact? Will heightening Psychologists’ awareness of their social context merely invite them to “ride the tide” of fashion? Is Thomas Kuhn’s idea about revolutions in science akin to Karl Marx’s idea about revolutions in society? Is the intelligentsia divided by tension along political left-right lines? Is it unusual in this respect (thus, the police might be thought to be on the right)?

If you find these ideas interesting, you might enjoy:
ANOTHER ARTICLE BY ALLAN BUSS (1978)
The structure of psychological revolutions. *Journal of the History of the Behavioural Sciences 14*, 57-64.

A2: Social Psychology as History

Journal of Personality and Social Psychology 26, pp. 309-320

Kenneth J GERGEN (1973)

"Positivism" is the view that facts are everything. The struggle with it has been a continuing one in Psychology. Its opponents have been able to show it up as even more ideological in Social Psychology than in other branches. One important part of the onslaught was the present article, which argues that the timeless laws of social psychology that positivists are seeking are a mirage —any theory is a child of its time. But we need to cover two articles, for Barry Schlenker made a swift rebuttal of Gergen's idea just a year later in the same journal.

If Social Psychology is a fact-accumulating science, it should be making *progress*. Yet Gergen argues persuasively that it is not, thus he says: "Knowledge [about human interaction] cannot accumulate in the usual scientific sense because such knowledge does not generally transcend its historical boundaries". Schlenker criticizes this view strongly: "This is analogous to claiming that no universal theories could have been developed in the natural sciences because ice changes into water ... or dinosaurs are no longer with us". As examples of social theories that are timelessly valid, Schlenker gives those known as social learning, social facilitation, social comparison and mere exposure. He also argues that if social processes were as transient as Gergen believes, it would be difficult to explain why the writings of Aristotle, Plato, Rousseau, Hobbes, *etc*, are still used.

There are further respects in which Gergen holds his discipline differs from science. He says, "The recipient of [Social Psychology] is...provided with dual messages: Messages that dispassionately describe what appears to be, and those which subtly prescribe what is desirable". For example to be categorized as an "introvert" is an implicit prod to become more outgoing. Similarly, Gergen puts Skinner's "scientific Law of Reinforcement" into a common sense perspective: Parents are accustomed to using direct rewards to influence the behaviour of their children. Over time, children become aware of the adult's premise that the reward will achieve the desired results and become obstinate.

Gergen makes a third, telling point. The goal of science is often said to be prediction, yet "In the same way that a military strategist lays himself open to defeat when his actions become predictable, an organizational official can be taken advantage of by his inferiors and wives manipulated by errant husbands when their behaviour patterns are reliable". Schlenker in his rebuttal argues this very "poverty of predictability" is a timeless truth. But isn't this like saying that the only timeless truth in Social Psychology is that it isn't a science?

Schlenker has described Gergen's position as being that the inborn/instinctual can be scientifically explained, while the learned and therefore cognitively modified cannot.

Social Psychology is perhaps the hardest branch to regard as science but the issue arises throughout the discipline, as discussed in *Psychology on the Couch* (S.M. Williams, 1988, Harvester Press). Freud had a controversial idea that women from early girlhood feel inferior to men. He called this "penis envy". Early psychologists developed what might be called "physics envy", believing they could only be respectable if they modelled themselves on physicists. The founding father himself, Wilhelm Wundt, acknowledged that there is an unscientific part of Psychology and wrote his huge *Völkerpsychologie* to make a start on it. Early Psychology in America, on the other hand, at least as represented by EB Titchener, discarded this side of Wundt's ideas. The more recent American, Fred Skinner, has been unequivocal: "The ease with which mentalistic explanations can be invented on the spot is perhaps the best gauge of how little attention we should pay to them ... It is science or nothing".

The Gergen-Schlenker debate has taken in further participants and I cannot convey in this space the richness of the original articles. Surely Gergen has indicated a vital difference between Social Psychology and traditional forms of scientific inquiry. It is because it is so hard to treat scientifically, that psychological study has neglected the social. Throwing the baby out with the bathwater …

In case you want the reference for:
SCHLENKER'S (1974) REBUTTAL OF GERGEN
Social Psychology and Science. *Journal of Personality and Social Psychology 29,* 1-15.

A3: The nature and limits of psychological knowledge: lessons of a century qua "science"

American Psychologist 36, pp. 257-269

Sigmund KOCH (1981)

Koch himself was once, in his own words, "a dauntless and virile rat-runner", before his horizons expanded. This expression means that he tested rats for their learning, as part of a behaviourist approach to Psychology. Behaviourism insists on third-person external observer accounts rather than first-person ones. This insistence is associated with the scientific approach to gaining knowledge, which stresses the public and the verifiable - and so, observable behaviour. It has been said that tackling behaviour is the *limit* of the scientific method. But by restricting the range of researchable mental operations to those relevant to establishing *verifiable* knowledge, behaviourism seems doomed to a lack of comprehensiveness.

The behaviourist approach was already advocated in a book by Galileo - *The Assayer* (1623); but early modern Psychology (that is, Psychology in the period starting in the second half of the nineteenth century) concentrated upon a "phenomenological" approach in which the investigator looks inward, the "method of introspection". Introspectionist Psychology proved incapable of achieving consensual agreement between different investigators; and a reaction against it, starting about the second decade of the twentieth century, was led by John Watson under the banner of "behaviourism". What Koch feels is that behaviourism itself has now been given a hearing for more than long enough. He says that the whole period since Watson is a history of behaviourism renouncing its central principle without saying so.

Clark Hull (1884-1952) was a pivotal figure in giving a new lease of life to behaviourism. He sought to deduce testable hypotheses about behaviour from a set of fundamental axioms, in the manner of Euclid's geometry. But his whole career, says Koch, was a demonstration that this approach is not feasible in Psychology. Even faithful members of the behaviourist school began to use the word "model" after the damage Hullian Theory did to the word "theory".

Koch believes the next word they should drop is "behaviour", which is over-abstract and encourages over-generalization. It shares these characteristics with "stimulus", "response" and "organism". The behaviourists were tired of the word "mind" as used by the older school, yet in practice, they only studied the sorts of behaviour that are produced by minds - they could not escape the word.

Watson and Hull and the younger Koch felt that they had the methods (of science), and knowledge would follow as an almost automatic consequence. But reading their publications, the older Koch wonders. At the end of the nineteenth century, science was the great success story, and throughout the world of learning researchers sought to adapt its methods to their fields. But the method came to precede the content - there was a retreat from the subject matter. Research hid away in those delimited areas that did seem to be yielding stable relationships of variables.

University Departments of Physics came about because of the successes of physics. University Departments of Psychology came about by edict, without proving themselves (without establishing new knowledge). But many forget this historical difference and assume publications in Psychology are as authoritative as those in Physics.

Moreover, the history of Psychology does not show visible *progress*, with new knowledge building and accumulating - rather, the efforts of the past drop away and are replaced by later fashion. All we know as psychologists is where we went wrong - that Hull's hypothetico-deductive approach does not work, and that the memory trace of a stimulus is not located within any specific region of the nervous system (the work of Karl

Lashley). But does this criticism merely reveal that expectations of Psychology have been too great? So long as Psychology holds our interest, does it matter that nothing much is being built?

What matters most to Koch is the role of scientific method in Psychology. He believes there is a strong chance that at some critical point of system-openness, boundary weakness, or mere internal complexity, scientific analysis ceases to work. What happens instead is imitation science, albeit this is a highly sophisticated skill.

Moreover, treating people as “subjects” or “participants” in an experiment is treating them like things, in a way that demeans and diminishes all concerned. Perhaps this is part of the appeal of experiments with animals rather than with people to investigators such as the young Koch. But should the lesson not be rather to avoid experiments altogether, to watch people in their normal environments much more?

With all his doubts and soul-searching, Koch does not write off Psychology completely. He does believe that there is a place for empirical methods, statistical analysis and mathematical models. Some branches of Psychology can even be regarded as parts of science. Koch himself edited a handbook of Psychology in many volumes. He is a radical critic, but a constructive one.

Koch concludes that the lesson of the last century has been that “extensive and important sectors of psychological study require modes of inquiry rather more like those of the humanities than the sciences. And among these I would include areas traditionally considered “fundamental” - like perception, cognition, motivation and learning, as well as such more obviously rarefied fields as social psychology, psycho-pathology, personality, aesthetics and the analysis of 'creativity'”.

If you are interested in these ideas, you might enjoy:

ANOTHER ARTICLE BY SIGMUND KOCH

Psychology as science. In S.C. Brown (Ed.) *Philosophy of Psychology*, Macmillan, 1974.

EXERCISES FOR SECTION A "ABOUT PSYCHOLOGY"

1 Is Psychology a science? If not, can it ever be one? Should it be one?

2 Identify some promises and pitfalls of a career in Psychology.

3 Does science identify causes? Is there any other way to do this than with an experiment? How then can it be known that carbon is the cause of global climate change?

B1: Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain

Journal of Comparative and Physiological Psychology 47, pp. 419-427

James OLDS and Peter MILNER (1954)

"Reinforcement" in the article title means that this stimulation will increase the frequency of the behaviour that precedes it, acting as some kind of *reward* for it.

This article by Olds and Milner is squarely in the behaviourist camp and depends on the assumption that what we learn about the brains of rats will help us to understand those of humans. Such knowledge can help in the understanding and treatment of neurological disease, though part of the motivation for doing this sort of research is pure scientific curiosity.

The experimenters implanted electrodes into the brains of rats, who were observed in the experimental apparatus known as a "Skinner box", set up so that electrical current was delivered to the brain so long as a lever was pressed. The rat thus stimulated its own brain by pressing the lever, and so this field of research is generally known as "electrical self-stimulation of the brain" or ESB.

The use of live animals for research, particularly for direct interventions into the brain such as this, has aroused great controversy. Some people feel that the medical and scientific benefits do not outweigh the harm done to the experimental animals.

However, the fifteen rats participating were killed at the end of behavioural testing, as is typical in experiments with rats. In this experiment, it was necessary to "sacrifice" (to use the standard terminology) the animals, to verify by microscopic anatomical examination of the brains where exactly the electrode tips had been placed.

Each electrode was a pair of wires cemented together in the form of a single needle. The needle was held still on the skull using a flange. Holes were drilled into the skull for the needle and four fixing screws for the flange. The operation of electrode implantation was performed under anaesthetic and the animals were given three days to recover, before ESB testing.

For this ESB testing, the voltage of the electricity was set initially at half a volt. If this did not produce any effect on the rat's behaviour, the voltage could be raised to up to five volts.

