2005

Psychological Sketches (7th Edition)

Vokey, John R.

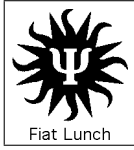
Psyence Ink

Psychological Sketches (Seventh Edition)

Edited by
John R. Vokey and Scott W. Allen
Department of Psychology and Neuroscience
The University of Lethbridge

A PsΨence™ Ink Book

Produced in LaTeX 2ε
on Apple Macintosh™ computers



PsΨence™ Ink
Department of Psychology and Neuroscience
The University of Lethbridge
Lethbridge, Alberta

To our students, past and present;
and to future students,
Christopher and Blake.

In memory of
M. G. Vokey (1924–1998) and Wm. B. Allen (1930–1999).

Preface to the Seventh Edition

> Lord Ronald said nothing; he flung himself from the room, flung himself upon his horse and rode madly off in all directions.
>
> – Stephen Leacock

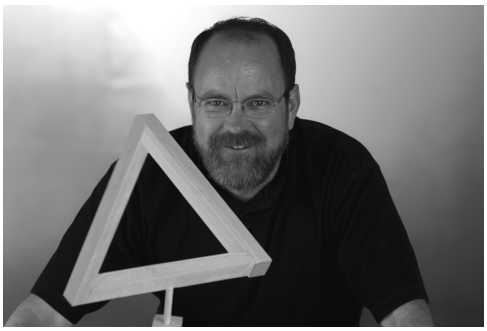
In 1912, the Canadian humorist and economist Stephen Leacock introduced the world to the town of Mariposa—a fictional small town (with *remarkable* similarities to Orillia) in Southern Ontario—in his book *Sunshine Sketches of a Little Town*. Through a series of short stories, or word sketches, about particular incidents and occasions in the town, Leacock led the reader to an understanding of the town and its people. In this textbook, and in the course that accompanies it, we have used the same strategy; through a series of short sketches of specific incidents or areas within psychology, we lead you to an understanding of the subject matter and methods of experimental psychology. Like Lord Ronald, however, these are sketches that ride off in all directions. Each of the chapters in this book is a short sketch of a particular topic in psychology. They can be read in any order as each chapter is meant to be a self-contained story. By the end of the book, we hope you will have learned what experimental psychology is about, what experimental psychologists do, and more specifically, what experimental psychologists do in the Department of Psychology and Neuroscience at the University of Lethbridge.

With both major additions and losses of faculty within the department, this edition of *Psychological Sketches* is a substantial revision of previous editions. This seventh edition of the book reflects these changes with the necessary, but ultimately sad removal of chapters previously contributed by now retired faculty, updates to some chapters from continuing faculty, and the addition of bright, exciting chapters from new members of the department. It also includes expanded subject and author indices to make it easier for the reader to locate specific topics and authors. We hope you enjoy it.

John R. Vokey and Scott W. Allen
July 6, 2005

The Contributors

Scott W. Allen

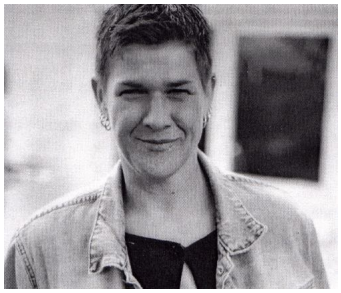


Scott W. Allen was born in Victoria, British Columbia, and lived in Ontario, Manitoba, Alberta, and B.C. before returning there to attend the University of Victoria and receive an Honours B.A. in Psychology. He received his Ph.D. in Psychology from McMaster University. He teaches cognitive psychology, and research methods and statistics, and has teaching and research interests in categorization and concept learning, memory, the acquisition of diagnostic expertise, and research ethics.

Jean Choi

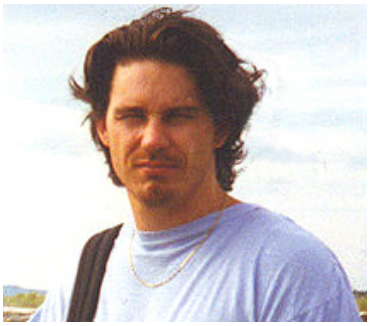
Jean Choi was born in Pusan, South Korea, and grew up in Toronto, Ontario. She received her Honours B.A. in Psychology from the University of Western Ontario, and her M.A. and Ph.D. from York University. She teaches child development, social psychology, and sex and cognition. Her research interests include the development of spatial behaviour and its evolutionary basis.

Margaret Forgie



Margaret Forgie was born in Vancouver, B.C. She received her B.A. from Simon Fraser University and her M.A. and Ph.D. from Concordia University in Montreal. She teaches human development, brain plasticity and memory, and hormones, drugs, and behavior. Her research interests include individual differences in brain organization and behaviour, in particular, the influence of hormones on the developing and aging brain.

Darren K. Hannesson

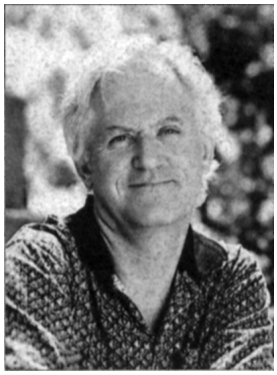


Born in Regina, Saskatchewan, Darren Hannesson obtained a B.A. from the University of Regina before heading to the Wet Coast. There, at the University of Victoria, he obtained an M.A. in Behavioural Neuroscience before fate drew him back to the University of Saskatchewan in Saskatoon where he obtained his Ph.D. in Experimental Psychology. Following a postdoctoral position at the University of British Columbia, he joined the faculty at the University of Regina before coming to the University of Lethbridge in 2004. His research interests are in the circuits, pharmacology, and molecular mechanisms of normal and abnormal memory function.

Gail Hepburn

Gail Hepburn was born in Toronto, Ontario, and grew up in the Ottawa area. She received her Honours B.Sc. in Psychology from Trent University, and her M.A. and Ph.D. in Psychology from Queen's University. After a two year post-doctoral fellowship at the University of Guelph she spent four years in Toronto at the Institute for Work & Health, a not-for-profit research organization. She joined the University of Lethbridge in 2004. Her area of specialization is Organizational Psychology. Her research interests include the effect of workplace factors—such as perceptions of justice or fairness, workplace aggression, and work-family balance—on employee well-being.

Bryan Kolb

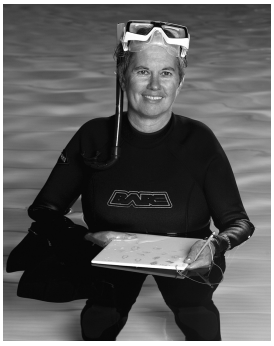


A native of Calgary, Alberta, Bryan Kolb received a B.Sc. and M.Sc. from the University of Calgary, and a Ph.D. from Pennsylvania State. He teaches neuroscience, neuropsychology, and brain plasticity. His research interests include brain development, plasticity, and brain behaviour. Among the books he has written is the widely-acclaimed *Fundamentals of Human Neuropsychology*. Professor Kolb is a fellow of the Royal Society of Canada.

Martin Lalumière

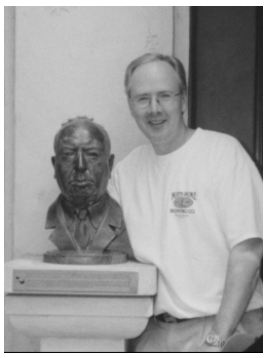
Martin Lalumière was born in Montréal, Québec and received his university degrees from the Université de Montréal (B.Sc., M.Ps.) and Queen's University (Ph.D.). He started his post-graduate research career at the Oak Ridge Division of the Mental Health Centre Penetanguishene (Ontario), a maximum-security institution for violent offenders with serious mental illnesses, and then joined the Centre for Addiction and Mental Health (Toronto, Ontario). Before moving to Lethbridge in 2004, he was an Associate Professor of Psychiatry and Criminology at the University of Toronto. When not fly-fishing, he conducts research on the etiology of sexual aggression, the development of atypical sexual preferences, the physiological assessment of sexual arousal, the effect of birth order on development, and the determinants of antisocial behaviour and risk taking.

Jennifer Mather



Jennifer Mather is a native of Victoria, British Columbia. She received a B.Sc. from the University of British Columbia, an M.Sc. from Florida State, and a Ph.D. from Brandeis. She teaches developmental psychology, perception, animal cognition, aging, schizophrenia, and the university capstone course. Her research interests include the behaviour of octopuses and psychopathology.

Robert J. McDonald



Robert J. McDonald was born in Calgary, Alberta, and received a B.Sc. in Psychology from the University of Lethbridge, and a M.Sc. and Ph.D. from McGill University in Montral. He currently teaches learning. His research interests include the organization of learning and memory in the mammalian brain, the etiology of Alzheimer disease, stroke and memory impairments, mechanisms of memory consolidation, and the link between learning and memory dysfunction and psychiatric disorders.

Gerlinde A. Metz

Gerlinde A. Metz was born in Kassel, Germany, and received a B.Sc. in Biology from the University of Giessen, and a M.Sc. in Biology and Neuroscience from the University of Giessen and the Federal Institute of Technology in Zurich, Switzerland. She continued at the Federal Institute of Technology in Zurich to receive a Ph.D. in Neuroscience and came to the University of Lethbridge in 1999 to perform post-doctoral research. She teaches courses on brain and behaviour, clinical neuroscience, and movement disorders. Her present research program focuses on investigating therapeutic strategies and physiological factors that influence recovery after brain damage.

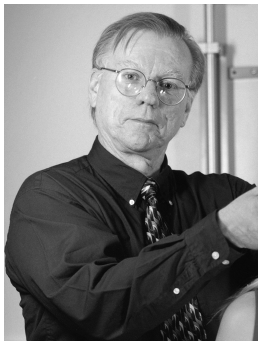
Sergio Pellis

Sergio Pellis was born in Trieste, Italy. He received his B.Sc., Dip. Ed., and Ph.D. from Monash University in Australia. He teaches statistics, comparative psychology, sex and human ethology, behavioural neuroscience, and behavioural development. His research interests include play and its control, and the organization of movement.

Glen Prusky

Glen Prusky is a native of Calgary, Alberta, and grew up in Claresholm, Alberta. An alumnus of The University of Lethbridge, he received a B.Sc. from there in 1986, and received an M.Sc. from Dalhousie University the same year, followed by a Ph.D. from Dalhousie University in 1989. Following a research postdoctoral position at Yale University for the years 1990–92, he joined the faculty at The University of Lethbridge in 1992. An avid musician, he is currently the drummer for the acoustic-rock duo, *Picks & Sticks*.

Robert Sutherland



Born in Toronto, Rob Sutherland received his B.Sc. from the University of Toronto and his M.A. and Ph.D. in Behavioural Neuroscience from Dalhousie University. He held a postdoctoral position at the University of Lethbridge before joining the faculty in 1981. From 1991 until 2001, he was a professor of psychology and neuroscience at the University of New Mexico (in Albuquerque) and rejoined the faculty at Lethbridge to become the University's first Alberta Heritage Medical Scientist in the new Canadian Centre for Behavioural Neuroscience. His interests include neuropsychology, cognitive neuroscience, learning and memory, functional brain imaging, neuropharmacology, and electrophysiology.

Drew Rendall

Drew Rendall was born in London, Ontario and after a spell there, spent the formative years in Halifax and Calgary. He did his B.Sc. at the University of Calgary, an M.A. and Ph.D. at the University of California-Davis, and postdoctoral work at the University of Pennsylvania. His research interests are in the evolution of social behaviour, vocal communication and cognition in human and nonhuman primates.

Paul Vasey

Paul L. Vasey was born in Toronto and went on to live in Edmonton, Yellowknife, Vancouver and Montréal with stops in between in rural Alberta, Ontario and Québec. He received a B.A. (Honours) from the University of Alberta, a M.A. from Simon Fraser University, and a Ph.D. from the Université de Montréal. He conducted post-doctoral research at Concordia University and York University. He teaches courses on animal behaviour, sex, sexuality, and gender. He conducts field research at Arashiyama, Japan, where he studies the development and evolution of female homosexual behaviour in Japanese macaques. He has a second field site in Independent Samoa where he studies gender atypical males, known locally as *fa'afafine*

John R. Vokey

Born in Winnipeg, Manitoba, John R. Vokey received an Honours B.A. in Psychology from the University of Winnipeg, and a Ph.D. from McMaster University. He teaches cognitive psychology, statistics and reseach methods, and has teaching and research interests in artificial neural networks, computer modeling of judgments of style in art and music, computer algorithms and numerical analysis, mathematical modelling of memory, models of consciousness, the recognition of faces, and the implicit learning and knowledge of structure.

Contents

List of Tables

List of Figures

Chapter 1

Facilitated Communication: Does It Enable Communication or Appropriate It?
Scott W. Allen

1.1 What is Facilitated Communication?

> Neil is 25 years old. He does not speak. Occasionally he screams for no apparent reason, makes moaning sounds, rocks back and forth, hits objects such as a desk, and bites the fat of his hand above the thumb. He spent the majority of his public school years in a special school for students with severe retardation. He is labelled autistic.
>
> With an adult supporting his hand at the wrist, he types with his index finger, revealing unexpected literacy. The method is called facilitated communication. One of Neil's facilitators recently asked him to complete the sentence, "When people see me they think I probably . . . " Neil typed "AM MEAN." When asked why people might think he is mean, Neil typed, "HOW I LOOK." Then Neil was asked to complete the sentence, "The thing I'd like others to know about my handicap is . . . " Neil typed, "I CANNOT HIDE IT."
>
> Neil's communication is not voluminous. He types mostly single-word and two- or three-word responses to questions. Yet his few words demonstrate attention to abstract concepts, awareness of others, and understandable feelings. (Biklen, 1992, p. 15)

Biklen was describing a man suffering from autism, a disorder that in its extreme forms is marked by profound communication difficulties. Those suffering from severe autism often do not speak (or, presumably communicate) at all, or if they do, they show echolalic speech, parroting back what they have just heard or constantly repeating inappropriate phrases such as commercial jingles or bits of songs. Severe autism is often diagnosed in conjunction with severe mental retardation. Thus, cases of people like Neil who appear to be capable of modes of thought far in excess of what they have previously shown are indeed intriguing and encouraging.

The magic key that appears to have unlocked the door to communication for Neil and others like him is a relatively new technique called facilitated communication. Developed by Rosemary Crossley in Australia in the 1980's (see Crossley, 1992; Crossley & Remington-Gurney, 1992) and popularized in North America by Douglas Biklen (see Biklen, 1990, 1992; Biklen et al., 1991; Biklen, Morton, Gold, Berrigan, & Swaminathan, 1992), this technique involves the use of a facilitator: an individual who supports the hand or arm of the autistic person in order to help that person point to letters or pictures or to type on a keyboard. The facilitator is not meant to guide the person's hand toward a particular response but to provide physical and emotional support to the autistic person. The physical support is meant to counteract such problems as poor eye/hand coordination, abnormally high or low muscle tone, difficulty in pointing with a single finger, and an inability to draw the hand back once a selection has been made (Biklen et al., 1991). Biklen and his colleagues summarize the method as follows:

> The method includes the following strategies: speaking to the student in the manner one would address a non-disabled student of the same age; explaining to the student that the hand-over-hand or wrist or arm support has proven effective for other students who have difficulty speaking or saying what they want to say; initially helping the student to not make mistakes, by pulling the student's hand back from a clearly wrong selection, for example if a student is pointing to a same letter for the third time; reminding the student to focus on the keyboard or other target; using structured questions, fill-in-the-blanks, simple match problems and similar 'set work' initially until a student achieves some success with typing; not 'testing' the student's competence; and giving students the opportunity to choose types of work they will do, including an opportunity for free expression (e.g., 'Is there anything else you would like to say?') (Biklen et al., 1991, pp. 163-164).

The procedure also involves (ideally) fading the physical support provided by the facilitator over time so that eventually the client can type independently.