The main finding of the experiment concerned four rats that were discovered on *post-mortem* examination to have been stimulated in the "septal" area, part of the limbic system of the brain that lies under the cortex of the cerebral hemispheres. During ESB testing, these rats all pressed the lever in the Skinner box very frequently. The frequency was between 3,000 and 7,500 times during the total twelve hours (spread over four days) that the rat was in the experimental set-up.

The strength of this sort of reinforcer has also been demonstrated by other later research. This later research showed that male rats will stimulate themselves using the press of a lever *in preference to* eating, drinking or having access to a sexually receptive female. These rats had been left hungry and thirsty. The area of the brain into which electrodes had been implanted was another one called the "lateral hypothalamus". But more research has shown that it is yet another area of the brain, called the "medial forebrain bundle" that provokes ESB most reliably. This has often been described as a "pleasure centre of the brain".

So the ESB reinforcing effect depends on the area of the brain stimulated. Olds and Milner themselves found that other animals, apart from the septal four, with different electrode sites, would not press the lever for electrical stimulation. Since some of these rats *would* press the lever, though less intensively when it did *not* cause electrical stimulation, this article suggested that there are "pain centres" as well as pleasure centres.

The rats would avoid pressing the lever when it was set up to administer a current to these areas, even though they would press with moderate frequency at other times.

Researchers today believe that electrical stimulation of the medial forebrain bundle is reinforcing because it activates the same system that is activated by natural reinforcers such as food, water and sex. The outside possibility that the electrode implantation is unpleasant in itself and ESB merely produces an alleviation has been excluded - the lever pressing represents the pursuit of a strongly rewarding stimulus.

Such research has potential for misapplication or at least premature application. There have been reports of the use of stimulation through electrodes of pleasure centres of people diagnosed as schizophrenic or mentally retarded. These reports came from the US and were not continued into the 1980s.

B2: Hemisphere deconnection and unity in conscious awareness

American Psychologist 23, pp. 723-733

Roger W. SPERRY (1968)

In 1981 Roger Sperry was awarded the Nobel Prize in Physiology/Medicine. The award (which was shared with researchers in another field) was above all for the work described in this article, though Sperry's citation does also mention his earlier work on salamander vision. A leading collaborator with Sperry on this deconnection work has been Michael Gazzaniga.

The nature of this work has been psychological testing of people who have had a certain sort of brain operation for the treatment of *epilepsy*. This disease of "fits" (or "convulsions" or "seizures") has been described since ancient times, but only much more recently has it been accepted that epilepsy is caused by abnormal electrical activity in the brain of the sufferer. This activity can be observed using the electroencephalogram.

Epilepsy can be an extremely serious disease. The situation is life-threatening if someone has a sudden fit while crossing a road. But also, with the progress of the disease, seizures can start to recur and continue in a condition known as *status epilepticus* which carries a risk of fatality. Pharmaceutical medication is the first line of defence in controlling epileptic seizures, but for people who have severe symptoms, there has been neuro-surgery.

One widespread form of neuro-surgical operation for epilepsy is called "temporal lobectomy" and involves the removal of grey matter in a temporal lobe of the brain, which has been identified as the *focus* of the epileptogenic electrical activity. But the people whom Sperry studied had had another operation, called "commissurotomy". This involves severing white matter called commissures, that connect the two cerebral hemispheres of the brain, to contain the epileptogenic electrical activity within one hemisphere.

There is more than one cerebral commissure. The main one is known as the *corpus callosum*, but there are smaller anterior and hippocampal commissures, as well as a structure lower down called the *massa intermedia*. Not all patients have had all commissures severed.

The operation of commissurotomy has been carried out since the 1940s at least, and naturally, doctors have been very concerned for many years about any possible unwanted effects from it. The prevailing view at first was that of Akelaitis and his collaborator Smith, who said there were no serious unwanted effects, so long as the operation spared non-commissural tissue. Sperry's work has had the consequence of revising this assessment completely.

Sperry used new forms of behavioural testing on these patients after their operation. His conclusion from these tests was that the usual sense we have of ourselves as a single consciousness has been altered in these patients. It is as though they now have two separate consciousnesses, one in each of the cerebral hemispheres.

Sperry's new tests depended upon anatomical and physiological knowledge about the "afferent" (sensory) and "efferent" (motor) nerves.

These nerves typically occur as a *pair*, one on the right and the other on the left. The anatomical picture is complicated, and different in detail for each type of nerve, such as the nerves for the eye, ear and finger. Some nerves travel back to the same side of the brain, but many "decussate", that is, they cross over from one side (right or left) to the other. In vision, the lens of the eye also makes information cross "contra-laterally", from one side of the outside world to the other side of the eye's retina.

There is a broad picture, though, which is that each side of the brain, right or left, has a special relationship with the *opposite* side of personal space.

When the two cerebral hemispheres are disconnected using the surgery, information no longer passes *so readily* from one side to the other (of course, the left and right sides of the central nervous system are still connected at lower levels). Sensory input to the right hemisphere will not pass to the left hemisphere. Since it is the left hemisphere that sends motor commands to the fingers of the right hand, these patients cannot point with the right hand to a copy of a picture presented in the left visual field. Similar phenomena are observed with every possible combination of lateralized input and output.

Another very interesting type of observation was made by Sperry: the patient cannot *name* an object presented to the right hemisphere, say through lateralized left-visual-field input. This accords with more than a century of neurological research indicating the left hemisphere has a special responsibility for some verbal functions.

Real life is never quite as simple as explanatory statements indicate, and reservations and qualifications have to be emphasized. Sperry's testing was of a very limited number of patients - he mentions eleven in this article - and even fewer, perhaps only two, showed a completely clean "disconnection syndrome".

It is in that light that we have to take the more anecdotal reports about such patients, such as the ones that the two consciousnesses of a patient can come into *conflict*, as when one arm is seen to restrain the other.

Behaviour consonant with physical deconnection, known in patients as the disconnection syndrome, was observed first in non-human animals operated upon with the same form of commissurotomy. Because of the additional problems of testing languageless animals, an additional operation on the optic nerve was required.

The claim to be able to test each hemisphere independently has led to a research industry with commissurotomized patients, trying to describe the differences between what each hemisphere does, as with the assertion that the left hemisphere is more verbal. Several writers have been tempted to draw conclusions that go way beyond the empirical evidence, a notable culprit being Ornstein in his book *The Psychology of Consciousness*. It is important not to allow the interest, indeed fascination, of such ideas to divert attention from the overriding need to relieve epilepsy and to assess the efficacy of one neuro-surgical operation with this aim.

If you want another source for this work then

A LONGER ARTICLE BY SPERRY DESCRIBING HIS SPLIT-BRAIN WORK IS

Mental unity following surgical disconnection of the cerebral hemispheres. In *The Harvey Lectures (1966-1967)* Series 62, Academic Press, pp. 293-322.

and AN IMPORTANT NOTE RELATED TO KIMURA'S ARTICLE IN SECTION C IS

Lateralized suppression of dichotically presented digits after commissural section in man.

Science 161 pp. 184-186, by *Brenda Milner, L Taylor and Roger Sperry (1968)*.

B3: Review of the book <u>Verbal Behavior</u> by B.F. Skinner

Language 35, pp. 26-58

Noam CHOMSKY (1959)

Too often Psychology is assumed to be a calm mill-pond. On the contrary, *controversy* pervades the discipline, and one place to see it readily is in reviews of major books.

Perhaps the most famous of such reviews is this lengthy one by Chomsky, a demolition of Skinner's effort to apply his behaviourist ideas to the understanding of *verbal* behaviour.

Chomsky is a noted theoretician in the field of Linguistics. His book *Aspects of the Theory of Syntax* (1965) had great influence, and penetrated the developing branch of Psychology known as Cognitive Psychology, to create a new field known as Psycholinguistics. He has also made a considerable reputation as a critic of American government policy - for example with his *Why are we in Vietnam?* These political writings are collected in his *Chronicles of Dissent* (1992).

Skinner's book originates in his doctrine that behaviour is shaped by the "contingencies" upon responses, that is, what happens as a result of them. He designed a box for experiments with rats, which has been named after him. The box has a lever and a tray for food in it, and it is automated so that the arrival in the tray of a pellet of food is made contingent upon a press of the lever.

The experimenter can also make the contingency a complicated one known as a "schedule". Thus, every five presses can be rewarded with a pellet (a fixed-ratio schedule), or else the first press after a five-minute wait (a fixed-interval schedule). These numbers can also be treated as *variables* with an average value (variable ratio and variable interval schedules). These four primary schedules can also be yoked together to form "conjunctive" schedules. Skinner observed with great care how lever-pressing behaviour changes consequent upon the schedule in force.

Skinner went on in books like *Beyond Freedom and Dignity* (1972) to make huge claims for the desirability of a society in which contingencies are set up to eliminate deviant behaviour like crime. But the whole outlook ignores the everyday fact that people *resent* the use of rewards to influence their behaviour. They see it as manipulative. The resentment is even greater for the use of punishment, which is seen as arrogant and cruel. Nonetheless, Skinner made a huge reputation within Psychology, outstripping Freud himself on citation count, even before he died in 1990. He published several autobiographical works, such as *Particulars of My Life* and *Cumulative Record*.

It is about time to say something about *Verbal Behavior*. According to Chomsky the goal in this book is to provide a way to predict and control verbal behaviour by observing and manipulating the physical environment of the speaker. What is held to be of overwhelming importance for verbal behaviour are the *inputs*. This means both the present stimulation and the history of *reinforcement* - the frequency, arrangement and withholding of reinforcing stimuli.

What Chomsky himself, on the other hand, holds to be fundamental in the study of language is its internal structure (or grammar). The verbal organism is not a "black box" between stimuli and responses but rather makes a complex *contribution* to language learning and performance. And if we think about it we will see that the input or stimulus is not, really, describable in separation from the organism's responses. It only *counts* as a stimulus if it evokes a response.

What Chomsky charges against *Verbal Behavior* is that it is nothing but a form of ritual:

"Although Skinner's concepts, such as reinforcement, are defined in a precise enough way in the laboratory with rats and pigeons, their meaning becomes crass and undifferentiated when applied to human social relations. Most damaging of all, the supposedly precise re-formulations of common-sense language are often only rescued from either incoherence or falsity by vague and metaphorical interpretations of concepts such as reinforcement; hence these terms become merely a poor substitute for ordinary usage”.