Biklen (1990) argues that the problem faced by autistic people is not that they cannot understand language but that they have a problem with generating language. They have some unspecified neurological problem that makes it difficult or impossible to put words together into meaningful units or even to put sounds together into meaningful words. There are accompanying motor difficulties that also make it difficult for these people to independently communicate by pointing to letters or pictures. Therefore their

understanding of the world around them lies hidden away because of their inability to communicate it. Their inability to generate language is seen by people around them as an inability to understand. Helping these autistic people to control their motor impairments enough to point to pictures or letters in a communicative way, or even to type on a keyboard allows them to finally communicate their thoughts to the outside world.

You are probably thinking by now that this sounds a bit too much like a miracle cure and, if so, you are not alone. There has been a lot of controversy over facilitated communication both in the popular press in Australia and North America and in the academic journals. In the remainder of this chapter I will summarize some of the arguments that have been made in support of facilitated communication followed by arguments that have been made against it.

1.2 The Case For Facilitated Communication

There are several important questions that must be answered in order to make a convincing case for the effectiveness of facilitated communication. These include:

1.2.1 Where did the autistic clients obtain their newly discovered skills?

If the autistic clients don't talk, how can they suddenly start using sophisticated words and sentence structures when they are typing? Where did they learn spelling and grammar? Biklen and his colleagues (Biklen et al., 1991) argue that these people pick up their knowledge of grammar and spelling the same way all children pick up much of their knowledge of spoken language, through being exposed to it. Children learn most of their language skills before they start formal schooling and without much effort toward teaching from the people around them. The students that Biklen studied reported (through facilitated communication) that they had learned reading and math skills from television, from being read to, and from seeing flash cards. Thus they had been exposed to spelling, grammar, mathematics and had incidentally learned a fair amount about it without demonstrating any knowledge at the time.

1.2.2 Who is the communication coming from: the client or the facilitator?

By now you have probably already considered this question. Is the information that is coming out through the facilitated communication process—a process that involves two people—really coming from the client? Are the communications the thoughts of the client or is the facilitator, knowingly or unknowingly, affecting the communications? Is this really a case of the client acting as a human Ouija board for the thoughts of the facilitator? This is really the main sticking point, theoretically, legally, and morally. Theoretically, this question is at the crux of the validity of the procedure, if the communication is not coming from the client the procedure is worthless and those who are using it are just fooling themselves. Legally, it is important to know whose thoughts are being communicated. For example, imagine that an autistic woman and her facilitator

witness an accident together. Both give evidence in court, the facilitator by herself and the autistic woman through facilitated communication with the help of her facilitator. Do you now have two independent witnesses giving testimony or is there really just one—the facilitator? Morally, the question is similar. Imagine that, in an effort to give the facilitated autistic person as much self-determination as possible, you directly ask the person the question "Do you want to live with your parents or in a group home?" If it is really the client who is answering the question then that person has just made a life altering decision for him or herself. It may not have been the best decision but, like the rest of us, it is that person's decision to make. If, on the other hand, it is the facilitator who is answering the question the facilitator has (knowingly or unknowingly) taken the right of self-determination away from the individual[1]. This is likely to be a conflict of interest situation when the questions are related to treatment or communication.

The best evidence that communications are coming from the client comes when clients have reached the point where they no longer need the assistance of the facilitator. Biklen (1990) reports one case where a student answered a simple question correctly with one keystroke without the facilitator present. However, the quality of this response was very different from the full sentences this child was typing when the facilitator was present. Biklen also mentions a dozen students who were seen to be typing phrases independently, but it is not clear what independent means in this case. It appears to mean "without hand support on the arm or hand"(Biklen, 1990, p. 311) so it is not clear whether these students were communicating without a facilitator present, in which case the communications could reasonably be assumed to be their own, or with facilitators touching them elsewhere (e.g., holding their sleeve or touching their shoulder) in which case there is still ample opportunity for subtle movement from the facilitators to influence the students' responses.

One of the principles of the facilitated communication approach, as mentioned earlier, is that clients are not to be tested. The rationale for this principle is that the stress associated with being evaluated will interfere with the supportive emotional atmosphere necessary for communication to occur. Not only must clients be supported physically, they must also be supported emotionally. They must be made to feel at ease and to feel that the facilitator, and any others present, believe they are capable of high-level communication and value what they have to say. When the clients are being evaluated these conditions aren't met and facilitated communication may not occur.

This principle of avoiding testing makes it difficult for proponents of facilitated communication to prove that communications are originating with the client and not the facilitator[2]. Biklen et al. (1992) list six arguments that indirectly support the idea that communications are coming from the clients rather than the facilitators. These are listed in Table 1.1 . None of these arguments proves that the communication is coming from the client but Biklen argues that together they provide at least some support for

[1] Of course you may not be able to avoid having a third party make important life decisions for an individual. However, if you are going to do so, you should at least be aware of it and make efforts to ensure that the person who is making the decisions is the person who is most able to operate in the best interests of the individual. This person is not likely to be the facilitator.

[2] Of course, if you respect this principle, it also makes it impossible to prove that the communications are *not* coming from the client. You may be familiar with this type of argument from a medium or a channeler: "There is an unbeliever present. Your negative vibrations are interfering with the cosmic continuum. There will be no visitations tonight. No, you don't get your money back."

1. The style, speed, and accuracy of a student's fine motor control movement to the letters or keys is fairly constant across facilitators.
2. Individuals make typographical errors that are unique to them. Some individuals fairly consistently hit more than one key at a time when typing.
3. Many individuals produce phonetic or invented spellings that are unique to them and do not appear in the work of others, despite the fact that several individuals sometimes share a common facilitator.
4. Some individuals type phrases or sentences that are unusual and would not be expected from the facilitators.
5. Individuals sometimes produce content that is not known to the facilitator.
6. Perhaps most impressive and satisfying, through facilitated communication individuals reveal their personalities.

Table 1.1: Arguments that the communication is coming from the client rather than the facilitator (from Biklen et al., 1992).

that notion. Points 1, 2, and 3 are similar in arguing that there are idiosyncrasies in the students' movements, errors, and unusual spellings that are unique to the individuals and constant across facilitators. This would be expected if the students were the source of the communication but if the facilitators were the source you would expect such idiosyncrasies to be unique to the facilitators. Point 4 argues that communications that you would not expect from the facilitators, such as 'my facilitator is a stupid jerk and I hate him', are evidence that the communication is coming from the student. Point 5 makes a similar argument that information that is generated that the facilitator claims not to know, such as what the student did on the weekend, is evidence that the student is the source. Point 6 argues that, because the students are showing evidence of unique personalities through their communications (personalities very different from what they show without facilitated communication) the communications must be theirs. This final argument, although satisfying to Biklen and his colleagues, is not very compelling. It is not hard to imagine that even if the communication were coming from the facilitator it would reveal a personality for the student.

Calculator and Singer (1992) have demonstrated the best evidence to date that communications are originating with the clients. They tested five children who had been using facilitated communication for at least three months. All were diagnosed as either mentally retarded or autistic. These children were given two different forms of the Peabody Picture Vocabulary Test-Revised (Dunn & Dunn, 1981) in two testing sessions separated by two weeks. In one of the sessions the facilitator was not present. In the other session the facilitator was present and facilitated the child's responding. However, the facilitator wore earplugs and headphones playing white noise so that he

or she could not hear what was going on in the session. All five children scored below the published norms for the test for their age when the facilitator was not present. In other words, they were off the bottom of the percentile scale. However, when the facilitator was present, three of the five children improved markedly. In fact, two of the children scored extremely well, one at the 95th percentile and one above the 99th percentile. Since the facilitator could not hear what was happening in the test situation, Calculator and Singer argue that the three children who showed a marked improvement in their performance in fact showed valid communication.

1.3 The Case Against Facilitated Communication

1.3.1 What was the pretreatment condition of the autistic clients?

When Biklen talks about students who communicate "independently" he gives little detail on what their communication abilities were like before they began facilitated communication or even whether these students were classified as autistic. There don't seem to be any documented cases where clients who were clearly classified as severely autistic have had their communication support faded to the point where they can type independently without a touch from a facilitator. This lack of documentation from proponents of facilitated communication (but see Calculator & Singer, 1992) is in marked contrast with most of the experimental research that finds no support for facilitated communication (e.g., Datlow Smith & Belcher, 1993; Eberlin, McConnachie, Ibel, & Volpe, 1993; Hudson, Melita, & Arnold, 1993; S. Moore, Donovan, & Hudson, 1993; S. Moore, Donovan, Hudson, Dykstra, & Lawrence, 1993; Szempruch & Jacobson, 1993; Wheeler, Jacobson, Paglieri, & Schwartz, 1993) where the pretreatment condition of the participants is well documented. The pretreatment condition of the participants is important because it isn't very impressive that they can communicate by typing if their language skills weren't really that bad to begin with (I suspect that no one is impressed that I typed this whole paper myself without help from anybody).

1.3.2 Who is the communication coming from?

This is really the main sticking point and there have been numerous experimental studies addressing this question. One of the most interesting is one conducted by Wheeler, Jacobson, Paglieri and Schwartz (1993). This study took place at the O. D. Heck Developmental Centre in Schenectady, New York where two of the authors were on staff. The staff at the Heck centre were using facilitated communication and seeing very satisfactory results with it. Given the controversy surrounding the procedure, they decided to do a controlled study to show how well facilitated communication worked. There were 12 participants ranging in age from 16 to 30 years. All were diagnosed as severely or profoundly mentally retarded with strong autistic tendencies and all but one were nonverbal, the remaining participant displayed limited vocalizations and gestures. These 12 participants and the 9 facilitators who worked with them were the most competent producers of facilitated communication at the residential centre. All had been using facilitated communication for between 5 months and 1 year.

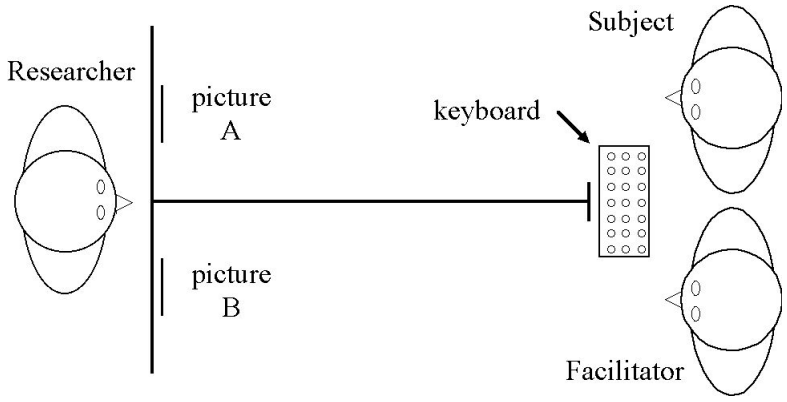


Figure 1.1: Experimental setup for Wheeler et al. (1993) study. Although the subject and the facilitator can type together, the subject can only see picture A and the facilitator can only see picture B.

The design of the study was quite elegant. The participants sat at the long end of a T-shaped barrier so that the client and the facilitator could each see down their own side but not down the other's side. See the illustration in Figure 1.1. At the far end of the T sat the researcher who displayed pairs of pictures, one on each side of the barrier. The participants' task was to name the pictures by typing the name. There were three conditions in this study: 1) The facilitated condition in which the subject was shown a card but the facilitator was not and the subject answered through facilitated communication, 2) The non-facilitated condition in which the subject was shown a card but the facilitator was not (as in the previous condition) but the facilitator was not allowed to touch the subject, although verbal prompts were allowed, and 3) the distracter condition in which both the subject and the facilitator were shown pictures and the subject answered through facilitated communication. This was a distracter condition because half the time the pictures shown to the subject and the facilitator were the same picture and half the time they were different pictures. The facilitator knew this was the case but of course did not know on which trials the pictures were the same. Each participant took part in 6 blocks of trials (2 of each of the three conditions). There were 5 trials in each block for a total of 30 trials per subject. Thus, across all 6 subjects there were 180 trials on which the subjects could show clear evidence of communication: all of the trials in the facilitated condition (12 subject X 10 trials each) and the half of the trials in the distracter condition in which the subject and facilitator saw different pictures (12 subjects X 5 trials each). Of these 180 trials the subjects showed communication on 0 (zero)! Furthermore there were about as many subjects who typed the picture the facilitator saw in the distracter condition when the pictures were different (12/60) and the subject never saw the picture they typed as when the cards were the same (14/60). Thus the subjects were not significantly more likely to correctly identify a picture they had seen than they were a picture they hadn't seen. The only sensible interpretation for these findings is that the communications were coming from the facilitators and not the subjects.

It is important to keep in mind that Wheeler and his colleagues were trying to *prove* facilitated communication and thus were very careful to run the experiment in such a way as to have every chance of success. The participants were given only 5 trials at a time so that they would not become distracted or fatigued. The subjects were paired with the facilitators with whom they showed the best communication in everyday activities. The test was explained to the participants as a way of demonstrating to others that the communications were coming from the subject (something the facilitators and, presumably, the subjects were eager to show) consequently there was no sign of distress among the subjects. Finally, the facilitators and the subjects were free to stop the procedure at any time if they weren't feeling in their best form. Despite all this, the only sign of communication appeared to be coming from the facilitators and not the subjects.

Wheeler and his colleagues are not the only authors to find such results. S. Moore, Donovan, Hudson, Dykstra, and Lawrence (1993) asked facilitators to generate 10 short-answer questions that the client would be able to answer (e.g., What is the name of your dog?). In one condition the facilitator asked the questions and the client answered using facilitated communication. In a second condition the questions were prerecorded and asked of both the client and the facilitator over earphones. In a third and critical condition the client heard the question over the earphones and the facilitator heard only music. Five of eight clients answered the questions in the first two conditions at better than chance levels (4/5 were 100% correct in the first condition and 3/5 were 100% correct in the second condition). Of those 5, 0 (zero) answered the questions at better than chance performance in the third condition.

One argument that has been raised against such findings is that autistic people have some sort of word finding difficulties so that asking them to name pictures or to give short answers are especially difficult tasks. Conversely, it may be that the tasks are so simple that the clients feel insulted and refuse to answer. Moore, Donovan, and Hudson (1993) responded to such criticisms by designing a study using a procedure suggested by two of the facilitators in the S. Moore, Donovan, Hudson, Dykstra, and Lawrence (1993) study. In the first phase of this study the facilitators left the room while the researchers presented an object of some sort to the client (e.g., a flashlight). The clients could touch the item and listen to a description of it. On returning to the room the facilitators could interact with the clients in any way they chose in order to find out what the client had seen. The second phase of the experiment was the same except that instead of showing the client an object, the client's primary caregiver talked to the client about a particular topic (e.g., the client's favourite television show). There were a total of 18 objects presented to 5 clients, 0 (zero) were identified. There were a total of 17 conversational topics presented to the 5 clients, 0 (zero) were identified. So there is no evidence of facilitated communication even in this study where the facilitators chose the task.

Let me quickly summarize some similar findings: Cummins and Prior (1992) report two studies, one with questions and one with patterns. When the facilitator could not hear the questions no client correctly answered a question through facilitated communication. When the facilitator and client were shown different patterns (unbeknownst to either) the pattern the client saw was reported through facilitated communication 0/36 times, the pattern the facilitator saw was reported 33/36 times.

Szempruch and Jacobson (1993) showed severely to profoundly mentally retarded clients pictures in the absence of the facilitators and then asked the facilitators to find out what was seen in their absence. In 127 trials with 23 clients there were 0 (zero) correct answers. Hudson et al. (1993) reported a single client study. When the facilitator could not hear the questions being asked of the client (questions generated by the facilitator) 0 (zero) of 26 questions were answered correctly.

The vast majority of the experimental studies show *no* evidence that any communication is coming from the clients and several studies present quite good evidence that the communication is coming from the facilitator.