This thesis about the book is elaborated at considerable length. Thus, Skinner says that a response like the utterance of a proper noun is “under the control of a specific person or thing (the stimulus)” - the presence of the stimulus increases the probability of the response. Think about it for a moment, says Chomsky, and this can be seen to be untrue, otherwise people would go around saying their own names all the time. As grammarians have always said, the proper noun *denotes* the person - to say that the person *controls* the noun adds nothing.

Skinner is also caught in a tangle when he tries to define the *strength* of a verbal response. All of us know that it is *not* the loudness, pitch and lack of delay of an utterance that gives it weight. All of us with the possible exception of strident behaviourists.

Skinner is quite wrong as well to say that children learn their language only through reinforcement by their parents. Far more comes from casual observation and imitation. Chomsky also refers to the phenomenon of “imprinting” in animals, which suggests a role for genetic factors not just environmental inputs in learning. We learn English rather than Mandarin Chinese, though, and a large vocabulary of English words. It would be stretching the case too far to say that the childhood environment has no role.

Skinner once asserted that he neither answered nor even read his critics since he had better things to do with his time than to clear up their misunderstandings. He concentrated on cultivating enthusiastic followers, and if you stay with Psychology, you will undoubtedly meet some.

EXERCISES FOR SECTION B “LIFE: BIOLOGICAL PSYCHOLOGY—PEOPLE AS ANIMALS”

1 Is it valid to compare human beings with other animals?

2 What is meant by saying that the brain is the organ of the mind?

3 Is the subject matter of psychology behaviour rather than the mind?

C1: a. Cerebral dominance and the perception of verbal stimuli.
b. Some effects of temporal-lobe damage on auditory perception.

Canadian Journal of Psychology 15, pp. 166-171, 156-165

Doreen KIMURA (1961)

These two articles printed consecutively in this leading journal had an immediate and lasting impact on the Psychology research community. They are “citation classics”.

The findings reported swiftly drew a re-interpretation from Inglis (see below), the following year in the same journal, a re-interpretation that Kimura rebutted flatly. Kimura went on to establish a formidable reputation in this field of “cerebral dominance”, with other seminal articles.

The phrase “cerebral dominance” is an old one, referring to the long-standing belief of clinical neurologists that one of the cerebra - the left one - is dominant over the other. Nowadays, most researchers in the field would be careful to specify that the dominance is selectively for a particular function of the brain - speech. We prefer terms like “hemispheric specialization” and “hemispheric asymmetry”.

Moreover, Kimura’s articles strove to illuminate hemispheric specialization *indirectly*, employing behavioural studies using normal subjects, not just neurological patients. Widespread concern over whether the logic of these indirect inferences is sound has prompted yet another re-naming of the field, as “laterality”. A journal with this title began publication in 1996.

Kimura’s 1961 studies took a new angle on some work of Donald Broadbent. That involved recording independently on channels 1 and 2 of a stereophonic tape recorder. Thus, simultaneous messages could be presented, over headphones, one to the left ear and one to the right ear, an experimental paradigm that has become known as “dichotic listening”. Broadbent was interested in the subjects’ *order of report*. In reporting back what they have heard, do they group stimuli by the time of arrival (i.e., simultaneous stimuli together) or by the ear of entry (first all the stimuli to the left ear, followed by all those to the right ear, or vice versa)? Kimura was more interested in a completely separate issue - the fact that subjects do better in reporting speech presented to the *right* ear - the right-ear advantage or REA as it has come to be known.

This observation was parallel with a 1952 report by Mishkin and Forgays that subjects report writing better from the right visual half-field. But far more so than the visual half-field asymmetries, the REA has proven to be a very *reliable* and *easily reproducible* effect.

Based as she was at the Montreal Neurological Institute, Kimura also used the opportunity to observe dichotic listening by patients with neurological damage (in various locations within the brain).

These patients had had the sodium amytal test to identify which hemisphere was responsible for speech. For most, it was the left hemisphere, and these people showed the usual REA, but for a few, it was the right hemisphere, and these people showed a *left* ear advantage. She concluded that the side of the ear advantage may be a way of diagnosing which is the dominant hemisphere. Dichotic listening is not physically invasive, in the way sodium amytal injection is, and may be less traumatic. Ever since, a debate has been continuing whether she was correct.

Her conclusion is shaky for the following reason. The patients with right hemisphere dominance also tended to have lesions in the left hemisphere. So poor performance on the right ear would naturally follow from the principle that each hemisphere represents mainly the *contra-lateral* half of space. Kimura herself showed that performance in dichotic listening on the ear contra-lateral to temporal-lobe damage drops following surgery on the lobe. Kimura argued against the view that the LEA in right-hemisphere-dominant patients is just a “lesion effect”, but there were so few patients involved that it was unconvincing.

Let me return to Inglis's reaction to Kimura's articles. He took up Broadbent's original observations on order of report. Broadbent's usual presentation for a trial was three successive pairs of digits. With appropriate intervals between the digits, subjects adopt what he called the "ear order of report", giving all the digits to one ear, more often the right, before all those presented to the other. Subjects can be told which ear to report first - so long as it is right and left equally often, an REA still appears. With this controlled order of report, it is possible to compare right and left both for the ear reported immediately and for the one reported with the inevitable delay. What Inglis said was that the REA is mainly for the delayed ear. Subsequent research has not borne out this claim with any consistency. The evidence for the Inglis hypothesis that the REA is more a storage-memory phenomenon than a perceptual one is weak.

Kimura in her 1962 rebuttal of Inglis said that she used a variant of the Broadbent paradigm in which subjects tend not to use an ear order of report, and report order could therefore not be controlled. She denied that a larger ear advantage for delayed than for immediate report would indicate it was storage rather than perception.

She also held this distinction to be a largely verbal one anyway. At the same time, she acknowledged the importance of detecting memory impairments in people with brain damage, which seems contradictory. The distinction between perceiving in a second and remembering over ten years is not a verbal one. Memory impairments require more subtle means of detection than perceptual ones, but they are nonetheless important for the patient and her rehabilitation.

I have pursued this distinction in research of my own, and believe that the REA is indeed a largely perceptual phenomenon, carrying with it little hope of permitting diagnosis of impairments of memory and understanding of speech. The hope of progress in this direction is to find a way to demonstrate REA with ordinary "monaural" stimulation. Great effort has gone into this sort of experiment, without, so far, reliable results.

There are other huge problems with the clinical use of dichotic listening. Most people are right-handed, and pick up a phone with one hand and put it to one ear - these facts cannot be ignored in the interpretation of the REA, though they were omitted from the neuro-psychological literature for twenty years. Moreover, the REA was found with opportunity samples of college students; in community samples, the middle-aged group may show an LEA, just as they show a left-ear preference for the telephone.

To end on a positive note, though. The lateral asymmetry also appears when the two dichotic messages are played back over speakers rather than headphones. This makes group testing feasible, and the REA can be demonstrated in a single session of one hour. The prospect is there for much more exploration of the various issues, based on experimental observation.

AN INFLUENTIAL THEORETICAL ARTICLE BY KIMURA (1967):

Functional asymmetry of the brain in dichotic listening. *Cortex 3*, pp. 163-178.

A CRITIQUE OF HER THEORY:

Advantage for speech from the right side of the sensory field: two facts and an interpretation. *Cognitive Systems 1*, pp. 187-205. By *S.M. Williams (1985)*

C2: Levels of processing: a framework for memory research.

Journal of Verbal Learning and Verbal Behavior 11, pp. 671-684.

Fergus CRAIK and Robert LOCKHART (1972)

The better-known of these two names is Craik, who was British before emigrating to a distinguished academic career in North America. Although he has also contributed a great deal in terms of empirical studies to *knowledge* about human memory, this article was a landmark contribution to *understanding* rather than to knowledge.

The article stands in direct line from two earlier articles, which would have been included in the present collection if it could have been longer. "The magical number 7 plus or minus 2" (1956) was an article by George Miller - a founding father of Cognitive Psychology - emphasising how there is often a *limit* to our capacity for processing information, of about seven items. "Two storage mechanisms in free recall" (1966) was an article by Murray Glanzer and Anita Cunitz identifying a separate *memory store* which is responsible for this capacity limit.

These researchers prepared lists of sixteen words (so, longer than the memory limit of seven) intended to be unrelated to one another. The lists can be read aloud, at a measured pace, to participants who are required to report the words back (in any order). Subjects can retrieve at most seven words from their "short-term store" and any more have to come from what Glanzer and Cunitz called the "long-term store". (The short-term store can be emptied, by adding a distracting task at the end of the list - the distraction *displaces* words from the short-term store. The long-term store can also be emptied, for example by speeding up the rate at which the words are read). This two-store model was enhanced by positing yet a third memory store, called the "sensory register", which, even more than the short-term store, holds information in a very *literal* form, like an echo or a photo.

The term "long-term store" is a bit of a misnomer - it is long only in the context of brief laboratory experiments. Also, the way we recognize pictures, faces, voices and melodies from long ago has some of the characteristics of the *short*-term store - it is literal, too. Putting aside this difficulty, it remains a fact that the three-store model proved extremely influential as a framework for research.

This article by Craik and Lockhart set itself the task of replacing the three-store model, which they subjected to a searching critique. The core idea of their replacement model is that a memory trace is a by-product of *perceptual* analysis. Trace persistence increases (the trace becomes "harder"), the greater the *depth* to which the stimulus has been analysed.

This "replacement model" strikes some people as no more than putting the same ideas into different words, but there is no doubt that it has altered the way memory researchers think about their findings.

Let me clarify the notion of analysing a stimulus in *greater depth* with an example: the understanding of speech. From physics, you will know that the sound of someone speaking is carried on the air as vibrations, called an "acoustic wave-form". The first task for the hearer is to identify the individual letters, or "phonemes" (as the sounds of letters are called). In English there is a restricted set of around forty phonemes, so we describe the first level of speech understanding as "phonetic categorization".

There is a lot of evidence that the trace of the waveform decays very rapidly before it is categorized as a particular phoneme. Deeper levels of understanding speech include the combination of phonemes to identify words in the hearer's vocabulary (or "mental lexicon"), the combination of words according to rules of grammar (or "parsing the sentence"), and finding a meaning for the sentence. The deeper such lexical, syntactic and semantic processing goes, the more durable the trace becomes.

One sort of experimental evidence for all this is the following. It is possible to give participants a task that requires no more than phonetic categorization of the stimulus words, a task such as *what words rhyme with (cat, cheese, etc)?* If, afterwards, they receive an unexpected test of their memory for the stimuli, they do not do particularly well. By contrast, if the task is one demanding *semantic* processing, such as whether the word is a particular type (animal, food, etc), subjects show quite good *incidental learning* of the stimulus words.