1.4 So is Facilitated Communication Real or Not?

There you have it. You have seen a number of arguments that facilitated communication exists and is a very useful tool for helping autistic people communicate. On the other hand, you have seen a number of arguments that facilitated communication does nothing to help the clients communicate and that any communication produced by the procedure comes solely from the facilitator. The question that remains is fairly simple: Who has the better evidence?

To close, I'll leave you with a moral argument to consider. Donnellan, Sabin and Majure argue that in the absence of conclusive data one should go ahead and try facilitated communication based on the argument that one should make decisions in such a way that if an assumption turns out to be incorrect it will have the least dangerous effect on the client.

> The danger of not offering the option of facilitation to an individual is far greater than the danger of raising false hopes for families, as long as the individual with autism is not blamed for "failing" to communicate with this methodology. Indeed, families have survived disappointments far more costly in terms of their lives, fortunes, and the welfare of their children than they will if facilitated communication turns out to be a relatively ungeneralizable phenomenon (Donnellan, Sabin, & Majure, 1992, pp.71-72).

On the other hand, Wheeler and his colleagues argue that one must exercise great care in accepting communications when the source is unclear.

> One of the most important issues to be addressed in connection with facilitated communication deals with whether one recognizes as valid, and therefore acts upon, communications involving (a) accusations of abuse or mistreatment; (b) statements of self-determination and personal preference regarding living conditions, daily activities, and social relationships; and (c) self reports of health and medical relevance. There is potential for great harm here, both in accepting as valid communications that are not valid or, conversely, ignoring authentic expressions made by individuals who are dependent on others for their care. Current practice in the field of developmental disabilities emphasizes the primacy of individual choice.

> Accepting a communication subject to facilitator influence as representing client preferences undermines person-centered planning (Wheeler et al., 1993, p. 59).

Chapter 2

Categorizing Objects: The Usefulness of False Friends

Scott W. Allen

2.1 What is Categorization?

When a dog comes walking toward you on a street in broad daylight, you easily, and almost effortlessly, categorize it. That is, you decide that particular animal is a member of the category "dog" even though you have no previous experience with that specific dog. This is such a natural and fluent response that it is easy to overlook that you are even making a decision. It is only under conditions of missing or ambiguous information that you are likely to become aware that you are making a decision: Is that animal 20 feet away on a rainy night a poor forlorn kitty or a skunk? The behavioural consequences of such a decision are obvious.

2.1.1 Why do I care?

My main interest and the topic of this chapter is the process by which we make these decisions. There are at least two reasons why one might be interested in such a topic. The first, and by far the most important in my view, is that such questions are intrinsically fascinating. If you are a completely pragmatic type and cannot see the sense in investigating our mental processes simply because they are neat, then you might care about how people categorize things because there are formal categorization tasks that certain groups of people must learn. In your high school math and physics classes you will have learned to recognize categories of word problems in order to apply the appropriate formulae to solve them. In your research methods and statistics courses you will learn to recognize different sorts of experimental designs in order to apply the appropriate statistical analyses. Medical students must learn to diagnose disease categories

in order to apply the proper treatment. The more we know about the process by which people make such categorization decisions, the more effectively we should be able to teach these tasks.

Incidentally, this chapter is a specific example of a general theme in cognitive psychology: the search for mental processes. Cognitive psychologists are interested in the processes underlying such cognitive tasks as memory, reading, problem solving, and attitude formation. The question that is common to all these areas is "How do we go about performing these tasks, what is the process?"

2.1.2 How could one, in principle, learn about mental processes?

One obvious concern with trying to study mental processes is that you can't see them (of course that hasn't stopped physicists from studying black holes and quarks, or chemists from studying molecular structure). Because there is no way to directly access the processes in which you are interested, you must concentrate on the results of such processes. Some sets of results are consistent with some posited processes and others are not. Thus, mental processes can be studied by setting up experiments that are designed to allow you to choose between alternative possible explanations. For example, if you can devise an experiment that will give you one result if process A is at work and a different result if process B is at work then you can determine which process must have been operating simply by examining the results you obtained. The trick, and it is no mean feat, is to design such experiments. We will come back to this after looking at some theories of how we go about classifying things.

2.2 General Theories of Categorization

E. E. Smith and Medin (1981) introduced a categorization of categorization theories that has become standard. Or to put it another way, Smith and Medin introduced a classification of theoretical approaches to concept learning that has become standard.

2.2.1 The classical view

Early research in concept formation took the form of requiring subjects to induce rules from presentations of items with feedback (e.g., Bruner, Goodnow, & Austin, 1956). The subject would be presented with a series of examples, usually consisting of a set of four discrete dimensions (e.g. shape of items, number of items, colour of items, number of borders) each with two or three discrete values (e.g. circle, triangle, or square; 1, 2, or 3; green, red, or black). The subject's task was to try to classify each of the items as it was presented. With practice, and feedback on every trial, the subject was meant to induce the underlying classification rule. This procedure was seen as a reasonable model of how people learned to classify objects. The specific point of argument in classical theorist's work was that people were accomplishing the classification task by a process of hypothesis testing rather than the process of accumulating differential habit strength to stimulus elements, as advocated by the behaviourists of the time (but that is another chapter). But the underlying assumption common to both views was that

learning a concept was mainly a process of learning what elements of the stimuli were essential for correct categorization.

Bruner, Goodnow, and Austin's ambitiously titled book *A Study of Thinking* contended that both adults and children learned to apply categorization rules to specific examples and that through the presentation of many examples with feedback one could induce the underlying rules. Thus their research dealt with how people could learn different types of rules, what strategies were used, how difficult different logical forms of rules were and so on. The rules were defined by a set of criterial features that were necessary and sufficient[1] for classification, for example, an item was in the category if it was black and had triangles, or an item was in the category if it had two items or 1 border but not both. This type of representation would allow for elegant, efficient representation of concepts. This view of concept learning and representation has been called a classical view because the emphasis on discovering the necessary and sufficient conditions for membership in a concept has historically been a dominant concern in both philosophy and psychology (E. E. Smith & Medin, 1981).

There were a number of problems with the classical view (see E. E. Smith & Medin, 1981 for a detailed discussion). One of the most notable was the consistent failure to find necessary and sufficient features for many natural categories (try, for example, to think of a set of features, that you can see from the outside, that will tell you that an animal is a dog). If the process of classification is based on these features, they should be available, however that appears not to be the case. Further problems for the classical view come from the existence of a graded typicality structure (e.g. a robin is considered a more typical bird than an ostrich) and unclear cases (e.g. Is a rug furniture? Is a piano?, Medin & Smith, 1984). According to a classical view all members are equally good members providing they possess the singly necessary and jointly sufficient features for membership. This feature of classically defined categories also excludes unclear cases—if the necessary and sufficient features are present the item in question is a member of the category, if they are not present it is not a member.

2.2.2 The probabilistic view

In response to the problems of the classical view with natural categories, Rosch and her colleagues (e.g., Rosch, 1978; Rosch & Mervis, 1975) developed a probabilistic view of natural categories. Instead of requiring adequate knowledge of a concept to be a set of necessary and sufficient defining features, people were believed to normally learn a large set of features that were characteristic of, but not necessary for the concept (e.g., singing and flying are characteristic of birds but neither are necessary for an animal to be a bird). The basis for categorization was argued to be the accumulation of a sufficient number of these characteristic features to distinguish the target concept from other plausible concepts. This form of knowledge would account for the fact that people normally acted as if some concept members were better than others, behaviour that was difficult to account for if category membership were determined by singly necessary and jointly sufficient features.

[1] As an example of necessary and sufficient conditions, in order to be a bachelor, it is necessary that one be male. It is also necessary that one not be married. Being male and unmarried are together sufficient to make one a bachelor.

A popular form of model based on Rosch's work was that, through exposure to many items, some form of average or prototype was abstracted that consisted of all of the characteristic features. New items were assigned to a category on the basis of their relative similarity to the prototype of various categories. As with the classical view, subjects in learning experiments conducted under this view were never given classification rules when studying category learning, in this case because of the assumption that prototype-based category learning was automatic and natural. The probabilistic framework led to investigation of topics such as the difference in classification performance on typical versus atypical category members (e.g., robins *vs* penguins), what level of category was easiest to learn (e.g., mammals *vs* deer *vs* mule deer), and what level was learned first by children (e.g., animal *vs* bird *vs* robin).

For the purposes of the distinctions I want to talk about in this chapter, the important notion in this view is that of implicit abstraction of the prototype and, as in the classical view, a loss of specific episodic information with continued learning. That is, as people become more familiar with a category, they rely more on the abstracted prototype and less on their memory for specific examples. For example, Homa, Sterling, and L. (1981) took this position when they suggested that specific examples are used for generalization only early in a training sequence, before the learner has had an opportunity to abstract a prototype. In their work, the structure being learned was ill-defined, rather than the explicitly defined rules thought to be learned according to the classical view, but the assumption is that the influence of particular prior cases quickly falls out while the abstract information carries the weight of the decisions.

2.2.3 The exemplar view

More recently, as a response to the abstractive emphasis of the classical and probabilistic views, some investigators have emphasized the specific exemplars of categories (e.g., Brooks, 1987; Medin & Ross, 1989). Although there are, as in the probabilistic view, a number of specific models, the common logic is that rather than abstracting the essence of a category at the time of experience with the category exemplars, either by inducing a rule or by abstracting a prototype, one may simply make use of memory for specific prior examples to make a categorization decision for a new item. A new item is categorized on the basis of its similarity to individual prior cases rather than on its similarity to an abstracted prototype: This critter is a dog because it looks just like Rin Tin Tin (or Lassie, or Benjie) and Rin Tin Tin is a dog—not because it matches some prototypical or ideal dog.

According to an exemplar view there is no need for abstraction, either of a prototype or of defining rules. Most of the data that favoured a prototype explanation over a rule explanation could also be explained by a nonanalytic use of prior exemplars for classification. The presumed benefit of having information in this form is that such knowledge would provide a more flexible response to complicated, interacting conditions in the world. Once again, the research in this framework did not tend to give subjects rules for classification when examining their behaviour.

To summarize, a classical view of categorization suggests that through experience with items you infer a rule based on individual characteristics that are necessary for any item to belong to a category. For example, birds must have feathers and beaks and

FURIG	NOBAL
FEKIG	NEKAL
FUTEG	NOTEL
PURYG	POBYL
FYRIP	NYBAP
KURIT	KOBAT

Table 2.1: Training display from Whittlesea (1987, expt. 6). Note that each of the five items under each prototype is exactly two letters away from the prototype.

so on. A prototype view of categorization suggests that through experience with items you devise a prototype: a sort of average or ideal member of the category. You classify new birds on the basis of their similarity to that bird prototype. Finally, an exemplar view suggests that you don't derive any sort of rule or average through experience with items. All you do is store away all the items you've seen in memory. You classify new items on the basis of their similarity to specific items you've seen before.

2.3 Distinguishing Between Different Theories

As I just mentioned, an exemplar theory of categorization can generally explain all the same results that a prototype theory can explain but for different reasons. For example, take the finding that a typical category member (e.g., robin or sparrow) is generally classified faster and more accurately than an atypical category member (e.g., penguin or ostrich). According to a prototype theory, this makes sense because robins and sparrows have lots of characteristic features of birds—they are small, you find them in trees, they sing, they fly—while penguins and ostriches have few of these characteristic features. Therefore, when you go to match these items against your bird prototype (the average bird you have abstracted from all your experience with birds) robins and sparrows match better and are thus classified faster and with fewer errors.

According to an exemplar theory you make your classification of new items based on their similarity to specific items you have seen before. Unless you live in a zoo, most of the birds you have seen before will be fairly typical birds. Thus you will have stored away lots of examples of robins and sparrows and relatively few examples of penguins and ostriches. Having lots of stored examples of typical birds means that when you encounter a new typical bird there will be lots of similar examples leading to fast accurate classification of the new item. When you encounter a new atypical bird there will be fewer similar examples leading to slower and more error-prone classification.

This brings us back to the problem, mentioned earlier, of discerning what processes must be operating. Remember that we are interested in the process, so that, even if two different proposed processes like classification by prototypes and classification by reference to specific prior examples often predict the same consequences, we still want to be able to discriminate between them. This requires that one design an experiment where the two processes will predict different outcomes. An experiment by Whittlesea (1987) is an excellent example of such a design.

Whittlesea was interested in discriminating between classification based on prototypes and classification based on similarity to prior examples. In order to test which process the participants in his study were using, he set up experimental conditions where the distance of new items from the prototypes varied independently of their distance from prior examples.

In a training phase Whittlesea gave subjects two lists of pseudowords as shown in 2.1. He told the subjects that there were two categories as defined by the prototypes at the top of each list and that the other members of the lists all were the same as the prototypes except for two letters. In a later test phase he tested how well the subjects could perceive three types of items: The old items seen in training, a set of new items that were two letters different from the prototypes and two letters different from the training items, and a set of new items that were three letters different from the prototypes but only one letter different from the training items (see 2.2).

Whittlesea's design generates different predictions from the two theories of categorization. A prototype theory predicts that subjects will perform better on the new items that are two letters away from the prototype than on items that are three letters away from the prototype because they are closer to the prototype. However, an exemplar theory predicts that subjects will perform better on the new items that are three letters away from the prototype because they are only one letter away from the old items whereas the new items that are only two letters away from the prototype are also two letters away from the training items.

Whittlesea found that subjects performance in the test phase was best on the training items (76% correct), next best on the items that were only one letter away from the training items even though they were three letters away from the prototype (72% correct) and worst on the items that were two letters away from the training items, even though they were only two letters away from the prototype (70% correct). These differences, while they were not large, were statistically significant and showed that, at least under the circumstances of this experiment, people were classifying new items based on their similarity to prior examples not on their similarity to a prototype.

2.4 Combinations When We Are Learning: The Role of Practice Items

Virtually all of the research carried out on categorization until very recently has concentrated on requiring subjects to classify items without any explicit knowledge of the category structure. That is, in situations where there is a rule that defines the categories, the subjects in the experiments are not told the rule. This is because according to a classical theory of categorization once the subject has been told the rule, the action is all finished. Because the subject is thought to be inducing the rule from his or her experience with the individual items, once the general rule has been discovered, there is no more induction to be done and the subject will always give the correct answer. The same conclusion does not follow from an episodic view of categorization. By this view, even when the subjects know a rule, there will be a strong influence of the specific items that have been seen. My main interest in categorization is what happens in

2 away from prototype 0 away from training (training items)	2 away from prototype 2 away from training	3 away from prototype 1 away from training
FEKIG	FUKIP	PEKIG
FUTEG	PUTIG	FYTEG
PURYG	FURYT	PURYT
FYRIP	FYREG	FYKIP
KURIT	KERIG	KURET
NEKAL	NOKAP	PEKAL
NOTEL	POTAL	NYTEL
POBYL	NOBYT	POBYT
NYBAP	NYBEL	NYKAP
KOBAT	KEBAL	KOBET

Table 2.2: Test items used by Whittlesea (1987, expt. 6). The prototype for the top set is FURIG and for the bottom set is NOBAL.

this situation where there are two forms of information that people can use: the rule information and information about specific items. What happens when people are given an explicit rule for classifying items and then given practice with specific items?

One can imagine two extreme answers to this question. The first is that with practice the information about specific items quickly drops out and the person becomes faster and more accurate at applying the abstract rule. If this were the case, then apart from an initial training period, there should be no effect of specific prior examples on classification performance. The opposite extreme answer is that after practice with a sufficient number of examples, a library of instances is built up in memory which is used in classifying novel examples on the basis of analogy to these known items. If this were the case, then once a sufficient storehouse of instances is built up, the specific prior episodes will be doing all the work and the original rule would have no part in classification performance. The use of specific episodes would be mimicking the effects of rule use.

One of the few researchers investigating categorization (loosely speaking) where subjects are provided with a rule followed by practice with examples has been Brian Ross (e.g., Ross, 1984; Ross & Kennedy, 1990). Teaching his subjects to solve probability word problems, Ross (1984) investigated their use of memory for prior examples when applying the rules they were learning to novel problems. He found that subjects' use of "remindings" of prior examples (where a "reminding" is an explicit retrieval of a prior problem) was affected by the similarity of the prior example to the current problem and that remindings had a substantial effect on the successful solution of the problems. Subjects performed better on a novel problem when it had a story line similar to a problem in the training set that required a similar solution than when it had a neutral story line. Furthermore, subjects were worse than in the neutral condition when the test problem had a story line similar to a practice problem that required a different solution. So, at least in the early stages of learning rules for classifying word problems, people successfully use their memory for similar prior examples to help them

solve new problems.

Research that is more closely related to categorization is work pertaining to medical diagnosis. Medical diagnosis is, after all, simply a categorization of patients on the basis of the disorders they exhibit. Allen, Norman, and Brooks (1992) investigated the influence of the practice examples that were used when medical students were being taught to diagnose dermatological disorders. The medical students were given rules for six diagnostic categories. Examples of the rules are shown in 2.3. After studying the rules, the students were given practice applying the rules with 24 training items (four examples of each of the six different disorders). The training items were colour photographic slides of dermatological lesions chosen from the slide collection of a teaching dermatologist. The students saw the training set five times with corrective feedback each time. Following this training they were given a set of new slides to diagnose. The nifty trick in this study is the relation of the test slides to the training slides. All the students saw the same set of test slides, however there were two different training sets. While each training set contained four examples of each of the six diagnostic categories, the specific examples used were different in the two training sets. The items in one training set were chosen to bias the diagnosis of certain critical items in the test set in one direction while the items in the other training set were meant to bias the diagnosis of the critical items in a different direction. 2.1 shows a critical (chameleon) item and similar training items from the two different training sets. The critical item in 2.1 is actually a case of Lichen Planus. However, note that according to the diagnostic rules in 2.3 you can make a reasonable argument for either Lichen Planus or Contact Dermatitis for this item. That is why we called these critical items "chameleon" items, they could be seen in more than one way. The interesting question is whether the medical students' propensity for diagnosing this item as a case of Lichen Planus varies depending on what they saw in the training series. That is, do students who saw a training series containing cases of Lichen Planus similar to the chameleon item (training set A) call the chameleon Lichen Planus more often than students who saw a training series containing cases of Contact Dermatitis similar to the chameleon item (training set B)? If they do, it means that they are using their memory for the specific items they saw in training to help make their diagnostic decisions since the rule information was identical for all students.

The Contact Dermatitis items that are similar to the chameleon are the false friends that were alluded to in the title to this article. They are false friends because, although they are similar to the chameleon item (which, remember, is a case of Lichen Planus) they belong to a different category. Thus, if you rely on your memory for these items to help you classify the chameleon item, you will be led astray. It is these false friends that allowed us to tease apart the students' reliance on rules from their reliance on prior examples. Without these items and their unusual relationship to the chameleon items we would be stuck with the more common and uninformative situation where similar old items give you the same answer as the rule.

Table 2.4 shows the results of this experiment. There are two ways the critical items could be biased by the training set: Toward the correct diagnosis (as in training set A for the critical item shown in Figure 2.1) and toward a specific incorrect diagnosis (as in training set B in Figure 2.1). Thus there are three ways in which a student could diagnose the critical items: a) Correctly, b) incorrectly but consistent with the incorrect

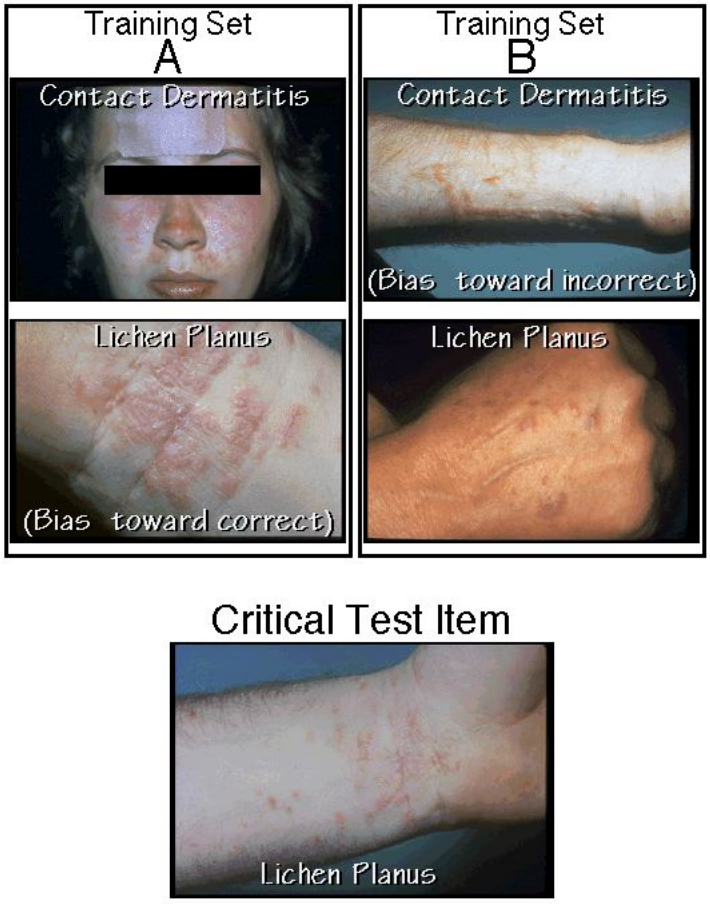


Figure 2.1: An illustration of the basic experimental manipulation used by Allen, Norman, & Brooks(1992). In this illustration, the critical test item, seen by both groups, is a case of lichen planus. For the subjects trained on set A the most similar looking training item is also a case of lichen planus. For subjects trained on set B the most similar looking training item is a case of contact dermatitis. Good quality colour slides were used in the experiment.

Rules

Contact Dermatitis – Well-defined areas of coalescing vesicles, may be red, with surrounding erythema, irregularly shaped, showing pattern of contact. Prone to secondary infection. Most common on areas where skin is exposed to contact with an irritant.

Lichen Planus – Pink or purplish coloured, shiny, polygonal-shaped papules which may coalesce to form plaques with a slight scale and areas of healthy skin in between. They may also have milky-coloured streaks. The most likely areas are the wrists and forearms.

Definitions

Erythema: Redness resulting from vasodilation.

Papule: Small circumscribed superficial solid elevation up to 1 cm in diameter.

Plaque: A disc-like lesion formed by the coalescence of papules or nodules.

Vesicle: Papule containing fluid, .1 to 1 cm in diameter.

Table 2.3: Two of the six rules used by Allen, Norman, & Brooks (1992) along with definitions of terms that were supplied.

Bias in Training	Diagnosis consistent with		
	Correct Bias	Incorrect Bias	Other Incorrect
Correct	89	7	4
Incorrect	42	37	2

Table 2.4: Diagnostic accuracy (%) for medical students (from Allen, Norman, & Brooks, 1992).

bias, and c) some other incorrect diagnosis. For the critical item shown in 2.1 the corresponding diagnoses would be: a) Lichen Planus, b) Contact Dermatitis, and c) one of the remaining four diagnoses. As shown in 2.4 when the training set biased the students toward the correct answer they were correct 89% of the time and answered with the incorrect bias only 7% of the time. However, when the training set biased them toward a particular incorrect diagnosis they were correct only 42% of the time and answered with the incorrect bias 37% of the time. Thus there was a substantial difference in their performance that could only be due to the specific items on which they received practice. A second experiment showed that this effect remained even when there was a one week delay between the training set and the test set. Thus the effect of similarity to training items appears to be relatively long lasting.

Although the previous study demonstrated a strong effect of previous examples on medical students' diagnoses, it did not suggest whether more expert diagnosticians would show a similar sensitivity to prior cases. To answer this question the same

authors (Brooks, Norman, & Allen, 1991) used a different procedure to look at the performance of more expert diagnosticians. In the first phase of this study first year family medicine residents and practising family physicians (averaging 15 years from graduation) were shown a series of 30 photographic slides of dermatological lesions. They were told the correct diagnosis for each slide and asked to rate how typical each lesion was of the diagnostic category it represented. The main purpose of this phase was to expose the physicians to correctly labelled examples of dermatological lesions. In the subsequent test phase the physicians were asked to diagnose 60 more slides. For our purposes there were two important types of items in the test phase. Both types of items belonged to the same diagnostic categories that were shown in the first phase. The difference was that some items, the same-diagnosis/similar items, were not only in the same diagnostic category as one or more training items but were also visually similar to one of the training items. The other type, the same-diagnosis/different items, were in the same diagnostic category as one or more training items but were not visually similar to one of the training items. Thus, any difference in performance on the same-diagnosis/similar and same-diagnosis/different items must be due to the earlier presentation of the similar items and not due to a difference in application of the formal rules known to these physicians.

For both the residents and the family physicians there was a strong effect of the similar training items on their subsequent diagnoses. The residents diagnosed more of the same-diagnosis/similar items correctly (65%) than they did the same-diagnosis/different items (53%). The same held true for the practising family physicians (70% for same-diagnosis/similar items and 52% for same-diagnosis/different items). The importance of similar prior cases is clearly not confined to only those people who are first learning a rule as was the case with the medical students in the Allen, Norman, and Brooks (1992) study. It remains even with experienced physicians with an average of 15 years of dealing with skin disorders. It is also worth noting that half the residents were tested immediately following the first phase and half were tested following a two week delay. There was no reduction in the difference between the similar and different items with the delay suggesting, as in the one week delay condition in the previous experiment, that the effects of similar prior cases are quite long-lasting.

2.5 Conclusion

It appears that specific examples used in practice do not fade quickly leaving the abstract rule to run classification performance. Rather, the specific examples used in training appear to have a life of their own, hanging around to influence the classification of similar items up to two weeks later. Furthermore, individuals one would expect to have exceedingly well learned rules and experience with thousands of cases of dermatological lesion (the practising physicians) continue to show classification performance that is strongly affected by similar prior cases. This suggests that even experts use their memory for prior cases to help them classify new items. Thus, even in situations where there are well-learned, formal classification rules, a person's classification performance will not be based solely on those rules but will be influenced by specific prior cases. That is, even in situations that should favour a process consistent with a

classical theory of classification, the classification process is at least partly based on memory for specific prior examples—a process consistent with an exemplar theory of classification.

Chapter 3

Attachment Theory
Jean Choi

Recall the last time that you encountered a particularly traumatic event in your life. Who was the first person you wanted to call? Now think about your relationship with that person. Do long separations evoke a need for contact? According to attachment theory, you have established a bond with that person, and the nature of your relationship with that person is influenced by the first relationship in your life—that between you and your primary caregiver.
Attachment theory was first developed by John Bowlby (1969), and since its inception, its effects have been far-reaching throughout various areas in psychology. In developmental psychology, it has influenced theories of parent-child relationships; in social and personality psychology it has influenced theories of peer, family, and romantic relationships in adults; and in clinical psychology it has influenced theories of the disruption of attachment. Attachment theory comprises two principal components: a normative component and an individual difference component (Simpson & Rholes, 1998). The normative component addresses modal, species-specific patterns of behaviour, whereas the individual difference component addresses the variations from these modal behavioural patterns (see Figure 3.1. The first empirical studies of attachment theory addressed the patterns of attachment in young children, emphasizing individual differences in the quality of attachment (Ainsworth, Blehar, Waters, & Wall, 1978).

3.1 Bowlby's Attachment Theory

The development of John Bowlby's attachment theory was influenced by his clinical work with institutionalized boys. This work suggested that disruptions between a child and his primary caregiver were associated with physical and social problems later in development. In addition, Bowlby was influenced by the observation that across human cultures and in several primate species, infants displayed a consistent sequence of reactions when separated from their caregivers. Immediately following separation,

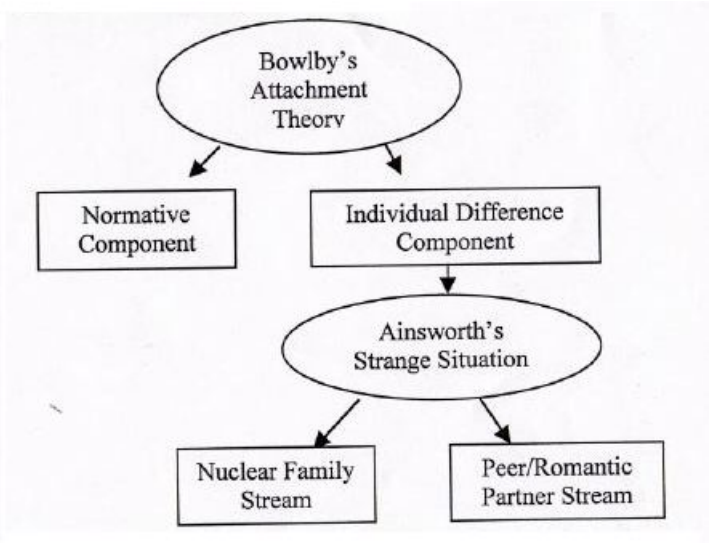


Figure 3.1: Research traditions of attachment theory (from Simpson & Rholes, 1998).

infants cried and screamed for their caregivers, characterizing the protest stage. These protests were thought to be a good strategy to promote survival as they often brought the caregivers back to the vulnerable infant. If persistent protests did not retrieve the caregiver, infants would enter the second stage—despair—characterized by inactivity and the cessation of protests. If protests failed to retrieve caregivers, the next best survival strategy may be to avoid excessive movement and loud protests, behaviours that may draw predators, for long periods of time. Following despair, infants not reunited with their caregivers would enter a third and final stage, detachment. During this phase, infants resume normal activity and begin to act independently. Detachment may act to clear the way for the formation of new affectional bonds with other potential caregivers willing to provide resources that the infants would require to survive.

Attachment originates from innate behaviours that are important for the newborn infant's protection and survival, particularly during the time of human evolution; Bowlby referred to this time period as "the environment of evolutionary adaptedness". Because infants are born completely dependent upon their caregivers, they require a set of mechanisms for maintaining proximity to those caregivers. An infant's set of behaviours—such as crying, sucking, clinging, orienting, and later, smiling and babbling—communicates its needs to others, elicits caregiving, and facilitates the formation of an attachment bond between the child and caregiver.

The attachment bond is a specific type of affectional bond . An affectional bond is the attraction of one individual to another that is persistent and involves a specific individual. The affectional bond may be characterized by four fundamental attributes: specificity, endurance, emotionality and biological functionality. *Specificity* refers to the target of the attachment behaviour; the attachment behaviour is directed to a specific individual within the proximate environment. The loss of that specific individual

is distressing despite the presence of other close individuals. *Endurance* refers to the persistence of the attachment bond; the bond is presumed to remain influential throughout the individual's life. *Emotionality* refers to the emotional intensity associated with the attachment bond, which often arises during its formation and maintenance. Any threat to the bonded relationship is accompanied by explicit protest and anxiety. Distress vocalizations elicit the primary caregiver's attention, signalling the caregiver of the infant's emotional state. Cries and calls function to lead to a quick reunion and protection from any impending danger. This protection from danger comprises the bond's *biological functionality*. Separation protests have become incorporated into the infant's behaviours and resemble fear responses associated with unfamiliarity.

A unique characteristic of the attachment bond that does not necessarily apply to other types of affectional bond is the security and comfort associated with the attached individual (Ainsworth et al., 1978). The caregiver acts as a reliable base for exploration so that the child may actively explore its environment with the security of possessing a safe base to which the child may return when reassurance is needed. Glances to the caregiver, who is also maintaining surveillance for potential perils, reassure the child, allowing the child to explore the environment, leading, in turn, to the development of a sense of autonomy.