Craik and Lockhart point out that “hardening the trace” is not the *only* way we retain stimuli. We can also cycle them round and round at a single level of processing, an activity known as “rehearsing”, which involves keeping the stimuli in consciousness. But rote learning through repetition does not work unless we move from *rehearsing* to *hardening*.

The levels-of-processing framework makes comprehensible another sort of finding in experiments on free recall. It is possible to graph a *serial position curve* of performance through the list (averaging recall of words at each position). It is reliably found that subjects report the *beginning* of the list of words well. This is because they have time to process these items more deeply. Also, though the *final* items are reported well too, at first, they are only processed shallowly, and in an unexpected second test of recall, they are remembered *worse* (the “negative recency effect”).

Memory research continues as a very active field of Psychology, and you may also find interesting the “working memory” framework of Alan Baddeley and Graham Hitch.

IF YOU WOULD LIKE TO READ THE KIND OF EXPERIMENT ON WHICH CRAIK’S THEORISING IS BASED, YOU COULD TRY:

Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General 104*, pp. 268-294. By Fergus Craik and Endel Tulving.

C3: Memory scanning: new findings and current controversies

Quarterly Journal of Experimental Psychology 27, pp. 1-32

Saul STERNBERG (1975)

The three articles covered in Section A *About Psychology* all expressed serious reservations about adopting a *scientific* approach to studying the mind. Saul Sternberg, who gives as his affiliation for this article Bell Laboratories, is one worker who has succeeded more than most in making it seem an *appropriate* approach - for a particular area, anyway.

The work that we associate with Saul Sternberg's name (there is another well-known psychologist of intelligence called Robert Sternberg) has been with the following experimental paradigm. A human subject is given a set of items to memorize (often the items have been digits). The *size* of this memory set varies - in the case of digits up to a maximum of ten, of course. Next, a single digit is briefly presented - flashed on a screen. The participant is required to decide whether or not this "probe" digit is within the memory set *and to decide as quickly as possible*. Typically, she presses one of two response keys held in either hand, one key for "Yes" and the other for "No". This paradigm is generally known as "high-speed memory scanning".

The decision time (between presentation and response) is measured in milliseconds. Sternberg found this time span rises in a straight line when graphed against the size of the memory set. The finding is reliable and indicates that subjects scan through the memory set mentally one item after another. That is, the processing is *serial* rather than *parallel*.

This distinction between serial and parallel processing is clearer than many in psychology, and one that has been applied widely. For example, the left hemisphere is often said to be a serial processor, by contrast with a right hemisphere which can perform plural activities at the same time (*i.e.,* in parallel).

Certainly, Sternberg's finding is restricted to particular experimental tasks. Ulric Neisser has shown a contrasting result when subjects search through printed lists of letters for a member of a letter memory set. For this task, the size of the memory set makes no difference - a great deal of parallel processing appears to be possible. Nonetheless, many experiments have shown the serial processing effect.

There is a second basic finding in Sternberg's paradigm It is possible and natural to graph the decision times for each size of memory set *separately* for "Yes" decisions and "No" decisions. When this is done, it is reliably found that the slopes of the two lines are about the same. This, too, has an interpretation in terms of the nature of memory scanning.

This time the interpretation is a little more complicated. It goes as follows. Suppose a serial scan through the memory set terminates when the probe is successfully matched. Then this termination will occur on average halfway through the set. On the other hand, "No" responses can only be given after a complete and *exhaustive* scan of the memory set. So as a consequence, the slope of the "Yes" line should be half that of the "No" line. But the slopes are not related like this, they are the same. So the scan does *not* terminate on a successful match. Rather, the scan is exhaustive, *whether the probe is in the memory set or not.*

This finding is much less reasonable intuitively - an exhaustive scan even after a match is made wastes some processing. Nonetheless, Sternberg makes a plausible case that it might happen. A *decision* to terminate a scan takes a significant amount of time - the task, repeated many times in an experimental session, can be more highly automated by forgoing this sort of decision.

By choosing response time as the measure, experimenters can answer legitimate anxiety throughout Cognitive Psychology about the exclusive use of *accuracy* as a measure of performance. Accuracy measurement, say in the verbal memory experiments reviewed by Craik and Lockhart, involves forcing *errors* from the participant. He may have feelings of frustration about making mistakes, and these feelings will add variance or “noise” to the experimental data.

Sternberg's *chronometry* goes back some way, for example, to Franciscus Donders in the 1860s, and the research can be improved by modern equipment. Timing or “latency” measures have been used in many experiments on laterality, such as those following on from the Kimura article described earlier. The results for latency are not always in harmony with those for accuracy - intuitively it seems likely that there will be a “trade-off” sacrificing accuracy for speed. The results can be confusing, but as Psychology progresses some of the confusion is dispelled.

Other workers have replicated both Sternberg’s main results - I have done myself in the article cited below. One intuitive result, one counter-intuitive - perhaps that is the appeal.

IF YOU WOULD LIKE TO READ THE KIND OF EXPERIMENT DONE WITHIN THE STERNBERG PARADIGM, YOU COULD TRY:

A response-type reaction time effect found in the S. Sternberg high-speed memory scanning paradigm. *Acta Psychologica 75,* pp. 279-292. By Stephen Williams, Colin Cooper and John Hunter (1990).

EXERCISES SECTION C "THOUGHT: COGNITIVE PSYCHOLOGY—PEOPLE AS THINKERS"

1 Is there a difference between perception and memory?

2 Is "human information processing" a helpful description?

3 Does philosophy have anything to say to psychology?

D1: On the social psychology of the psychology experiment: With particular reference to demand characteristics and their implications.

American Psychologist 17, pp. 776-783

Martin ORNE (1962)

Experimental studies, of the kind reviewed in the last two sections, have flourished for several decades. During that period "Can I design a new experiment?" was a more important question for Psychologists than the real content issues of psychology. In colleges in America, which has been the centre of World Psychology, Introductory Psychology majors crowded into laboratory cubicles. They had to take part in experiments as part of their course requirements, and they formed the bulk of participants in experiments.

Psychology staff, with their colloquia and conferences, would act the part of the dedicated scientist in a white coat, meticulously checking the details of the apparatus. The questioning by the audience would be insistent and even hostile, as befitted the attention of scientific peers. But the questions would be such as to share implicitly the assumptions of the experimenter.

It was in the 1960s that there began a more wholesale reaction against experimental methods, and the three articles of Section D were key ones in that reaction. These objections were less easy to accommodate through a fresh statistical analysis, or with a further experiment.

Orne's point is fundamentally a simple one. He introduces it with the following observation. He once asked some casual acquaintances to do him a favour - when they said yes he asked them to do five push-ups. They reacted with amazement, incredulity, and the question "Why?". But when he asked a similar group to take part in a short *experiment* before asking them to do five push-ups, their only question was "Where?".

Agreeing to take part in a Psychology experiment is an action with meaning. Participants know that they are not to inquire about the purpose of the experiment and that they may need to tolerate boredom and even discomfort. Yet in the experiment, this active decision by the participant is forgotten, and she is treated as a passive responder to stimuli. And it is forgotten that she is bound to *wonder* about the purpose of the experiment.

Orne points out that though subjects have various reasons, such as course requirements or payment, for agreeing to take part, a major reason tends to be their belief that they are contributing to science and so ultimately to human welfare. But this gives them a *stake* in the outcome of the experiment, they want it to "work". Much practical experience shows that there exists a concern to detect and validate the experimental hypothesis - and to avoid "fouling up" the experiment.

This is precisely the reason why participants are commonly *not told* the purpose of the experiment. But to assume that they are wholly ignorant of that purpose is to ignore the many cues to that purpose that are there for them to glean. Such cues are dubbed by Orne the "demand characteristics" of the experimental situation.

These demand characteristics include information from several sources: campus rumours about the research, the information conveyed while inviting a potential subject to take part, the person of the experimenter, and the setting of the laboratory, as well as all explicit *and implicit* communication during the experiment. Thus, if a test is given twice with some treatment intervening, any participant can work out that a *change* is expected. A few participants deliberately adopt the mindset of an obedient and unquestioning servant, but many make an active attempt to respond appropriately to the totality of the experimental situation.

While the participant's detection, through demand characteristics, of the experimenter's hypothesis will lead to some concern by the participant for the experiment working, the participant may also lean over backwards to be honest in her report. But that compounds the damage - all this is going on in the participant's mind independent of the specified experimental variables and so it is a big problem for the interpretation of the experiment.

In practical terms, an experimenter mindful of the demand characteristics problem should always ask the participant at the end of the session whether he detected the purpose of the experiment. Bear in mind the convention that the participant is not told the purpose of the experiment. This question should be asked in a manner that imputes no guilt to him for having worked it out, and in a manner open to any possible response.

Orne discusses in his article the importance of demand characteristics in particular experimental paradigms, such as hypnosis studies and sensory deprivation studies.

Though posing a radical challenge to the work of the Psychology establishment, Orne himself, whose affiliation for this article is Harvard Medical School, was not treated as a maverick and has been recognized for his achievements by the American Psychological Association.

D2: Interpersonal expectancy effects: the first 345 studies

The Behavioral and Brain Sciences 3, pp. 377-386, plus open peer commentary to p. 415

Robert ROSENTHAL and Donald Rubin (1978)

My first project in Psychology was a comparison of left- with right-handers on the Wechsler Adult Intelligence Scale. This is a battery of eleven tests which take a couple of hours to administer. My supervisor suggested a sample of sixty subjects, and I shared the work with another undergraduate, Eric Wynn. We did statistical comparisons of all the tests and found some suggestive handedness differences, but the most striking effect statistically was that my subjects scored higher IQ than Eric's.

This is a form of "experimenter effect", something on which Robert Rosenthal has stamped his name, by doing a large number of systematic studies of it. Many attributes of the experimenter, such as his/her age or gender, or the wearing of a white coat, can affect the results of an experiment, but the article under review considers another particular aspect. It considers the way experimenters tend to obtain the results they *expect.*

They do not necessarily do this simply because they have made correct experimental predictions, but rather because they have helped to *shape* the results *through* their expectations. When psychologists expect certain results from their human (or animal) participants, they unwittingly treat them in such a way as to increase the probability that the subjects will respond as expected. The predictions make *themselves* true.

This is not confined to psychology experiments - it has far greater general and social importance. Teachers, employers and therapists all affect the performance of their pupils, employees and clients, by way of their expectations of them. This has been called the "Pygmalion effect", after an ancient Greek myth. A more up-to-date description would be the "My Fair Lady effect".

A 1966 study by Rosenthal went as follows. He gave an identical project to several different postgraduate students. Some of them were told the results would come out one way, others were told they would come out the opposite way. What he found was that all students had a strong tendency to produce the results they had been told would happen - contradictory results, of course. Rosenthal put this down to unintentional communication to the participants by the experimenters of their expectancies, through "paralinguistic or kinesic cues".

In a way, the experimenter expectancy effect is another illustration of the power of authority, which Stanley Milgram is best known for exploring experimentally (see *Classic Article* E2). Both Rosenthal's postgraduates and the subjects in their experiments were alert to show obedience.

Rosenthal has been criticized by Theodore Barber, for *doing* experiments to weaken the credibility of experiments. He depicts this as circular. Barber made other points. He said Rosenthal's experiments were methodologically unsound. He also said that expectancies work in other ways as well: through experimenters deviating from the stipulated procedure, mis-recording data, and even fabricating data. The last possibility evokes memories of the notorious charges against Sir Cyril Burt, which included the charge that he had fabricated data. If one of Britain's most eminent psychologists can do it, why not a student? (But the charges against Burt are still contested).

The central point that experimenter expectancies do have a role remains valid. As the title of this article by Rosenthal and Rubin suggests, the *weight* of the evidence is compelling. There is as much support for an experimenter expectancy effect as for almost any other effect in Psychology.

The 345 studies are not just a tedious re-working of the postgraduate student experiment. And they fall into eight broad categories of research: reaction time, inkblot tests, animal learning, laboratory interviews, psycho-physical judgements, learning and ability, person perception and everyday life situations. In other words, they run the full gamut of Psychology, from hard to soft.

What can Psychologists do to counter Rosenthal's problem? One precaution that seems to be indicated is to have more than one experimenter, an option requiring some bother that has rarely been adopted. More popular with the advent of microprocessors is the idea of trying to eliminate the experimenter and *automate* the experiment more and more. A problem with this is that it weakens the researcher's feel about what is going on in the experiment. Also, widespread apparatus like one-way mirrors and video cameras raise issues about the invasion of privacy.

At the end of the day, the issue is about the *competence* of the experimenter. Good experimenters should lean over backwards to avoid biased confirmation of their hypotheses. Experimental ability is a *skill* that is acquired and developed with practice. Like many skills, it cannot be fully explained in words. No individual experimenter can establish beyond doubt that she possesses it. This is true even in the natural sciences, where the variables to control are fewer.

Accusations of incompetence are rife in Psychology, and it is often no more than bad-mouthing. The "pioneers" of work on *plant emotion* dismissed some contradictory work on grounds of incompetence (they said the wrong electrode material had been used). Be careful about dismissing work as incompetent. He who is without sin should cast the first stone.

ANOTHER VERY IMPORTANT METHODOLOGICAL ARTICLE BY ROSENTHAL
The file drawer problem and tolerance for null results.
Psychological Bulletin 86, pp. 638-641.

ALSO BY ROSENTHAL AND WORTH READING
The interpretation of levels of significance by psychological researchers.
Journal of Psychology 55, pp. 33-38, by *R Rosenthal and J Gaito (1963)*

D3: Dependence of empirical laws upon the source of experimental variation.

Psychological Bulletin 66, pp. 488-498

Robert GRICE (1966)

A problem that plagues Experimental Psychology is the difficulty in repeating the results of other researchers - the lack of the *replicability* of work that characterizes a genuine science. This is one of the motors behind accusations of experimental incompetence that poison the atmosphere.

Grice's article discusses one reason why failures to replicate are so common. This contribution is also a good introduction to a fundamental methodological issue in Psychology: whether an experimental variable is designed for "repeated-measures" or "between-groups".

In a repeated-measures experiment, a participant is compared with herself. For example, she can be tested for intelligence at eleven years of age, and then again two years later. In the between-groups version of this experiment, students of one grade level would be compared with students of a later one. (In practice, *both* types of experiment would probably involve *groups* of participants, to do statistical analysis.)

It is much more *convenient* to compare the two grade levels, rather than to have to wait two years by following the eleven-plus set through. The repeated-measures design, however, is more elegant, efficient and economical. Only half as many participants are required. The results should reflect a pure effect of ageing, rather than being *confounded* with the difference between two year groups, that is, between two different sets of people. We say that subjects are serving as their own controls. The limit of this approach is the single case study or "*N*=1" found throughout the medical literature.

The choice of a type of design, however, is not so straightforward, for all that some established researchers dismiss between-groups studies as inherently flawed. The point they do not take into account is that the subject's experience of a prior experimental session may *transfer* to the later one. This transfer may take the form of a practice effect (improvement) or a fatigue effect (deterioration). Such effects are just as liable to confound the age comparison as participant differences are.

With other independent variables than age, say season of the year, comparing winter and summer, it is possible to *do* something about the transfer problem. Simply have half your participants do the test in winter first, while the other half do it in summer first. This is called *counterbalancing* the design. It can be taken to considerable lengths when variables have more than two levels, with "Latin-square designs" and so on. The intention is that a practice effect will even out across the winter and summer conditions that you are comparing.

This is certainly an increase in experimental sophistication, but it is not the end of the story with transfer. There remains a theoretical possibility that previous practice in summer may be helpful, but not previous practice in winter. This would mean previous practice would still confound the comparison of seasons. Even in today's Psychology, hardly any researchers are doing anything about this possibility of *asymmetrical transfer*.

Grice built his experimental reputation on behavioural research using animals. This work is the basis for his confident assertion that the two different types of design, repeated-measures or between-groups, may lead to different results. For example, he describes such different results with work on "stimulus intensity" and "stimulus generalization". Anyone reviewing such findings in terms of the experimental variable only, without taking into account the type of design, will conclude there has been a failure of replication. So, discord and demoralization go on in Psychology.

For animal research, experimental subjects are plentiful, and Grice recommends doing *double* experiments in which a repeated-measures procedure is directly compared with a between-groups one.

If you come to use computer statistical software for analysing data in terms of variances (the *F*-test) you will find it needs to have each variable identified as repeated or not.

Two measures on a single subject tend to be correlated. The *error variance* against which experimental variance is compared is thus reduced. This is just putting into different words what is the advantage of repeated-measures designs. However, the correlation undermines an assumption underlying the analysis of variance (the technical description of this assumption is “homogeneity of the variance-covariance matrix”). This assumption should not be accepted in repeated-measures designs without *testing* whether it is approximately true. Though the steady improvement of statistical software means that this test can sometimes be done, in many published experimental studies it has not been done, and their results are therefore statistically suspect.

In many areas of Psychology, participants are difficult or expensive to come by. This will always prompt recourse to repeated-measures designs. Researchers who have adopted such designs themselves do tend to favour them. But between-groups designs are no less preferable inherently. The only methodological kudos goes to those who adopt Grice’s prescription of combining both types of design into a single experiment.

FOR AN ARTICLE FOLLOWING UP ON THE GRICE PRESCRIPTION TRY:

Design and analysis of experiments contrasting the within- and between-subjects manipulation of the independent variable.

Psychological Bulletin 84, pp. 212-219, by *Al Erlebacher (1977)*.

EXERCISES FOR SECTION D "RESEARCH: THE METHODS USED TO FIND EVIDENCE FOR CONCLUSIONS"

1 Does the method of laboratory experiment present special problems in Psychology?

2 Will Psychology eventually accumulate enough facts to build into sound theories?

3 Has Psychology neglected to watch people?

E1: Asking only one question in the conservation experiment

Journal of Child Psychology and Psychiatry 25, pp. 315-318

Judith Samuel and Peter BRYANT (1984)

The conservation experiment is associated with the name of, above all, Jean Piaget. This Swiss psychologist has made more impact on the Anglo-American empiricist mainstream than any other outsider, although Psychology is vigorous in continental Europe. There are other significant Europeans, like Alexander Luria, and just as the English philosopher Alfred Ayer “discovered” the work of the Vienna Circle, so there may be a future fusion of the European and the empiricist traditions in Psychology.

Piaget began his intellectual life as a philosopher, with a specialist interest in “epistemology”, the philosophical study of knowledge. The insight that was to guide his life’s work was that adult knowledge can be understood better by studying the *development* of cognition in the child. Unfortunately, his writing style is a difficult one even for French speakers, made harder still by the translations, and a popular exposition such as Flavell’s is the best way into his ideas.

The conservation experiment can act as a test for “operational” thinking in a child. Piaget divides childhood into four stages. The first is called “sensorimotor” - the child up until two years of age is preoccupied with its ability to perceive and move. Then the child enters a “pre-operational” stage, in which it is building up to being able to perform “operations”. A stage of “concrete operations” lasts from seven to eleven years. Finally, the child enters the stage of “formal operations”.

The classic version of the conservation experiment uses glass containers of different shapes, either short and fat or tall and thin. The experiment starts with equal amounts of water in two short fat containers, the child is asked whether they are the same amount, and agrees. Then water is poured, in front of the child, from one of the short fat containers into a long thin one. Only a child who has reached the operational stages will agree that the amount of water in the long thin container is the same.

Such a child can recognize the *identity* of the water (one operation) and *compensate* for the change of shape (a second basic type of operation). The operational-stage child will also agree, that if the water were poured back into the short fat container it would be the same amount as the other short fat container, an operation called *reversibility*.

Piaget was stronger on theory than on experiment, and his work has been criticized as being unscientific and as not conforming to the mould of hypothesis testing. Professor Peter Bryant spent most of his career at Oxford University putting such work on a sounder scientific footing. In this article, with Judith Samuel, he identifies a feature of the *procedure* of the conservation experiment that is critical.

The point is that the child is asked *twice* whether the two amounts of water are the same, the second time after the experimenter has, under its nose, *changed* something. It is natural for a young child to imagine that the experimenter wants a changed answer. Children tend to do what pleases adults. “Unless the adult wants me to say something has changed - what could be the point of the experiment?”

So Samuel and Bryant tried the conservation experiment again, asking the question whether the water was the same amount only once - after the water had been poured. Many more children now answered the question correctly, and these results showed clearly that many children who *have* mastered the relevant operations merely *look* as though they haven’t in the classical conservation experiments.

Samuel and Bryant repeated their experiment with the conservation of number (rows of counters) and mass (pieces of plasticine). The results were reliable, in a large sample of 252 children. They used analysis of variance for the statistical treatment of their data. In all these respects they went beyond Piaget's much less formal studies.

What they have established, to conclude, is that Piaget's norms for the age of onset of cognitive operations were substantially too high.