3.2 The Development of the Attachment Bond

As infants develop, their attachment behaviour passes through four phases, becoming increasingly refined and targeted to the primary caregiver.

3.2.1 Phase 1: Nonfocused orienting and signalling

The infant begins life with a set of innate, proximity-promoting behaviour patterns that orient the infant toward others. In the newborn, these include crying, making eye contact, clinging, cuddling, and responding to caregiving efforts. These behaviours function to draw individuals toward the infant. At this stage, there is little evidence that the infant is attached to the parents; attachment behaviours are emitted, but without direction toward any specific individual.

3.2.2 Phase 2: Focus on one or more figures

By 2 or 3 months of age, the infant becomes discriminative and directs attachment behaviours to familiar individuals. The infant may smile more at familiar individuals and less at strangers, indicating recognition. However, in this phase, attachment has not fully developed. The infant still favours several people with proximity-promoting behaviours, and no one person has yet become a safe-base. Infants at this phase show no special anxiety at being separated from their parents and no fear of strangers.

3.2.3 Phase 3: Secure base behaviours

By 6 or 7 months of age, the infant normally has formed a genuine attachment. For the first time, the infant uses the attached individual as a safe base from which to explore the environment. By this age, the infant is able to crawl more freely and can move toward the caregiver or entice the caregiver to come to the infant. Attachment behaviours shift from mostly *proximity-promoting* to *proximity-seeking* behaviours. These proximity-seeking behaviours lead the infant to the caregiver and include approaching and following the caregiver. In this phase, the attachment figure has become the infant's stable social-emotional base. A 7-month-old, for example, may explore a novel environment, but periodically look to his or her "mom" as if seeking reassurance that all is well. The presence of the primary caregiver allows the infant to explore and learn about the environment from a safe base. Also characteristic of this phase is the development of stranger anxiety in which strangers begin to evoke alarm or withdrawal in the infant, as well as increased separation protests, which includes the infant's distress when separated from the primary caregiver. These patterns of behaviour correspond with the infant's increased mobility, which may lead it to wander from its protector.

3.2.4 Phase 4: Partnership

By 18 months to 2 years of age, the child acquires insight into the primary caregiver's feelings and motives, and a reciprocal relationship forms between the child and caregiver.

3.3 Quality of Attachment

Although attachment behaviour is deemed universal, the nature of attachment relationships is subject to individual differences. Mary Ainsworth, a student of Bowlby, developed a laboratory technique, the Strange Situation procedure, in order to assess the type of attachment relationship in one- to two-year-olds (Ainsworth et al., 1978).

This technique consists of eight stages:

1. Parent and infant are introduced to the experimental room. 30 seconds
2. Parent and infant are alone. Parent does not participate while infant explores. 3 minutes
3. Stranger enters and talks to the parent, then approaches infant. Parent leaves inconspicuously. 3 minutes
4. First separation episode: Stranger's behaviour is geared to the infant. 3 minutes
5. First reunion episode: Stranger leaves, parent greets and comforts infant, then leaves again. 3 minutes
6. Second separation episode: Infant is alone. 3 minutes

7. Continuation of second separation episode: Stranger enters and gears behaviour to the infant. 3 minutes

8. Second reunion episode: Parent enters and greets the infant; stranger leaves inconspicuously. 3 minutes

The infant's behaviour during each stage comprises the basis for classifying the infant's relationship with the caregiver into one of three attachment categories (Ainsworth et al., 1978):

Secure attachment. In the experimenter room with the caregiver, the infant uses the caregiver as a secure base; the infant actively explores the environment, occasionally glancing at the caregiver during this time. The infant may or may not cry when the caregiver leaves, but shows some signs of distress. With the return of the caregiver, the infant seeks the caregiver and is easily soothed. This type of relationship characterized approximately 63% of Ainsworth et al.'s (1978) sample.

Insecure–avoidant attachment. When in the experimenter room, the infant tends to ignore or avoid the caregiver, playing in isolation. The infant does not show distress when the caregiver leaves. The infant continues to ignore the caregiver in each of the reunion episodes by looking or turning away. This type of relationship characterized approximately 29% of Ainsworth et al.'s (1978) sample.

Insecure–ambivalent attachment. When in the experimenter room, the infant demonstrates ambivalent behaviour, ranging from clinging to the caregiver to hitting and/or pushing the caregiver away. The infant fails to explore the new environment, preferring to stay with the parent. When the caregiver leaves the room, the infant shows a great deal of distress. The infant remains upset and angry when the caregiver returns and is difficult to console. This category characterized approximately 8% of Ainsworth et al.'s (1978) sample.

A fourth category—Insecure–Disorganized Attachment—was introduced after Ainsworth et al.'s (1978) original study, characterizing high-risk caregiver-child pairs, whereby confusion, apprehension, and undirected movements are displayed by the children (Main & Solomon, 1990).

3.4 Attachment in Adults

The primary focus of attachment research has been the relationship between the child and caregiver. However, Bowlby (1969) has always maintained that attachment bonds are sustained throughout the life span. Research on attachment in adults has followed two research traditions: (1) attachment to peers or romantic partners and (2) attachment in nuclear families.

3.4.1 Attachment to romantic partners

According to Zeifman and Hazan (1997), the infant-caregiver relationship and the pair bond between two romantic partners share the same underlying mechanism: attachment. During development, the importance of attachment between the infant and caregiver shifts to the pair bond. Pair-bonding as a reproductive strategy is proposed to have evolved due to the adaptive advantage conferred through shared care of the young (Lancaster & Lancaster, 1987). Attachment is defined as the flexible, pre-existing mechanism that ensures that "two individuals would be highly motivated to stay together and resist attempts to separate them" (Zeifman & Hazan, 1997, ; p. 246). In other words, attachment has been co-opted by evolution to serve a second function in addition to protection of the infant—the maintenance of long-term pair bond relationships. Parallels drawn between the infant-caregiver relationship and adult pair-bond relationship provide evidence that attachment is the common mechanism in both types of relationship. Some parallels include: (1) the similarities of the adult's response to separation and loss of a pair-bond individual to the infant's response to separation from the caregiver; for example, the sequences of reactions to disruptions of a pair-bond relationship follows the "protest, despair, and detachment" sequence of a disrupted child-caregiver relationship; (2) the significance of familiarity as a selection factor for both adults and infants in their choices for attachment figures; there is evidence for assortative mating in adults (i.e., we tend to like what is familiar) and infants' enhanced attraction to the familiar caregiver; (3) evidence that attachment bonds are most frequently limited to parents and sexual partners (Hazan & Zeifman, 1994); (4) the common need for physical contact in both infant-caregiver relationships and pair-bond relationships; and (5) the essential biological requirement for attachment, as evident in the studies of institutionalized children(e.g., Kaler & Freeman, 1994) and health decrements associated with disrupted attachment(see Zeifman & Hazan, 1997).

3.4.2 Attachment to peers

Aspects of attachment theory have also been applied to peer relationships. Some propose that early attachment relationships form the prototype for future attachment relationships, including peer relationships, via the internal working model. The internal working model is the set of expectations about the reliability of attachment figures (i.e., "Can I depend on the attachment figure being there for me when I need her or him?"). The early working model is assumed to be a function of the infant-caregiver relationship that becomes internalized over time and comprises the prototype not only for future relationships but for the individual's personality. The internal working model consists of concepts of one's self and the concepts of other individuals, and the interaction between these concepts influences personality development. Kim Bartholomew (1990) developed a model of peer relationships based on these components of the internal working model, which continue operating throughout one's life, affecting adult relationships.

3.4.3 Attachment in the nuclear family

The focus in the nuclear family research tradition addresses the inter-generational transference of the infant-caregiver attachment relationship, measuring the way in which adults discussed their past relationships with their own parents and how these relationships are linked to the attachment classifications of their own children. The first and most widely used interview measure of adult attachment, the Adult Attachment Interview (AAI) (George, Kaplan, & Main, 1985), is a 60-minute interview that gauges the feelings adults have about their own childhood relationships. The AAI was specifically developed to predict the Strange Situation behaviour of the respondents' infants. The principle goal of the AAI is to determine how information about past attachment figures is structured, organized, and stored. This information is presumed to reflect the childhood internal working model. Research addresses both the long-term stability of attachment patterns across the life span and the role that early care played in the formation of attachment bonds. Research results have suggested that there is some correspondence between the maternal internal working model and the quality of attachment in the next generation such that mothers who describe characteristics of a secure relationship with their parents are more likely to have a securely-attached child themselves (Benoit & Parker, 1994).

Chapter 4

Sex Differences in Spatial Abilities in Humans: Two Levels of Explanation
Jean Choi

Think about how you find your way around when you're visiting a new city. As you travel from place to place, do you take note of buildings and their respective locations, or do you generally try to keep track of which direction you're going and how far you've gone? Similarly, when you read a map, do you notice the landmarks and use them to find your way, or do you look at the scale to code distances and use the direction indicators? These different methods we use to find our way have been found to vary as a function of sex. Males tend to use more cardinal directions (e.g., north and south) and distances (e.g., miles or kilometers) than females, whereas females tend to rely more heavily on landmarks (e.g., buildings) and their relative directions (e.g., right and left). This example illustrates the application of spatial abilities, and how differences in spatial abilities become evident in everyday activities.

4.1 Spatial Abilities

In general, spatial ability is defined as the "skill in representing, transforming, generating and recalling symbolic, nonlinguistic information" (Linn & Petersen, 1985, p. 1482). It is required in many facets of our daily activities, including map-reading, navigating, and remembering where objects are located in space (such as your keys!). Spatial abilities have been studied extensively in both humans and nonhumans using a variety of methods, and perhaps the most consistent finding in this large body of research is the prevalence of sex differences. In humans, spatial research, for the most part, has been based on various pencil-and-paper tests (see Figure 4.1).

Based on a meta-analysis of almost three hundred individual studies using various

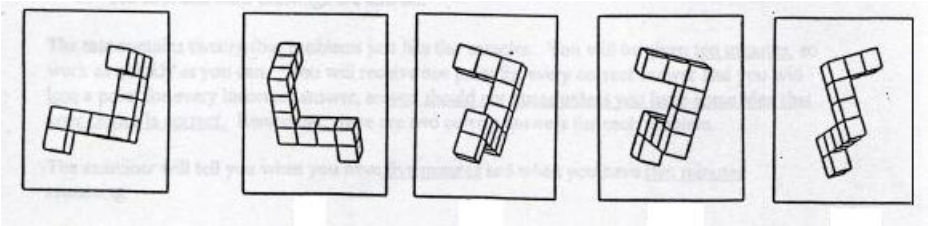


Figure 4.1: Sample item from a spatial rotation test. Participante are asked to mark the two drawings from the 4 on the right that match the drawing on the left.

spatial tests, Voyer, Voyer, and Bryden (1995) concluded that spatial abilities comprise three dissociable categories: mental rotation, spatial perception, and spatial visualization. Mental rotation is the ability to manipulate, rotate, twist, or invert two- or three-dimensional objects. Testing this ability involves asking an individual to visualize objects in different positions by mentally rotating them (as in Figure 4.1). Spatial perception refers to the ability to determine spatial relationships despite distracting information. The water level test, for example, requires one to indicate how the top of the water line would look in a tilted bottle. This test requires one to ignore distracting information such as the tilt of the bottle. Spatial visualization is the manipulation of complex information through several stages. A prototypical test item requires one to mentally transform a two-dimensional object into a three-dimensional object. In general, males tend to show an advantage in all of these spatial categories, although the size of the differences between males and females depends on the category of spatial abilities. Mental rotation tests yield the largest and most reliable sex difference, followed by spatial perception tests, which show a moderate effect size for sex (i.e., a moderate sex difference), and spatial visualization tests usually yield the smallest effect size. However, another category of spatial abilities that has recently received attention is object location memory. This is a unique category in that it is the only one that produces a female advantage (Silverman & Eals, 1992). Object location memory assesses how well an individual remembers the location of objects.

But, why are there sex differences in spatial abilities? The answers to this question are varied and can be addressed using two levels of explanation.

4.2 Levels of Explanations

Two levels of explanation can be considered when accounting for spatial sex differences: proximate and ultimate explanations. Proximate explanations address the direct mechanisms affecting behaviour; these may include hormonal or neural mechanisms. Ultimate explanations address the adaptive significance of the behaviour; that is, how might the behaviour lead to the survival and/or reproduction of the individual. Ultimate and proximate explanations are complementary, as opposed to competing, explanations. An example of how behaviour may be explained at these two levels can be seen when considering why male songbirds sing in the spring. An ultimate ex-

planation would focus on the possible adaptive nature of the male's singing. Over evolution, a singing male may have attracted more females than a male who did not sing; thus, singing males would have had access to more mating opportunities, increasing the chance of having offspring. Or, it may be that the male's singing may have warned other males to stay away from the singer's territory. This may have led to fewer fights for the singing male, increasing his chance of survival and ultimately, reproduction. Either way, those males who carried the genes that led them to sing had a greater chance to reproduce and leave more offspring than the males who did not sing. This would lead to a greater proportion of the singing males' genetic material in the next generation. Over time, the genes that led to singing would be pervasive among males and finally, become characteristic of the population. A proximate explanation of singing behaviour in males would emphasize the environmental and physiological stimuli underlying the behaviour. For example, a proximate explanation for the male's singing would include the effects of increased daylight in the spring and the increased production of testosterone, triggering singing behaviour. As you can see, the ultimate explanation is directed to the survival or reproductive value of the behaviour, whereas the proximate explanation addresses the direct mechanisms of the behaviour.

4.2.1 A proximate explanation of spatial sex differences

Differential effects of gonadal hormones comprise one proximate explanation of spatial sex differences. The prevailing hormonal theories presume that differences in the relative concentrations of estrogen and testosterone affect how the brain is organized, which, in turn, may lead to differences in various cognitive abilities, such as spatial abilities. The lateralization of function (i.e., differences in the degree to which each side of the brain is specialized for cognitive processes, such as spatial abilities) may be related, at least partially, to differential exposure to gonadal hormones (Hampson & Kimura, 1992). Given that females have higher concentrations of estrogen and progesterone, and males have higher concentrations of testosterone, these hormonal differences may affect the degree of lateralization. In most individuals, particularly in right-handed people, language is represented to a greater degree in the left hemisphere, whereas the right hemisphere seems to be more involved in nonverbal processes, such as spatial abilities. Many theories presume that sex differences in lateralization are related to sex differences in cognitive abilities, specifically, that adult males tend to be more lateralized than females. That is, poorer performances on spatial tests in females are due to either bilateral representation of language, interfering with development of spatial abilities (e.g., McGlone, 1980) or bilateral representation of spatial ability, leading to greater inefficiency than unilateral representation (Harris, 1978).

Although there is little direct evidence of the effects of gonadal hormones on spatial abilities in humans, there is much indirect evidence based on a variety of populations and methodologies. Hampson and her associates, for example(Hampson, Rovet, & Altmann, 1998), investigated spatial abilities in children with a syndrome called congenital adrenal hyperplasia (CAH). This condition is characterized by excess levels of prenatal androgens due to the overproduction of a testosterone-like androgen, androstenedione. CAH females, therefore, become relatively more masculinized than other females, due to exposure to higher levels of androgens as fetuses. Although

generally treated immediately after birth through surgery and hormonal therapy, the exposure to prenatal hormones appears to have an enduring effect on the brain. When compared to control girls (those girls who have not had exposure to excessive levels of androgens), CAH girls demonstrate superior spatial abilities, supporting the notion that prenatal hormonal levels have an enduring effect on the organization of the brain, despite surgical and hormonal intervention after birth.