E2: Some conditions of obedience and disobedience to authority

Human Relations 18, pp. 57-76

Stanley MILGRAM (1965)

Two of the articles in Section D touch on the issue of authority within the psychology experiment. But the definitive exposition on authority is this work of Milgram, which is probably the best-known experiment in the whole of Psychology. The time in our lives, when obedience to authority is most important, is childhood, and that is why this article appears in Section E.

What Milgram did has aroused considerable controversy. He recruited volunteer subjects to take part in an experiment on verbal learning. When they came to the laboratory they found a row of levers marked according to the number of electric volts up to 450, with another lever marked “XXX” at the very end of the row. They were told the purpose of the experiment was to investigate the effect of punishment (with electric shock) on learning.

Though their role was to *do* the punishing, they were given a mild shock themselves to convince them of the reality of the punishment. They heard the learner (a confederate of Milgram’s) reporting his memory of word lists. When they could see (from their copy of the list) that the learner was wrong, they gave him the lowest level of shock. The punishment increased with successive errors.

Occasionally the learner who was pretending to suffer this punishment protested. The protests could include the learner complaining he had a heart condition. The key question was how many participants would go all the way through the levers beyond 450 volts in obedience to the experimenter. Milgram’s astounding result was that no fewer than two-thirds of them did so. Even experienced psychiatrists, asked to predict the degree of obedience that would be shown, underestimated this figure.

There was no doubt the participants did become emotionally involved in the experiment. They were observed to sweat, tremble, stutter, bite their lips, groan and go into nervous laughing fits. Milgram’s intended deception of them worked. They frequently voiced concern over what they were doing and asked to stop. Milgram urged them to continue with a standard set of prods such as “The experiment requires that you continue”. Most did continue.

So far so bad. The experiment encourages a gloomy view of human nature. But Milgram explored further. He investigated two aspects of *immediacy* in obedience to authority. First, the victim is immediate, who could be close enough to touch, or further away, even in a different room. The experiments confirmed that there was more disobedience when the learner was close. Second, there is the immediacy of the experimenter. There was more disobedience when Milgram gave his prods over the telephone.

There were other ways to encourage disobedience. Participants took Milgram as a genuine experimenter from his lab coat and were more recalcitrant when he wasn’t wearing it. Another influential factor was having more than one participant experimenting at the same time. By itself, this gave the participant more confidence. But if another participant provided a *model* of disobedience, the primary participant became much more likely to resist.

Some other variations on the experiment do *not* affect the willing obedience shown. Milgram thought that a recruitment advertisement from Yale University might have convinced those answering that the experiment must be scientifically valuable, but found that participants obeyed no less for downtown “Research Associates of Bridgeport”. Children obey readily in this experimental paradigm, as do adult participants in many other countries. Milgram found no evidence of any difference between male and female participants.

There are ethical questions about this dramatic work of Milgram. He began it in the shadow of the Second World War and Hitler's *Endlösung* (Final Solution) of "the Jewish problem" - the terrible Holocaust of more than five million people. In those years it was common to vilify as evil monsters the ordinary Germans who connived in these murders. Milgram felt it helped international understanding to present an alternative picture of them, a picture of obedient underlings.

Even today, right-wing politicians say we should be readier to condemn and less ready to understand. Many viewers of the film Milgram has made of his experiments do attribute evil to the obedient participants. This temptation runs very deep. It's a hot potato. In any case, the British in 2014 may be less obedient than Milgram's participants.

How about the ethics of putting participants in the experiment through such a harrowing experience? Milgram was careful to come clean and explain to subjects the true purpose of the experiment at its end ("to debrief"), and they had an opportunity then to complain. Some may simply have felt foolish. But a further problem is that deception has become quite common in a range of similar experiments; many potential participants have come to be leery of a hoax, and so Milgram has queered the pitch for all Psychology experimenters.

The Milgram experiment started a fashion for outrageous experiments, and one follow-up in this paradigm involved *actual shocks* to the "learner". It seems like a re-visit of Hitler's Germany in more than one way (for hideous experiments on human participants are now known to have been conducted there).

Milgram did other important work, for example using the "lost letter" technique to measure altruism, and on the "small world problem". He was a sad loss to Psychology, dying young at 51.

E3: A conception of adult development

American Psychologist 41, pp. 3-13

Daniel LEVINSON (1986)

An idea that has only entered popular consciousness since the War is that of the "mid-life crisis". A seminal source was a book by journalist Gail Sheehy called *Passages,* and Sheehy acknowledged her debt to Levinson as the professional psychologist who inspired her thinking.

Levinson is one who has taken Developmental Psychology away from its former exclusive concentration on children. Child Psychology must, of course, have a special place in the developmental branch. Childhood is the foundation for the rest of our lives ("the child is father to the man"), and it is the time when we are changing fastest. Nonetheless, modern Psychology extends its attention to the whole lifespan.

The course of life has recurrently in literature been compared to the seasons of a year, beginning in spring and continuing through summer and autumn into the winter of old age. The mid-life crisis is an idea about autumn.

Nowadays we are more aware of the culture-boundedness of our ideas. In some parts of the world, there are just two seasons: dry and wet; while in the tropics the climate changes little through the year, and the *absence* in African literature of this metaphor of seasons has been noted.

Levinson says the life course is divided into several stable periods, such as early adulthood, and transitional periods between them. He locates the "mid-life transition" at about 40-45 years, between early adulthood and middle adulthood.

In Levinson's view, the central components of a person's life are marriage, family and occupation - though he acknowledges that wide variations occur in the relative weight of these and the importance of other components.

Levinson studied the life course using in-depth interviews, beginning with a sample of men (though he went on to study women). This particular article is mainly a theoretical framework for material gathered from the interviews and presented in a 1978 book *Seasons of a Man's Life.*

For these men, the mid-life transition had involved *re-appraisal* of their marriages, families and occupations so far in their lives. Earlier life often included the formation of a "dream" about what later years might contain. In mid-life the reality was reviewed by comparison with the dream. The dream, too, having been based initially upon an effort of imagination, was reviewed in the light of real experience.

Something else was happening in mid-life. Time for the accomplishment of the dream started to seem limited. - a man had to come to terms with being "halfway through". For some, transition became crisis. More frequently than is widely known, this may mean a mental crisis. At mid-life, it's not *that* a man's going to die, but *when* a man's going to die. This is a particular problem for those who, forgetting for a good while about mortality, are suddenly brought face to face with it, through, say, a bereavement.

Though for most the crisis is not psychiatric, life events that touch nearly everyone - parents ageing and dying, children growing up and leaving home, divorce, redundancy - can hardly have no effect.

Yet another source of crisis in mid-life has to do with our physical powers. For longer than we can remember they have been ceasing to expand. Whether they are actually in decline is scientifically controversial and complex, and it depends a lot on the individual. But the perception of possible decline begins, and perception may be more powerful than reality. In dark moments worries about our physical strength, our looks, our memory, and our sex drive all nag us.

For women there is a definite biological change, the end of the fertile period - the menopause. Though nothing changes in men's reproductive capacity, the phrase "male menopause" is widespread too. To worry and cope with it is normal, but a failure of healthy adaptation doctors call a "neurosis". There is nothing neurotic about the psychological consequences of the serious physical illnesses that some suffer at this age.

It is obligatory to express two serious reservations about the whole idea of a "mid-life crisis". First, Levinson's interviews were conducted with middle-class Americans and the extension of the idea to people in other cultures, particularly the developing world, may well be inappropriate. Second, no statistical evidence at all is presented, even to support the claim that mental crisis clusters around mid-life, although psychiatric records should speak to this.

EXERCISES FOR SECTION E: AGE

1 Do people change in the course of their lives?

2 How is childhood different from the rest of life?

3 Are our childhood years formative to some extent?

F1: Population density and social pathology

Scientific American 206, pp. 139-148

John CALHOUN (1962)

Seeking an explanation for mental disturbance, many writers have pointed to the fact that most of us nowadays live in cities—there has been a global trend of *urbanization*. One most apparent distinguishing feature of cities is the extent in them of *crowding*.

Psychologists have worried about the effects of crowding on us, from the time it became clear that it is conditions of overcrowding that cause the mass “suicide” of the Norwegian lemmings. Every year these rodents migrate to the sea, to the deep fjords cut into the coast, where many of them drown. In a research study, a herd of deer living isolated on an island in the Chesapeake Bay was monitored for numbers. The number of these deer rose generally but tumbled dramatically every so often as they bred above the limit that the island could support. Do such examples have relevance to human beings confined on our planet?

It was John Calhoun who brought the study of animal crowding into the laboratory. His basic experiment was to confine breeding colonies of rats in limited spaces. He showed that individual animals in such colonies developed major pathologies of behaviour. He designed a colony’s living space so that part of it was more readily accessible. It was in this area that the animals chose to eat the food he provided. He observed that they preferred to eat in each other’s company, and this was another reason why crowding was more acute in this area. The animals living in this area of especially acute crowding showed the pathologies more sharply, and he described it graphically and memorably as a "behaviour sink".

Behavioural pathologies varied from animal to animal, and fall into three main categories:

1 deviations in sexuality
2 complete passivity - this was the main pathology for female animals
3 hyperactivity

An even more noteworthy finding was that infant mortality was nearly twice as high in the behaviour sink. Since about half of the infants died even in the better area of the colony, this meant that hardly any survived past weaning in the behaviour sink.

Before the colony expanded into a crowded one, male animals gathered a harem and mated within the harem. Though they defended a territory, they didn’t fight much; though they roamed freely, they didn’t mate outside the harem. Females, who didn’t fight at all, built nests and raised their young, resisting advances from anyone but their male. This was the normal social order, that disintegrated under conditions of high density.

Calhoun’s conclusions were based on close observation for more than a year. The animals were usually Norway rats, though he has also worked with mice. He repeated the basic experiment several times with variations in the procedure. This was at a time when public anxieties about animal experiments were much more muted and he did not face strong ethical objections.

Other studies have shown that the pathological behaviours described by Calhoun can also be demonstrated in non-crowded conditions.

Nonetheless, the findings are foreboding when we consider that the human population on Earth is rising fast. In the West, improved sanitation and successes in medical research are two factors that have brought death rates down. This is also happening in the developing world. But while, in the West, parents have limited the size of their families in response to the declining death rate, elsewhere this drop in birth rate has yet to show through sufficiently to curb population growth. In other words, the “demographic transition” has yet to occur.

The problem of crowding in our cities is allied to the question of an individual's ability to establish his or her *territory*. Each of us needs a space of our own within which we have freedom of action and some privacy. Clearly, under conditions of overcrowding, we have to work much harder to establish such a territory. The topic has been the subject of much empirical research beyond this.

F2: The social re-adjustment rating scale

Journal of Psychosomatic Research 11, pp. 213-218

Thomas HOLMES and Richard RAHE (1967)

Perhaps you have encountered at some time or another a checklist of stressful life events that are given points. So, for example, an injury or illness is given 53 points, the recent death of a spouse is given 100 points, sex difficulties make 39, a move to a new house 20, the changes at Christmas 12, and so on. The participant completes the checklist, adds up his score, and finds that it lies in a band of scores for which the advice is: “You are under considerable stress at the moment. Do not undertake any major new projects.”

The idea, which goes back to those doctors of ancient times, Hippocrates and Galen, is that stress causes illness. Today, certain specific conditions have been identified as having a *psychosomatic* component (that means, they are bodily diseases caused by stress). These conditions include asthma, some skin disorders, hypertension (high blood pressure), ulcers, irritable bowel syndrome, migraine and low back pain.

There is also now reliable research evidence that there is a certain type of personality called “Type A”:- hard-driving, ambitious and hostile to others - and especially prone to *coronary heart disease* (the most frequent cause of death in Britain). Doctors are sufficiently convinced about the Type A idea to have tried prescribing breathing exercises and relaxation and meditation techniques to help anybody with this personality profile manage his or her stress better.

In the light of figures showing huge losses of production in the economy due to such stress-related illness, there has even grown up a new occupation of “stress manager”.

Stress is also behind a lot of mental disturbance. It is for depression that the best evidence exists for a causal role of stress. George Brown and Tirril Harris did well-known and careful studies (reported in a 1978 book) that established the case. They excluded the possibility that stresses merely act as a "trigger" for a certain sort of personality. Eight or nine out of ten diagnosed depressions had been preceded by stressful life events in the previous nine months. This compared with only three out of ten control subjects without depression experiencing such stress. Many other investigators have now reported similar findings.

So we need ways to identify and measure stress. The idea of stress point counts is regularly aired in high-circulation magazines, and though popularization often leads to distortions and inaccuracies, the idea itself is based upon legitimate research, the landmark article being this one of Holmes and Rahe.

Research has moved on a great deal since 1967 and no doubt the authors themselves would be the first to admit the crudity and oversimplification of this first article. The full list of life events chosen and estimated for stressfulness is given at the end. Scrutiny of them shows obvious datedness. One - “Wife begins or stops work” - assumes a man rather than a woman is completing the checklist. Also, the dividing line between a small and a large mortgage/loan would need revision now.

By emphasizing *events* so much, rather than ongoing conditions such as being unemployed or living in poor housing, major potential sources of stress are ignored. Also, it would be a better test instrument for separating questions of lifestyle from the more dramatic one-off occurrences.

Further, how current or recent this particular re-adjustment has been in her experience is going to affect an individual’s estimate. It is a *retrospective* perception of stress, and memory is unreliable.

Holmes and Rahe drew up their list “empirically from clinical experience”. They then asked a sample of 394 subjects to estimate the amount of re-adjustment required by each life event (with the reference point being Marriage = 50). This task of magnitude estimation has been extensively used in rigorous research on “psycho-physical scaling”.

Surely these magnitude estimates would depend a lot on the individual? Holmes and Rahe appear to argue that individual variation is small. They divided their sample into groups, e.g. 179 males and 215 females, and found a high correlation (greater than 0.9) between groups. To me, this is counter-intuitive and casts some doubt on their empirical work.

Also, this survey was investigating *lay* theories of stress. It has been said that on no other issue are lay and professional perceptions more sharply at variance.

Please turn to the next page for the full list of life events requiring readjustment.

Holmes-Rahe social re-adjustment rating scale

Death of spouse	100
Divorce	73
Marital separation	65
Jail term	63
Death of close family member	63
Personal injury or illness	53
Marriage	50
Fired at work	47
Marital reconciliation	45
Retirement	45
Change in health of a family member	44
Pregnancy	40
Sex difficulties	39
Gain of new family member	39
Business readjustment	39
Change in financial state	38
Death of close friend	37
Change to a different line of work	36
Change in number of arguments with spouse	35
Mortgage over $10000	31
Foreclosure of mortgage or loan	30
Change in responsibilities at work	29
Son or daughter leaving home	29
Trouble with in-laws	28
Outstanding personal achievement	28
Wife begin or stop work	26
Begin or end school	26
Change in living conditions	25
Revision of personal habits	24
Trouble with boss	23
Change in work hours or conditions	20
Change in residence	20
Change in schools	20
Change in recreation	19
Change in church activities	19
Change in social activities	18
Mortgage or loan less than $10000	17
Change in sleeping habits	16
Change in number of family get-togethers	15
Change in eating habits	15
Vacation	13
Christmas	12
Minor violations of the law	11

F3: On being sane in insane places.

Science 179, pp. 250-258

David L. ROSENHAN (1973).

On the face of it, we can recognize readily enough whether someone is mentally disturbed. Television and the newspapers regularly give us the information from which we form our picture of mental disturbance. We are told with the authority of the medical profession that there are diagnosable mental illnesses such as psychopathy, bipolar disorder and schizophrenia.

A belief holding great sway is that there are two kinds of people:- a majority who are normal, and a minority who are mad. Less easy to fit into this belief are people whose abnormality only shows itself in particular circumstances:- when a spider appears, for example. Such people used commonly to be diagnosed by psychiatrists as having a *neurosis.*

This celebrated study by David Rosenhan placed a big question mark against psychiatric diagnosis. His basic idea was to tell *normal volunteers* to present themselves to psychiatric hospitals with a single symptom such as "I hear a voice saying 'Thud'". Thereafter they were to act totally as normal. What he found was astonishing. The hospitals detained *every volunteer* for several weeks, and then discharged them with the diagnosis "schizophrenia in remission".

There has been an "anti-psychiatric" critique, by disillusioned doctors such as Ronald Laing, David Cooper, Aaron Esterson and Thomas Szasz, of what they call the "medical model" of mental illness. They criticize particularly the physical treatments, such as anti-psychotic drugs, electroshock and psycho-surgery, that are used under the medical model. The view of mental disturbance as a disorder of brain biochemistry is part of the medical model. Anti-psychiatrists compare the medical model to a television viewer sending for the TV repairer because he dislikes the programme he sees on the screen. The anti-psychiatrists seized upon Rosenhan’s study with delight.

For anti-psychiatry, the problems of mental patients were due to their distress in an inadequate society. Now psychiatrists and mental health service professionals generally could be identified as one particularly distressing feature of that society.

The term “schizophrenia” is not comparable to “heart disease”, say. Anti-psychiatrists describe it as a form of character assassination, used to invalidate and attach a lasting stigma to people and to what they say. It is a modern replacement for the term “witchcraft”. In the Soviet Union, for many years, the old tyrannical society abused psychiatry. Many people were detained and given physical treatments forcibly, for no more than criticizing the political regime. The more public their criticism the more serious their condition.

Rosenhan’s study was not based on the experiences of a single individual. Eight different people gained admission to twelve different hospitals. After phoning for an appointment, they told the admitting psychiatrist about their voice - what it was saying varied. Like any psychiatric patient, they entered knowing nothing about when they would be discharged. They found absolutely nothing to do in the psychiatric wards, which increased the considerable stress of the whole experience. All hoped for discharge almost immediately, but they were detained from 7 to 52 days (on average, for 19 days). They tried to engage others in conversation and wrote down their observations on the ward and its patients and staff. It was the other patients, not the staff, who detected that they were different, and accused them of being there to check up on the hospital.

The hospitals segregated staff and patients strictly, in terms of dining facilities, bathrooms, and places to assemble. Staff came out of their quarters for specific care-taking purposes, to give medication, conduct a therapy or group meeting, and instruct or reprimand a patient. Otherwise, the staff kept to themselves.

In a second phase of the study, another research and teaching hospital whose staff had heard these findings was informed further of these “pseudo-patients” would attempt to gain admission. Although no further pseudo-patients were sent, 41 out of 193 genuine patients were judged to be pseudo-patients.

Psychiatrists are indeed trained in medicine, and what they respond to this sort of attack is that though Rosenhan may have shown that diagnosis for mental illness is unreliable, diagnosis for physical illness is sometimes pretty shaky as well. The community requires them to deal with the problem of mental illness, they have a legal responsibility, and all Rosenhan’s study showed was that some individual psychiatrists were playing very safe. Perhaps they were putting the rights of the community higher than the rights of the individual - this is an ethical issue that much discussion has failed to resolve.

What is important is not to stereotype people simply for being a mental patient, or a psychiatrist, or a mental nurse, or, indeed, a psychologist. All individuals act within a social system, and any individual can do her or his best within the given system. Take people as they come.

EXERCISES FOR SECTION F: MENTAL HEALTH

1 Are problems in mental health genetic in origin?

2 Are doctors of medicine the best people to deal with problems in mental health?

3 Do you agree with Dr Thomas Szasz that it is bodies that are ill, not minds?

G1: A study of prisoners and guards in a simulated prison.

Naval Research Reviews 30 (9), pp. 4-17

Craig Haney, Curtis Banks and Philip ZIMBARDO (1973)

The task of social psychologists is to unsettle us, to shake us out of the cosy assumption that “everything is for the best in this best of all possible worlds”. This experiment by Philip Zimbardo, often called the Stanford prison experiment, undermines any too-benign view of human nature. Though famous, it has often been questioned and criticized on ethical grounds, to an extent because some people dislike the findings so much.

Prison is a side of our society that many prefer not to look at. A fundamental philosophical divide separates those who believe prison is purely to punish and those who believe it can and should also *rehabilitate* the offender. Both schools of thought have to face the fact that of those released from prison, a very large proportion offend again (this is called “recidivism”). Moreover, incarceration in prison breeds a serious alienation from authority and the established order of society. We also have to come to terms with the fact that prison is an exceptionally costly social expenditure. In Britain, it is the Howard League that presses for Penal Reform.

The riots that are endemic in the prison system are only the tip of an iceberg of violence and brutality. Often such behaviour is put down to something about prisoners and warders *as people*. This is the assumption that Zimbardo sought to challenge. It is not just that prisoners are the anti-social element anyway, not just that warders tend to be uneducated, insensitive and even sadistic.

Another motivation for the research, as the title of the journal suggests, was to understand better the way captured military personnel can be brainwashed by the enemy holding them. This phenomenon was seen for example in the Allied pilots captured during the Kuwait conflict, who were exhibited on television by Saddam Hussein's Iraq.