In addition to organizational effects of gonadal hormones, there is also evidence of activational effects. For example, changes in gonadal hormones in adults seem to have an effect on spatial abilities, as demonstrated by a study of transsexuals undergoing hormonal therapy before a sex change operation (Goozen, Cohen-Kettenis, Gooren, Frijda, & Poll, 1995). Women who were undergoing androgen therapy, and men who were undergoing anti-androgen and estrogen therapy were tested twice: before hormone therapy and three months after the onset of their therapy. Spatial scores of women increased, presumably due to the exposure to androgens, whereas spatial scores of men exposed to estrogen decreased. This pattern of results suggests that small increases in androgens enhance spatial abilities in women, whereas decreases in androgens and increased exposure to estrogen depress spatial abilities in men. Other research has focused on estrogen using normal hormonal fluctuations associated with the menstrual cycle in women. For example, Silverman and Phillips (1993) demonstrated that scores on a mental rotation test varied within the same individual, depending on when she completed the test. Women performed more poorly on this spatial test when estrogen levels were elevated during the midluteal phase (defined as 16-23 days after the onset of menses), compared to performances during menses, when levels of estrogen were relatively lower. However, despite the growing evidence that testosterone and estrogen affect spatial abilities in humans, the exact nature of the inter-relationships between estrogen, testosterone and spatial abilities remains unclear.

4.2.2 An ultimate explanation of spatial sex differences

In the previous section, we found that differential levels of gonadal hormones appear to affect spatial abilities and may account for some of the differences between males and females—an example of a proximate explanation. The ultimate explanation addresses the origin of these spatial sex differences. Recall that ultimate explanations address the adaptive nature of specific behaviour. Evolutionary psychology, as outlined in other chapters of this textbook, emphasizes the ultimate explanations of behaviour, using the principles of natural selection. In human evolutionary psychology, the mind is conceptualized as a set of specialized information-processing mechanisms that developed during the environment of evolutionary adaptedness (Tooby & Cosmides, 1992). Humans spent the majority of their time in evolutionary history living in a hunter-gatherer social structure, in which group members gathered plants for sustenance, and supplemented them by hunting animals. The skills required to solve particular problems faced by our ancestors during the environment of evolutionary adaptedness, such as the necessary skills to gather plants and hunt animals, would thereby be favoured by selection pressures. Natural selection would have caused the human mind to evolve specific mechanisms to effectively solve adaptive problems, leading to a set of specialized cognitive modules. These modules have evolved over generations, leading to function-specific

adaptations. Spatial behaviour is an example of a product of differential pressures.

The evolutionary origin of sex differences in spatial abilities has spawned several theories; one that is specific to human evolution is the hunter-gatherer theory. During the environment of evolutionary adaptedness, males and females worked cooperatively in small kin groups, in which sexual division of labour was pervasive in all facets of daily activities (Tooby & Cosmides, 1992). The division of labour persisted over evolutionary history during which selection pressures favoured different spatial skills in males and females. This division of labour would be particularly evident in the acquisition of food whereby males primarily hunted and women gathered. The resulting products of this division of labour were exchanged, with males providing meat, a highly valued source of protein and fats, and females providing plant foods, comprising much of the daily sustenance rations.

Because males were involved in hunting, scavenging, and travelling over large areas for food, effective skills related to navigation over a large area were essential for males. Hunters would have had to roam over a large territory, subsequently finding their way home after pursuing prey. Superior spatial skills related to traversing through a large environment and directed motor tasks such as projectile accuracy would have been subject to selection pressure in males. In support of this supposition, Watson and Kimura (1991) found that males were more accurate in an overhand-throwing task and were better able to intercept a flying ball, even after accounting for differences in sports-related experiences. Further support for the hunter-gatherer theory was found in our study focusing on navigation. In this study, males and females were led through a circuitous path through a wooded area and required to "find the shortest route back to the origin" (Silverman et al., 2000). Participants were also asked to indicate the direction of the point of origin at various points in the route. Males chose the shortest route more frequently than females and were more accurate in indicating the absolute direction of the point of origin. These navigation abilities were associated with three-dimension mental rotation ability, suggesting that the ability to mentally rotate objects and tracking in an unfamiliar environment void of distinguishable features are mediated by the same underlying mechanism, or at the least analogous mechanisms.

in the environment of evolutionary adaptedness, females tended to be involved in foraging and searching for non-mobile plant foods that had to be identified and localized within a larger array of vegetation (Silverman & Eals, 1992). Foraging and gathering by females were performed in a smaller area, compared to hunting, and would have led to selection pressures associated with object location memory within a static spatial array. Silverman and Eals (1992), for instance, demonstrated that females were better able than males to recall the locations of objects. In their test, participants were asked to "examine the objects" for a minute, and then given a second array in which some of the objects had been moved. Females were better able to indicate which objects in the second array had been moved than were males. In a further study, undergraduates were asked to wait for two minutes in a tiny office that was contrived to resemble a graduate student's office. The office was outfitted with various objects such a pair of glasses, books, and a coffee mug. As expected, females were better able than males to freely recall the location of various objects in the office. More recently, Neave, Hamilton, Hutton, and Pickering (In press) replicated and extended these findings. They found that females were better than males at recognizing target plants within a larger array of

plants that varied in size and complexity.

4.3 An Application of Spatial Sex Differences: Map Reading

Given that males tend to perform better in mental rotation, spatial visualization, and spatial perception tests, and females tend to perform better in object location memory, the question arises as to whether these sex differences emerge in everyday behaviour. We explored this question using a direction-giving paradigm, based on a novel map (see Figure 4.2)

We asked individuals to learn the shortest route from a given point of origin to a point of destination, and to provide written directions from memory to other adults (Choi & Silverman, 1996). In their directions, females tend to refer more to landmarks and relative directions whereas males tend to use more distances and cardinal directions. An example of a set of directions containing primarily distance and cardinal directions would be "Go southwest for about a mile, then go south for another 5 miles. Then turn west, and travel for another 4 miles". Directions with more emphasis on landmarks and relative directions would be "Go left until you see the high school, then turn left. Go past the bank and across the railroad tracks. Turn right at the mall, and go past the car rental place". Although males and females used different methods to give directions, each method was equally effective. We inferred that females' better object location memory may lead to greater emphasis on landmarks and relative directions when reading a map, and males' better mental rotation ability may lead to more emphasis on distances and cardinal directions.

These sex differences have been replicated using a three-dimensional virtual environment, in which the locations of landmarks were either manipulated or remained the same (Sandstrom, Kaufman, & Huettel, 1998). This navigation task required participants to find a hidden target as quickly as possible. Females and males performed at a comparable level when landmark information was available and reliable. However, when landmarks were moved in an unpredictable fashion, the ability to navigate in the virtual environment was more disrupted in females than in males. Females also required more time when there were no landmarks and only geometric cues, such as the shape of the room, providing further evidence that males and females depend on different methods to navigate. This greater female reliance on landmarks is consistent with our findings, as described above, that males were better able to navigate in an environment with very few distinguishable landmarks (Silverman et al., 2000). Because the ability to navigate in this type of environment was correlated with higher scores on the mental rotation test, it suggests that there is some specific ability that is shared by both tasks. This association also provides some indication that performances on individual spatial tests are manifested in everyday behaviour.

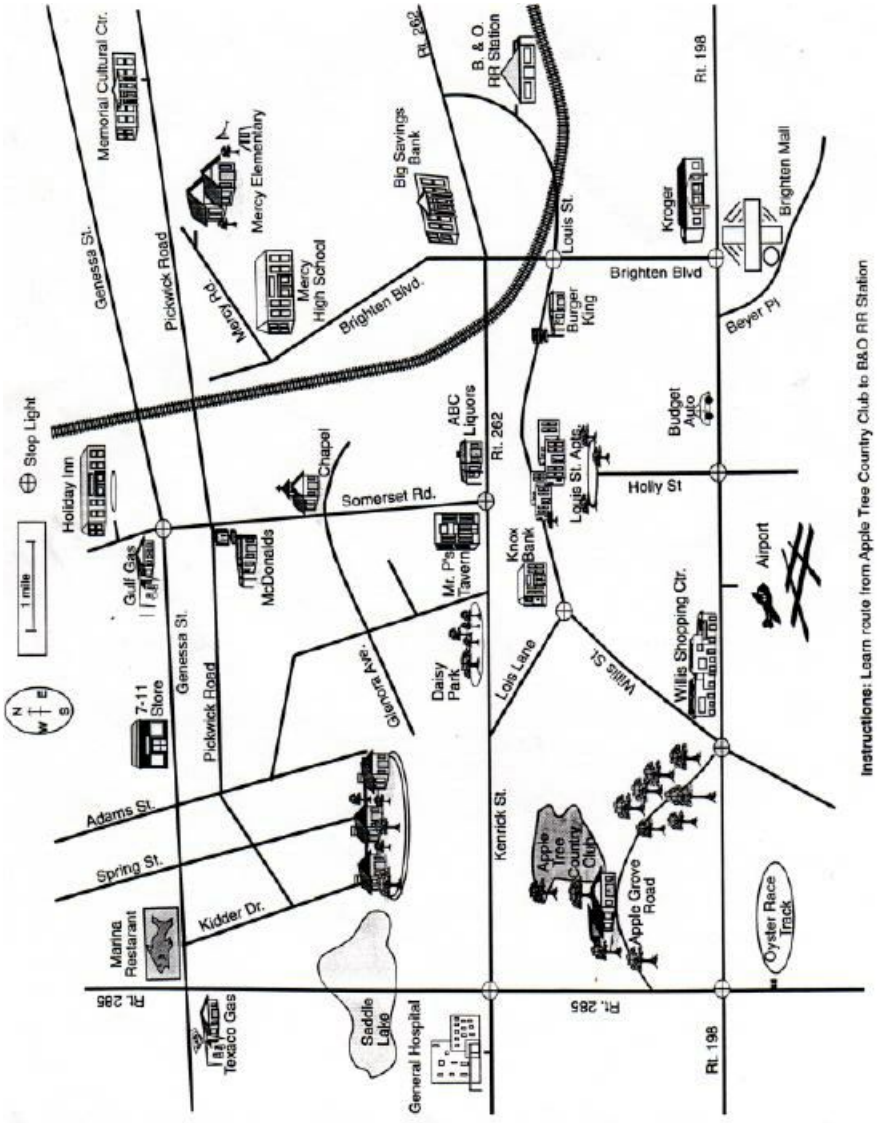


Figure 4.2: Sample map used for route-learning test.

4.4 Conclusion

Spatial abilities comprise a broad cognitive capacity, which consists of dissociable categories. There is evidence that the extent and direction of sex differences depend on the spatial category; for example, males perform better on mental rotation tests and females perform better on object location tests. There are also related differences in map reading and navigation methods, suggesting that individual differences in spatial tests are reflected in everyday behaviour. The origin of spatial sex differences can be addressed using two complementary levels of explanations; proximate explanations address the direct mechanisms underlying spatial abilities, whereas ultimate explanations address the evolutionary origins of spatial sex differences. At the level of proximate explanations, research suggests that sex differences in spatial abilities are associated with differences in gonadal hormones. At the level of ultimate explanations, such differences are argued to have evolved due to different selection pressures related to the division of labour.

Chapter 5

Albert — Portrait of a Cheese Freak
Margaret Forgie

> . . . It's a matter of instinct;
> it's a matter of conditioning;
> it's a matter of fact.
> You can call me Pavlov's dog,
> ring a bell and I'll salivate —
> how'd you like that . . .
>
> — from "Brian Wilson"
> by Barenaked Ladies (Words and music by Steven Page, 1992).

Perhaps one of the most famous discoveries in all of psychology (after all, it's in the lyrics of a popular song!) was that of classical conditioning, a form of learning in which an animal comes to make an association between two previously unrelated events. This discovery is credited to the Russian scientist, Ivan Pavlov (Brennan, 1986; Pavlov, 1960). In the popular version of this "experiment", Pavlov is said to have trained dogs to salivate to the ringing of a bell by consistently pairing the bell's ring with the presentation of food (a stimulus that reliably produces salivation in the first place).[1]

This principle, that animals can learn to associate two stimuli and to respond to a cue or signal that was previously unimportant to the animal, is a fundamental form of learning that can be observed in virtually all animal species, from the simple sea

[1] Despite this popular conception, it should be noted that Pavlov was neither a psychologist, nor specifically studying learning at the time. His discovery was serendipitous and occurred during the course of studying digestive processes in dogs, work for which he was awarded the Nobel Prize in 1904 (Brennan, 1986). He also did not discover the phenomenon by ringing a bell, but rather by observing the reaction of his dogs' salivary output to the sound of the approaching lab worker whose job it was to feed the animals.

slug *Aplysia californica* to human beings. In fact, the process of classical conditioning is everywhere in daily life, as in, for example, the fact that you and I, like the dogs in Pavlov's original studies, have learned to salivate to the sound of supper being prepared.

But classical conditioning is not just about drooling for food. It is a general process that is in place to prepare us to deal with recurrent situations in our environment. In order to understand just how powerful the phenomenon can be, and how insidious (particularly when the behavioural response being conditioned is one that is undesirable), we need to examine it in more detail; and I can think of no better subject for this exercise than to introduce you to Albert, cheese freak.

Somewhere between the lowly sea slug and humans lies the domestic cat (sometimes I think closer to sea slugs than other times). I have lived with cats since my childhood, and, also being a long-time student of behaviour, I have spent many hours pondering their various behaviours. Albert, who has lived with me for about 10 years now, is a very large (15 to 17 pounds depending on the season), male, tabby version of the species. And Albert has attitude, lovin' attitude. He's the biggest, marshmallow cat I have ever met, and more than a bit dumb, to boot. He is also just plain large. And so, when he wants something, he usually gets it (in fact, many of my friends are afraid of Albert! It's the size of the eyes apparently).

Albert has a lot of interesting behaviours, but I think some of the most interesting have to do with "people food". If it is Albert and food, then the food is probably cheese. The moment that the cheese block is removed from the refrigerator with its accompanying tell-tale crinkle of the cheese wrapper and Ziploc bag, Albert turns those big eyes on you, and begins his approach. Over time, Albert has come to respond to all sorts of crinkling plastic no matter where he is in the house at the time. Every time the sequence is the same: open refrigerator—make crinkle—and Albert orients and appears out of nowhere. Did I mention Albert was large and friendly…? And terribly hard to get rid of once that approach has started.

5.1 Classical Conditioning

So what's going on here? Albert is demonstrating a *conditioned response* to plastic wrap in much the same way as Pavlov's dogs demonstrated a conditioned response to a ringing bell. The process by which this happens is outlined in Figure 5.1. Initially, Albert experienced cheese and began to approach when he detected its presence. In psychological jargon, these two events are referred to as the *unconditioned stimulus* (UCS—the cheese) and the *unconditioned response* (UCR—the approach behaviour). The term "unconditioned" refers to the fact that the stimulus reliably produces the response before conditioning starts. The *conditioned stimulus* (the crinkle of the plastic wrap), so-called because it is the one to which the animal will become conditioned and is unrelated to the response initially, is present each time the cheese is produced. In this case, because I am in the habit of buying a big block of cheese and keeping it in a Ziploc bag, the conditioned stimulus reliably predicts the arrival of the unconditioned stimulus (a situation that is necessary for conditioning to occur). After many repeated experiences of crinkle⇒cheese, crinkle⇒cheese, the crinkle of the wrapper alone (or the presentation of the conditioned stimulus) is enough to bring Albert running. Al-

bert's approach behaviour in these latter cases is referred to as a conditioned response because it is elicited by the conditioned stimulus.