In Zimbardo’s experiment, an attempt was made to simulate prison life as closely as possible. Some aspects could not possibly be reproduced, such as the practice of physical beatings and the ambience of homosexual and racist behaviour. Above all, an experiment cannot possibly continue for the length of time of most prison sentences. Nonetheless, the basement of the Stanford Psychology Department was converted into something physically approximating a prison, with secure cells. Warders wore uniforms, and prisoners, who were known only by a number, wore prison clothing, with an ankle chain. Zimbardo presents plentiful evidence that the simulation of prison was taken as real by the participants; particularly as the experiment went on and they entered into their roles more fully, forgetting that it was just an experiment.

The key point is that both prisoners and warders were normal volunteers. They were a minority from a pool of male college students volunteering, selected for their stability, maturity and lack of anti-social tendencies. Also, these volunteers were assigned *randomly* to the role of prisoner or warder. There are some parallels with the study by Rosenhan of psychiatric hospitals, reviewed just previously. Both studies are explorations of what have been called the “total institutions” of society.

The effects of the prison simulation on those participating were dramatic. What Zimbardo found was that all their encounters with each other became negative, hostile and dehumanising. The prisoners became very passive, the warders very active, happily issuing forth the commands that, along with insults, were the main form of verbal interaction during the experiment.

Several of the prisoners simply could not take their ordeal, and the experiment had to be terminated ahead of schedule. This was because of the prisoners’ adverse reactions, including in one out of the twelve of them an extensive psychosomatic skin rash. The experiment had lasted six days, and the warders were already living their roles enough to regret giving up the power and control that they had enjoyed for that time.

The prisoners' reactions could be described as a loss of their sense of personal identity, a surrender to the arbitrary control of the warders, and dependency and emasculation.

From an ethical point of view, it should be noted that Zimbardo had gained approval for this experiment from all the appropriate committees. He also instigated a thorough debriefing and follow-up programme and was satisfied that the psychological after-effects of the experiment were not long-lasting.

Finally, a criticism of the study is that it was wholly qualitative observation, with no statistical evidence presented.

G2: Bystander intervention in emergencies: diffusion of responsibility

Journal of Personality and Social Psychology 8, pp. 377-383

John M DARLEY and Bibb LATANÉ (1968)

The terrible crime of murder is all too common, but in the annals of crime, one particular murder stands out - because of the behaviour of bystanders. Very late one night Kitty Genovese was returning home from her job as manager of a bar in New York City. As she walked the short distance from her car to her apartment house, a man with a knife came up to her. She ran, but he gave chase, caught and stabbed her. She screamed for help, and many lights came on in the street. The attacker retreated. But when nothing more happened, he changed his mind and came back. Although she carried on screaming, he stabbed her until she was dead. The whole attack took more than half an hour, yet of thirty-eight witnesses who told the police that they had heard the screams, none at all intervened.

This was the case that inspired the experimental investigations of Darley and Latané. Rather than leave it that such occurrences are due to “moral decay” or “urban anomie”, they went deeper. They felt it was the sheer number of witnesses that had caused the lack of help. Everybody was leaving it to everybody else to do something.

For though we know it is right to help someone in this sort of distress, there are plenty of real or perceived reasons for avoiding involvement. There is the chance of physical harm, above all, but also that of losing days at work through being drawn into police procedures. There is even the chance of public embarrassment if the situation has been misperceived. In this conflict between moral sense and prudence, the knowledge that others are just as able to help as oneself tips the scales towards prudence. In the Kitty Genovese case, some of the thirty-eight witnesses may have assumed that somebody else *was* at least phoning the police (though none of them did, in fact).

So Darley and Latané created a bogus emergency in the laboratory, with a varying number of bystanders. The hypothesis was that as the number of potential helpers increased, so responsibility for helping would be *diffused*, and as a result, the likelihood of helping would decrease. Not only the responsibility but the possible blame for not doing anything would be diffused in a larger group.

When the undergraduate participants arrived for the experiment, they were told that they would discuss with fellow students some of the problems faced by those attending college in a high-pressure urban environment. To reduce embarrassment about revealing personal problems, each participant would be alone in a room, communicating with the others through an intercom system. The experimenter would not be listening. Each participant would talk for two minutes and then each would comment on what the others had said.

Participants were told that they were one of either two, three, or else six students taking part. In fact, they were alone, and the other supposed participants were just tape recordings. The first person to speak in the supposed discussion was the “emergency victim”. He said, sounding embarrassed, that he was prone to seizures, especially during stressful times such as exams. Then the two-minute talks were given, with the “victim” responding last. After making a few calm and coherent statements, this recorded actor then gave a convincing simulation of someone experiencing a seizure.

What Darley and Latané measured was the percentage of subjects who left their laboratory room while the “seizure” was going on to look for the emergency victim. This percentage dropped from eighty-five when they thought they were alone with the victim, to sixty-two when they thought there was one other additional subject, to thirty-one when they thought there were four other additional subjects.

The results, therefore, confirmed the hypothesis - the more bystanders, the less likely an individual is to help. Observation of the participants suggested their apparent apathy overlay a lot of emotional upset and confusion. Nonetheless, this phenomenon has become known as “bystander apathy”. Observation of the participants suggested that they all were taken in by the simulated seizure, whether they intervened or not. The success of the hoax was also confirmed by the de-briefing of the participants that had to take place on grounds of ethics.

It made no difference to the results whether the participants were male or female, or whether they thought the other bystanders were male or female.

This famous experiment was the beginning of a whole programme of research by Darley and Latané on what is variously called helping, altruism or “pro-social” behaviour.

G3: Twenty years of experimental gaming: critique, synthesis and suggestions for the future.

Annual Review of Psychology 28, pp. 363-392

Dean PRUITT and Melvin KIMMEL (1977)

Typical of modern Social Psychology as neat *demonstrations* of what we may have suspected about human nature are the two previous articles of Zimbardo and Darley and Latané. However, these are not typical of Psychology more broadly, where the *application of scientific method* has been the dominant ideology. To find an attempt to align Social Psychology with broader Psychology, we turn to this field of experimental gaming.

No one experimental study has had the sort of impact to make it a classic article - rather, the gaming idea just seemed to take hold gradually, and so it is an important early review article that is under consideration. Another early landmark was a book by Anatol Rapoport. The field of experimental gaming had already accumulated over a thousand published studies by the time this article by Pruitt and Kimmel came out, but there is little citation to these from the rest of Social Psychology. There has been enthusiasm, but not much cross-fertilization.

The phrase “experimental gaming” describes a method of research rather than a substantive topic, but the standard Social Psychology course is organized in terms of substantive topics. In such a course, this work will fall under the heading of “social exchange” “strategic interaction” or more familiarly, *cooperation*.

The basic idea can be given with the simplest sort of game that has been studied, one known as the cone game. Here, each of several participants, probably children, holds a string with a cone on the end. The cones are in a bottle - one cone can come out through the neck of the bottle at any one time, but not several at the same time. The task is to pull the cone out of the bottle, and children learn to take turns in pulling their cone, right? Wrong, what invariably happens is that all the children pull their strings at the same time and the cones all become jammed in the neck of the bottle.

In other words, the level of cooperation shown is very poor, and this is the main finding and theme of all the experimental gaming literature.

Other very simple games have been used by experimental social psychologists, such as the trucking game devised by Morton Deutsch and Robert Krauss, and the “minimal game” (or “minimal social situation”). But the most famous game used to study cooperation, what has been called “the archetype of the controlled laboratory experiment in Social Psychology”, is the *Prisoner’s Dilemma* game.

This is a game for two players, who have to put themselves in the position of two prisoners, each of whom has a choice of two courses of action. One course of action is to confess to the crime, while the other course of action is to “shop” the other prisoner for the crime. The *outcome* for each prisoner depends on what the other one does. If they both shop their partner, they both gain an outcome of 2 only. If they both confess to the crime, they both gain an outcome of 8. That is, the cooperative course of action works out better *if* the other player is also cooperative. So there is some incentive for co-operation. Neither player knows his partner’s decision before his own decision. The best outcome of all, however, comes from shopping your partner while he is trying to be cooperative. Then your outcome is 10 while his is only 1. What this means conversely, is that the worst outcome follows being co-operative with an uncooperative partner, so there is also some disincentive for co-operation.

One pair of decisions is called a trial of the experiment, and the experiment can be repeated for many trials without the cooperative pattern of both confessing emerging. In this respect, the experimental paradigm is consistent with the rest of the experimental gaming literature.

You can probably imagine the scope for variation of this basic paradigm over hundreds of published studies. The numerical outcomes (the "payoff matrix") can be varied, like the number of players, and whether players are given some sort of "veto" power. Do players show individual differences? It has been said that they are one of three types: the natural co-operators, the *competitors* who always want to beat the other player, and the *individualists* who have to maximize their outcomes though they do not need to do down the other player.

Defenders of this sort of research say that the basic result is highly reproducible. They also point to the many global issues that depend on promoting cooperation, such as conserving energy, limiting environmental pollution and containing population growth. Other relevant issues include whether to join a trade union and behaviour in panic situations. Collective action by committees includes such significant organizations as the United Nations (including the Security Council), the European Union, and trade cartels such as the Organization of Petroleum Exporting Countries. If research on the Prisoner's Dilemma game throws light on this sort of real-life cooperation, then it is justified indeed.

EXERCISES FOR SECTION G: UNDERSTANDING HUMAN INTERACTION—SOCIAL PSYCHOLOGY

1 Is there a place for laboratory demonstrations of social interaction in special circumstances?

2 Should Social Psychology advance using the application of scientific method?

3 Is social psychology what most people mean by “psychology”?

Author Note

The exciting time in my life was eleven years of Northern Ireland during the disturbances there. While a lecturer at the University of Ulster I wrote many academic titles, which now, many years later, place me ahead of four-fifths of other members of Researchgate.net (where most of these titles are available in full-text). With children reaching school age and a credible death warning from the IRA, I had to leave, taking up an offer to teach nurses and nursing students in England. When that job went after four years I set up as a psychologist in independent practice. Though the work was varied and at first fulfilling, my bread-and-butter was giving cognitive behaviour therapy to victims of road traffic accidents. I retired from eighteen years of this, though I continued as an examiner for International Baccalaureate. I write about Psychology and have always believed that it neither is, will, nor can be a science. When it comes to behaviour or neurocognition that aspiration may be less futile and self-destructive. The same may go for aspects of Environmental Psychology, a more recent speciality for me. Since retirement, I have taken a lot more interest in computer mathematics. I also practise my Christian faith more wholeheartedly, though many will see a tension between that and my divorce after 32 years of marriage.