5.2 Stimulus Generalization and Discrimination

Examining Albert's behaviour further demonstrates two other principles associated with conditioned responding: *stimulus generalization* and *stimulus discrimination*. These two principles help us to understand how conditioned behaviours can be increased or made more refined. First, stimulus generalization is a process whereby an animal no longer restricts its response to the specific conditioned stimulus that first elicited it, but rather to a variety of closely-related stimuli. The fact that Albert now responds to pretty much any crinkly plastic that generates a sound similar to the Ziploc bag/cheese wrapper combination is a good example of this principle. We can say that Albert's behaviour has *generalized* to all plastic wrappers (imagine his surprise, though, when the sound produces a bag of tomatoes!). Second, Albert also shows stimulus discrimination. In this case, an animal refines its conditioned response to include only the most likely stimuli to produce the desired result. Albert demonstrates this principle when he restricts his attention to crinkly plastic that is immediately preceded by the opening and closing of the refrigerator door. As he rarely makes a response if the bag did not emerge from the refrigerator, he has apparently learned to discriminate fridge-plastic from not fridge-plastic.

5.3 Operant Conditioning

Albert's behaviour also provides an example of another form of learning, that of *operant conditioning*. This form of learning is most often associated with the name of B.F. Skinner (Brennan, 1986; Skinner, 1938), and involves changing the frequency of a behavioural response (that is, increasing or decreasing the rate and persistence with which an animal will perform a behaviour). In this form of learning, the change in behaviour is produced by making rewards (e.g., the giving of food to a hungry animal) and punishments (e.g., the spanking of a naughty child) when a particular response is made. For example, when we require children to complete their homework before allowing them to watch television, we are providing them with a reward (television viewing) for performing a behaviour (doing homework). It is our hope that by offering the reward, the behaviour of homework-doing will increase. In these simple cases, rewards serve as positive reinforcers, increasing the likelihood that the behaviour will be performed again; whereas punishments serve as negative events designed to decrease behaviour. In Albert's case, when he makes his approach response, sometimes he gets a piece of cheese and sometimes he does not. When he does not (because the bag does not contain cheese or I just decline to give him some), he is usually surprised, but persistent, and this often leaves me in the ridiculous position of trying to show Albert that the bag does not contain cheese so he will go away. Thus, the times that Albert gets the cheese have reinforced the behaviour of responding to that crinkly sound.

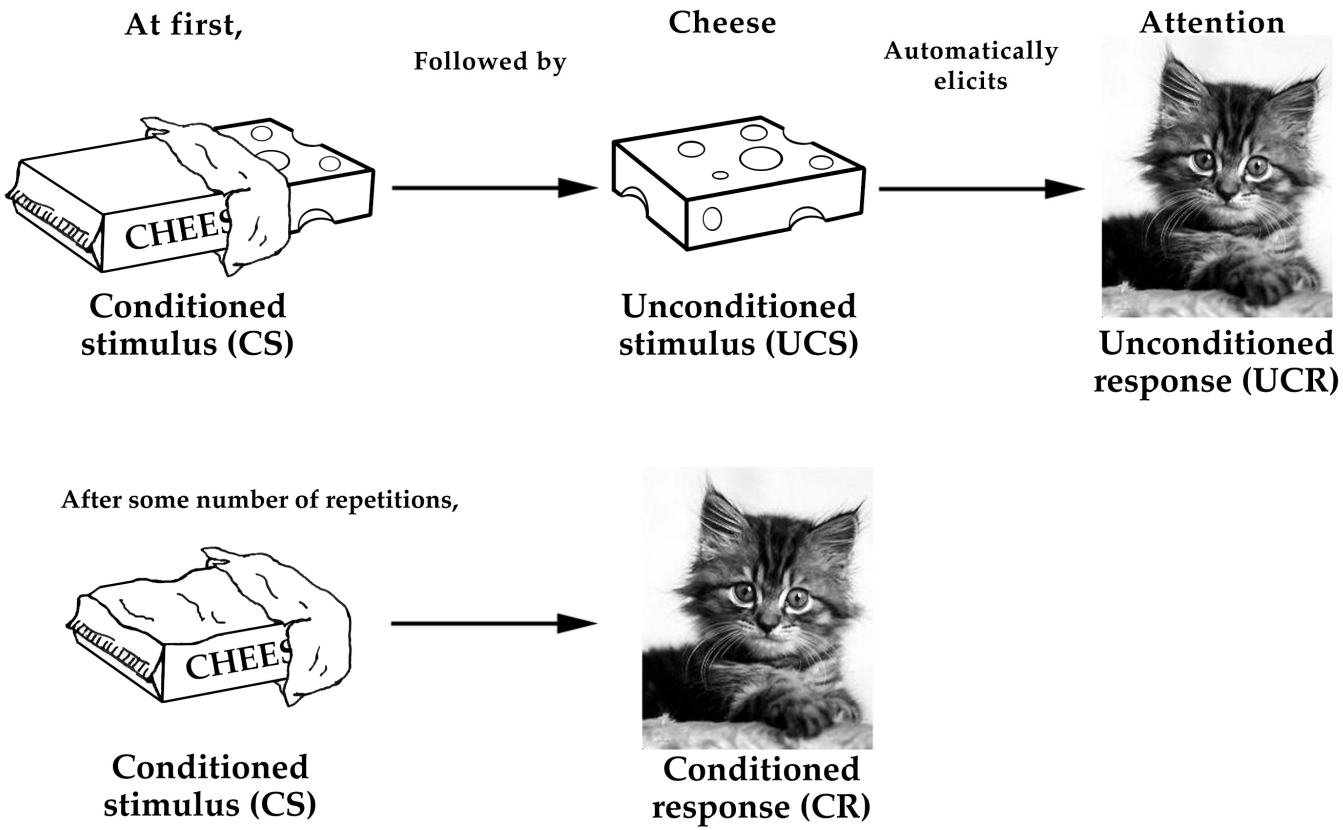


Figure 5.1: The process of classically conditioning Albert.

5.4 Schedules of Reinforcement

Albert does not receive the reward every time, however, and sometimes he even gets an active negative response when he is too persistent. These aspects of his behavioural interactions highlight two additional features of operant conditioning: the schedule or pattern of delivery of reinforcements, and the consistency with which they are delivered, are strong influences on the continuation of the behaviour. *Extinction*, or the decline and eventual disappearance of the behaviour, can be brought about by repeatedly and consistently not rewarding the response. When *partial reinforcement* occurs (such as when the bag only sometimes contains cheese, or when I only sometimes give him a piece), however, the behaviour is extremely difficult to stop. One example of this form of schedule of reinforcement maintaining an undesirable and difficult-to-stop behaviour (in some cases) is gambling (particularly devices such as slot machines and the like). Like Albert's approach response to crinkly plastic, each time you pull the handle of a slot machine there is a chance that you will win a jackpot. Most of the time this does not happen, but the few times that it does (particularly if you have a run of successes early on in the process) maintain the "hopeful" behaviour of "Maybe this time will be the next lucky break!". Thus, it is doubtful that Albert will stop seeking the opportunity to do cheese for a long time to come.

Although Albert's response to cheese makes for a humorous way to introduce the concepts associated with these two forms of learning, it also highlights the fact that complex patterns of behaviour can be produced and maintained in response to signals or cues from the environment that are not connected to the situation in an immediately obvious way. Furthermore, patterns of behaviour can come to be "cemented-in" by a combination of these cues and the rewarding properties of the stimulus of interest. In addition to characterizing all aspects of our daily lives from childhood to old age, these general phenomena are thought to play a role in a number of maladaptive (or undesirable) patterns of behaviour that we humans seem to want to engage in with vigour (e.g., those associated with administering drugs of abuse). In the lecture we will explore how these forces operate in the initiation and maintenance of the drug-seeking behaviour that is an important component of the addictive cycle of behaviour.

Chapter 6

Building Brains. . . Building People

Margaret Forgie

The role of experience in shaping the development of the nervous system, and thus of behaviour, has been addressed by several authors in this volume. There is no doubt that differential experience in our postnatal environment is one of the most powerful tools in shaping the individual architecture of the brain. This experience is crucial to the normal brain development of most species, and it is this individual brain architecture that gives each of us our own unique pattern of strengths and weaknesses in ability, personality, and so forth. As you can read in "Nurture's Little Secret"(see Chapter 22, page 167), experience produces this change by eliciting activity in the brain; and it is this activity, or lack thereof, that allows for the growth of brain cells and the pruning back and remodelling of connections between cells.

The effects of experience, however, do not occur in a unilateral fashion, affecting each of us in an identical manner. Why? The answer is simply that we all begin life outside the womb with different brains—the result of a long and complex chain of events that begins at the earliest phases of the development of a human child. Thus, although newborn human infants share many behavioural characteristics—as one might expect from members of the same species with the same basic "human" brain plan, each individual infant is also fundamentally unique. For example, all human infants appear to have a special attraction to human faces, and the characteristics of the infant's visual capabilities are such that when held in the arms the child is placed at the best viewing distance to take in that face (Berndt, 1997). These basic patterns of behaviour shared by all infants presumably prepare them for the more or less consistent aspects of the postnatal environment. On the other hand, infants are also behaviourally different, as is clearly evident in temperament (Berndt, 1997). Newborn babies differ in their activity (speed and intensity of behaviour), reactivity (reaction to stimuli), emotionality (depth

of emotional response), and sociability (willingness to associate with others). Thus, differences in behaviour that occur from the moment of birth, interact with the postnatal environment to produce an environment that is also unique for each child.

Both the prenatal foundation of nervous system development and the sculpting effects of postnatal environments and experiences are necessary to produce a functional human brain (and, thus, a behaviourally functional human being). To offer a simple analogy, we can liken the process of the development of an individual's behavioural repertoire to the building of a house. A house without a carpenter (and a brain without experience) is just a nicely formed block of cement: the basics of the plan are there, but there is nothing to complete the building or to decorate the finished product. On the other hand, it is also the case that a house without an adequate foundation (or a brain that has been somehow compromised in early development) will not allow the carpenter to construct a solid building that can withstand the stresses and strains of years of living. At a minimum, the layout of the foundation will constrain the layout of the rooms that the carpenter can build, and give the house its unique appeal. In this way, the foundation for individual differences in brain organization is put into place early in development, and is then modified by postnatal experience.

6.1 Prenatal Development of the Brain

The nervous system begins its development very early in the prenatal period and continues well into adolescence and young adulthood. For example, the frontal lobe of the cerebral cortex is a very long-developing brain structure that does not finish the developmental process until well after puberty (Kolb, 1995). How much of this developmental process is completed before birth differs between species (Kolb, 1995). Beginning at about two to three weeks post-conception, the first hint of the nervous system appears in the developing human embryo with the formation of the neural tube (K. Moore & Persaud, 1998). This tube-like structure gives rise to the central nervous system. Once this tube is formed, cells that will become brain cells begin to divide in the wall of the tube, and then migrate to form the basic structure of the brain. Once cells have reached their targets, they begin to form the characteristic shape of the type of brain cell that they are, and to hook up to other cells, and so start the process of making connections. Although the brain has begun to form the highest of brain control centres (the cerebral cortex) by the fifth month of pregnancy, in human beings the process of cell migration is still occurring at the time of birth. The process of the "blooming" of connections peaks at about one year of age (Greenough, Black, & Wallace, 1987; Kolb, 1995). From this point, throughout the remainder of postnatal development, the primary forces include the activity-dependent remodelling of those connections discussed in Chapter 22. Thus, events occurring in the prenatal period have potent influences on the basic construction of the brain.

6.2 Where do Individual Differences Come From?

Given the obvious complexity of the processes of prenatal development, it is easy to see that abnormal construction of the brain could contribute to developmental disabilities. But what about the normal variation in behaviour that characterizes the vast majority of the human population? Where do these more subtle, individual differences in behaviour come from? In short, they result from the same variations in the combined forces of genes and environment that can produce abnormal development, but on a less dramatic scale. All development can be seen as an interaction between the unfolding of a genetic program and the modification of the program by environmental factors. Many persons, however, persist in assigning causality to one or the other of these influences, viewing behaviours as arising from the singular influence of either a genetic or an environmental factor (the age-old "nature" [genes] vs. "nurture" [environment] debate). This dichotomy has some appeal, for it can be seen to simplify the incredibly complex issue of behavioural development. Merely restricting the argument to a simple nature vs. nurture debate, however, does not really do much to help to explain the exact nature of the phenomenon in question. In his 1949 book, *The Organization of Behaviour*, the noted Canadian psychologist, D.O. Hebb attempted to move away from this simple dichotomy in a discussion of the potential contributing factors to individual differences in intelligence. These factors are also useful in categorizing the potential contributors to all individual differences in behaviour.

The first of these factors is, of course, genes. Unless we are the member of a set of identical siblings, each of us has our own unique genetic endowment. The integrity of this genetic endowment is essential for normal development, but it is not a blueprint for the whole of the process. Although each cell contains the same genetic information, whether or not a gene will be turned on or off, a process known as gene expression, depends on myriad other factors. When a gene is turned on or off it changes what the cell can and cannot do. The factors that influence gene expression are enormous in number. Thus, the final product of development is genes + environment = a new form, a process known as *epigenesis*.

The second factor includes chemicals and other experiences occurring in the prenatal period. This period includes a host of physiological factors specific to the individual foetus such as hormones, growth factors, the presence of teratogenic agents (agents from the outside that are able to produce birth defects), diseases from the mother, etc. Among the most powerful agents in shaping the development of the individual nervous system (and indeed of the entire foetus) are hormones. In particular, the hormonal events that characterize the differentiation of the developing child into a biological male or female, have profound effects on the wiring of the central nervous system.

As with all developing structures, the brain shows critical or sensitive periods: windows of time during which it is especially susceptible to specific influences from the environment. The timing of these critical periods depends largely on the specific brain area concerned. Because of the length of time over which the brain develops, the whole of the prenatal period contains sensitive periods for some aspect of brain function. Furthermore, the exact nature of the effect of an environmental agent will depend upon

what the environmental agent is, and what particular process is occurring at the time the influence is received. Many of these influences are an expected and necessary part of development (such as the exposure of the brain to circulating hormones in the foetal and maternal blood), but others are unexpected and produced by exposing the foetus to harmful agents (or by depriving it of beneficial ones). For example, a deficiency in folic acid during and before the early phases of pregnancy is associated with a failure of the neural tube to close, leading to disorders such as spina bifida (Batshaw, 1999). Consumption of alcohol by the mother during pregnancy appears to disrupt cell migration in the foetal brain (among myriad other deleterious effects, Ferriero & Dempsey, 1999).

The remaining factors primarily influence the brain in the postnatal period. The third factor includes chemical agents that one might experience after birth. Many environmental toxins, such as lead, cause severe brain damage in developing children (Batshaw, 1999). It is not unlikely, then, to postulate that variations in our exposure to chemical agents throughout our lives will have an effect on our brain. In fact, such environmental exposure is implicated in diseases such as *multiple sclerosis* and *Parkinsonism* (Kolb & Whishaw, 1994). The fourth and fifth factors are the effects of experience in the postnatal period. Hebb (1949) proposed that both general experiences, such as might be encountered by all members of the species, and specific experiences that would affect a certain individual, help to shape the developing brain. An example of both of these classes of experience can be seen in language. The human brain is wired up in such a way that at birth (or shortly thereafter), children respond to language-related sounds regardless of what language is spoken around them; their brains are prepared for the general experience of language. Somewhat later, children begin to respond only to the sounds of their native language, and start the process of acquiring that native tongue. This is an example of specific experience: What specific language did you learn as your first language in childhood? Last, most of us will experience some type of brain damage at some point in our lives, whether due to injury, disease, or the aging process (Kolb, 1995). These traumatic events may alter the wiring of the brain, and the individual's behaviour, irrevocably (see Chapter 10, page 77).

Thus, the development of the brain is a cumulative process, and one that requires both a solid foundation of growth prenatally, and a safe, experience-filled environment postnatally. The nature of that foundation is unique for each of us, and it prepares us to respond in our own individual way to all the variations of experience awaiting us in our life outside the womb.

Chapter 7

Stress in the Workplace
C. Gail Hepburn

Every day there seems to be a news story about the rise of stress in the workplace. Workers are reporting higher levels of stress in the workplace than ever before and these stress levels have costs both for workers and for the organizations in which they work. After reading this discussion you will have an understanding of what stress is (and what it is not), what the costs of stress in the workplace may be, why workplaces are not completely full of 'stressed-out' workers, and how researchers have conceptualized the causes of stress in the workplace.

7.1 The Stress Process

It is important to distinguish between the events or situations that may be perceived as stressful and the workers' experience of stress. The term *stressor* is applied to any objectifiable event or situation in the workplace that has the potential to cause stress (e.g., a deadline, extreme heat or cold, a rude supervisor). *Stress* is a worker's psychological response to stressors (e.g., feeling threatened or challenged). *Strain* is another important term in the stress process—strain results from stress, think of strain as the consequences or costs associated with feeling stress (e.g., anxiety, stomach upset). In general, workers are faced with stressors, they feel stress, and then they experience strain. These terms will be elaborated on throughout this discussion.

7.2 The Costs of Stress

7.2.1 Stress affects workers

Strain comes in a variety of forms. Workers may experience *psychological strain*. Psychological strain describes those outcomes that are affective or emotional in nature

as well as those outcomes that are cognitive in nature. Individuals experiencing psychological strain will report such affective responses as feeling anxious, depressed, or overwhelmed. They may also report cognitive disturbances such as being distracted or having difficulty concentrating or remembering things.

Physical strain is another outcome of stress. Physical strain includes minor health complaints such as headaches or stomach upset. More serious health conditions are also associated with the experience of stress. They include coronary heart disease, high blood pressure, and asthma.

Workers also may experience *behavioural strain.* Behavioural strain outcomes are those behaviours or actions that are engaged in by workers who are experiencing stress. They can include nervous habits, situation avoidance, smoking, and alcohol consumption.

7.2.2 Stress affects organizations

One would hope that organizations are concerned about the health of their workforce and interested in reducing stressors in their workplaces. Unfortunately, concern is often not a sufficient argument to induce stressor reduction. However, organizations do agree that stress is associated with certain costs and these costs are of interest to organizations—they affect their bottom line. Costs of worker stress to organizations include, but are not limited to, reduced performance, increased absenteeism and turnover, and workplace accidents.

Worker performance may decline for those workers experiencing stress. Workers suffering from psychological strain may not be performing at their best: how well do you study when you are anxious or distracted? Workers experiencing stress may choose to avoid the stressor and stay away from their workplace. This form of behavioural strain will likely affect work performance and, if the worker quits, may force the organization to engage in recruitment activities; often a costly exercise. Researchers have recently linked worker stress to accidents in the workplace, but more research needs to be done to confirm this relationship. However, given what we know about strain it does seem a logical outcome.

We can agree that all workplaces contain stressors, those events or situations that we find stressful. Some workplaces are inherently stressful, they always contain events or situations that cause stress. For example, the work environment of firefighters, police officers, air traffic controllers, and dentists (imagine if most of your interactions at work were with people who were frightened of you) are inherently stressful or full of stressors. We should then also agree that it simply is not possible to rid all workplaces of all stressors. Why then do workplaces function at all? In fact, we know that some workplaces function extremely well.

7.3 Why Workplaces Function in the Face of Stress

7.3.1 Peoples' reactions to stressors differ

Not everyone reacts the same way to a stressor—the same stressful situation or event does not elicit the same strain reaction in everyone. People with certain personality types or with access to certain resources appear to be protected from the effects of stressors, they experience less strain than others as a result of the exposure. For example, people who have access to a strong social support network may be buffered from the effects of certain stressors. Coworkers may be able to help them with a job task while friends or family may provide emotional support in times of difficulty. These people will experience less strain than those with weak social support networks.

Similarly, people with certain personality types or without access to certain resources appear to be more vulnerable than others to some stressors, they experience more strain than others as a result of the exposure. For example, people who are rated highly on the personality trait *negative affectivity*—the general inclination to view the world negatively—are more vulnerable to stressors and they experience greater strain. People high in negative affectivity simply tend to react negatively to situations and experience negative emotions.

Therefore, one reason why workplaces are functional in the face of stress is the fact that not everyone reacts the same way to the same situation or event. Not everyone experiences the same level of strain in the face of the same event or situation.

7.3.2 People appraise situations and events differently

Not everyone views the same situation or event the same way. Lazarus and Folkman (1984) describe a process whereby people evaluate the situations and events that they face. The *Transactional View of Stress* includes two stages of cognitive appraisal: primary and secondary appraisal.

Primary appraisal is the first evaluation of a situation or event at work. When faced with a potential stressor at work we think about the possible costs or benefits to us of this situation or event. Is it irrelevant or benign *or* does it have the potential to be a threat (e.g., cause us harm or loss) or provide a challenge? It is only those situations that we consider to be potentially threatening or challenging that undergo further appraisal. Please note the use of the words threat and challenge. Not all potential stressors are considered negatively. For example, for most people a promotion at work would be considered a challenge and be associated with positive emotions.

Once a situation has been defined as potentially threatening or challenging we consider what we can do to cope with the situation. *Secondary appraisal* involves asking ourselves if we have the resources or abilities to cope with the situation or event. If we believe we do not have the ability to cope with the situation or event we perceive it as stressful. Until this point we would refer to the situation or event as a potential stressor, now it will be considered as stressor. *Stressors* then are those situations or events in the workplace that we perceive to be a threat or challenge *and* that we believe we do not have the ability to deal or cope with.

Therefore, another reason workplaces function in the face of stress is likely due to the fact that not everyone appraises the same situation or event the same way. Once again personality traits can play a role. For example, those high in negative affectivity may be more likely to appraise situations or events negatively (i.e., during primary appraisal they may be more likely to see events or situations as threatening). Alternatively, optimistic individuals may be more inclined to view events or situations in a positive way and see a challenge.

7.3.3 People cope with stressors differently

People can engage in a variety of coping strategies. Researchers have devoted much attention to two strategies proposed by Lazarus and Folkman (1984) and a third strategy proposed by Moos and Billings (1982).

Lazarus and Folkman proposed problem-focused coping and emotion-focused coping. People engage in *problem-focused coping* when they attempt to deal directly with the stressor, usually with the hopes of removing it. For example, if you must complete an assignment that requires the use of a computer program that you do not understand, problem-focused coping could involve taking a training session on how to use the software. People engage in *emotion-focused coping* when they attempt to deal with the emotions that accompany the stressor, rather than the stressor itself. For example, turning to your coworkers and discussing your concerns about your lack of experience with the program would be an example of emotion-focused coping.

Moos and Billings (1982) introduced appraisal-focused coping. People engage in *appraisal-focused coping*
indexcoping strategies!appraisal-focused when they attempt to deal with a stressor by denying that it exists or by reinterpreting it. Deciding to avoid the office and go to the movies instead of taking that training session is an example of appraisal-focused coping, as is rationalizing not taking the training session because you are convinced that you will never be able to use the software. It is easy to imagine that personality traits are useful in explaining why some people choose one coping strategy over another.

Therefore, a third reason why workplaces function in the face of stress is that not everyone chooses the same strategy to cope with the same stressor. However, it is important to be aware of one danger in interpretation of these coping strategies. No one coping strategy is necessarily superior to another. At first glance it may seem obvious that one should always attempt a problem-focused coping strategy. Why not remove the stressor entirely? But what if the worker has no control over the situation? If one has chosen an occupation where a stressor is inherent (and to remove it would mean no longer being in this occupation) an emotion-focused coping strategy may not be a bad alternative. Think about the coping strategies available to firefighters, police officers, and dentists.

7.3.4 Organizations deal with the negative effects of stress in different ways

Not all organizations deal with stressors in their workplaces in the same way. A detailed description of organizational strategies to combat the negative effects of stress

in workplaces is beyond the scope of this discussion. However, it is unacceptable to imply that the sole reasons why organizations function in the face of stress are due to individual differences in workers. Organizational leaders also have a role to play.

Organizations combat the negative effects of stress in their workplaces in a number of ways (Murphy, 1992). *Tertiary preventive interventions*are efforts to assist those workers who have been unable to cope with stressors in their work environment. Employee assistance programs such as programs to fight drug and alcohol abuse are examples of tertiary preventive strategies. *Secondary preventive interventions* are efforts to assist workers who currently are facing stressors in the workplace. For example, providing time management training for managers is a secondary preventive strategy. Such training may provide managers with a few more coping options. Although helpful, both tertiary and secondary preventive strategies tend to put the onus back on workers to deal with the stressors in their work environment. *Primary prevention*, the final intervention to be discussed, puts the onus back on the organization. Organizations use primary prevention strategies when they attempt to reduce or entirely remove stressors in their workplaces. For example, giving workers greater control over their jobs, or reducing workloads would be examples of primary prevention strategies.

Therefore, a final reason why workplaces function in the face of stress (and why some workplaces function better than others) may be the fact that organizations deal with the sources and the effects of stress in their workplaces differently. Not surprisingly, given the expense involved in many primary prevention efforts, organizations have historically made more use of tertiary and secondary preventive strategies than primary prevention strategies.

7.4 Stressors in the Workplace—A Researchers' Framework

Sauter, Murphy, and Hurrell's (1990) framework for describing potential stressors in the workplace is frequently used in the literature. In it they describe several critical stressors in the workplace. Please note that they do not include the physical work environment in their framework. However, the physical environment can be a stressor in the workplace. Think about the noise or heat levels that workers are exposed to in a manufacturing setting and how these levels may affect workers.

Sauter et al. (1990) consider:

Work Load and Work Pace: Work load consists of role overload, having too much work or having too difficult a job to complete; and role underload, having too little work or having too easy a job to complete. Both overload and underload as well as machine-paced work have been linked to health outcomes.

Work Schedule: Strain outcomes have been linked to the temporal scheduling of work (e.g., night shift, rotating shifts, 12 hour shifts).

Role Stressors: They are derived from Role Stress Theory (Kahn, Wolfe, Quinn, Snoek, & Rosenthal, 1964). They include role ambiguity—a lack of understanding of

what is expected of you, and role conflict—having incompatible work demands. Both are associated with strain.

Career Security: Job insecurity, or insecurity about the long-term job prospects have been linked to strain.

Interpersonal Relationships: The social context at work should be considered as a potential stressor. Social relationships can be positive—people receive social support, and negative—there can be conflict within and between work groups.

Job Content: The nature of the job or task should not be forgotten. It includes the skills and abilities that are required as well as the degree to which one is responsible for other people.

Control: Although not listed explicitly in their framework of workplace stressors, Sauter et al. (1990) discuss control over one's work in their list of factors that determine the impact of work on worker health. Greater levels of control over one's work are associated with less strain.

This list of potential workplace stressors is by no means complete, but these stressors could be considered core stressors in any workplace. For researchers studying workplace health the changes and trends in today's workplace are driving the need to expand such lists (Hepburn, Loughlin, & Barling, 1997). For example, the face of the workforce is changing and as more and more women enter the workforce issues related to work and family balance are emerging. Globalization has led to downsizing and mergers and the call for work teams. Technological advances are encouraging debate about "deskilling" of the workforce and the need for "reskilling". The effect of these changes on workers and their organizations must be considered. Trends show that more and more people are working part time yet much of the research in the field of work stress and health has focused on full-time employees. What does it mean to be part of this "contingent" workforce (e.g., part-time and contract employees)? These and other questions are "challenging" researchers interested in how people function in the workplace.

Chapter 8

Sleep: What's the point?
Darren K. Hannesson

When I was dating my wife-to-be, we once had an argument. Well, actually we had lots of arguments, but once we argued about what I was doing with my time. I had been playing a lot of hockey with my friends and, to be honest, spending less time with her than I usually had. She said that if she was important to me, more important than hockey, that should be reflected in the amount of time I spent with her. I conceded the point, "upped" my "wife-to-be" time by decreasing my hockey time ("sniff") and, two cats, two kids, and a minivan later, the rest is history, as they say.

I share this story with you not to lament my hardship, but to make the point that the relative importance of activities in our lives is often reflected in how much time we spend doing them. By this criterion, there is nothing we do that is more important than sleep. In fact, next time you are asked what you are doing with your life, I suggest the most honest reply would be: "sleeping, mostly". On average, humans spend about 1/3 of their lives sleeping—between 7–8.5 hours a day, 50–60 hours a week, 2500–3000 hours a year, or between 200,000–300,000 hours a lifetime.

Given the importance of sleep, you would probably expect that we have a very good understanding of what sleep is and why we do it. However, you'd be mistaken. Surprising to me at least, sleep is proving to be a great mystery on all fronts, although we have adopted some good operational criteria to define what it is. In this chapter, we will briefly explore what psychologists know about sleep, how it is defined, and why it occurs.

8.1 What Is Sleep?

Everyone feels pretty sure he or she knows what sleep is. Indeed, the difference between being full awake and fully asleep is like the difference between "day and night". But, if you are like me, upon being roused while watching a boring movie, you might

have asked more than once, “Was I asleep?”. Or, alternatively, you may have shaken a friend’s shoulder to get her to open her eyes, and then interrogated her with suspicion when she’s claimed she was really asleep (a duplicitous ploy my wife uses sometimes to make me answer the phone instead of her!). Such cases illustrate that sleep, and particularly the transition from waking to sleeping, is much harder to define with certainty than one thinks. In general terms, sleep is defined by a decrease in arousal associated with reduced responsiveness to sensory input and reduced motor output. But where is the line between resting quietly with your eyes closed and being asleep? What’s the difference between being in a coma and being asleep? Sleep researchers and clinicians have thus turned to operational criteria to define sleep precisely.

8.1.1 Sleep stages

The the main tools used to define sleep are the *electroencephalogram* (EEG), *electromyogram* (EMG) and *electrooculogram* (EOG). Each of these tools measures electrical activity and converts it into movements of a line up or down (see Figure 8.1), such that the record produced looks like a squiggly line. An EEG is recorded by placing electrodes on the scalp of a person (or other animal), which then detect the electrical activity of the cerebral cortex underneath the skull. The signal obtained reflects the summed activity of large populations of cortical neurons in the brain regions below the electrodes, and is analogous to the crowd noise that would be recorded by a microphone hanging from the bottom of a blimp over a football stadium. Sleep is reflected in characteristic patterns of brain activity with distinct appearances on the EEG record. EEG, therefore, is the primary tool used to define sleep. An EMG is recorded by placing electrodes on the skin overlying major muscle groups, which detect the underlying electrical activity associated with muscle contractions. Although EMG could be used to monitor a range of different muscle groups, in sleep research EMG usually is targeted at muscles in the body core (as opposed to the periphery), with neck muscles being the standard choice. An EOG is recorded by placing electrodes over the eye muscles, and thus detects eye movements. EOG is the primary criterion for defining the rapid-eye-movement (REM) stage of sleep, discussed subsequently.

Using these tools, researchers have discovered that sleep is not a simple continuous state, but rather a dynamic one associated with a number of distinct orderly phases, each characterized by particular patterns of EEG, EMG, and EOG recordings (see Table 8.1) (Hobson, 1989; Kavanau, 1994; Rechtschaffen, 1998). In fact, sleep can be differentiated into 4 distinct stages, termed stages I through IV, which are all distinct from an awake state.

When a person is awake, but quiet (e.g., reading), the EEG recorded consists of low amplitude (i.e., small) irregular waves reflecting high, but generally asynchronous brain activity (see Figure 8.2). Intermittent muscle activity and eye movements are reflected in high amplitude but irregular activity in the EMG and EOG respectively. Compared to waking, stages I-IV of sleep, in general, are associated with an increase in regularity and amplitude, a decrease in frequency of EEG activitym and a decrease in EMG and EOG activity. Thus, stage I, the lightest sleep stage, is associated with reductions in EEG amplitude and frequency and reduced EMG and EOG activity (although some slow eye movements are evident in the EOG) relative to waking. The frequency of EEG

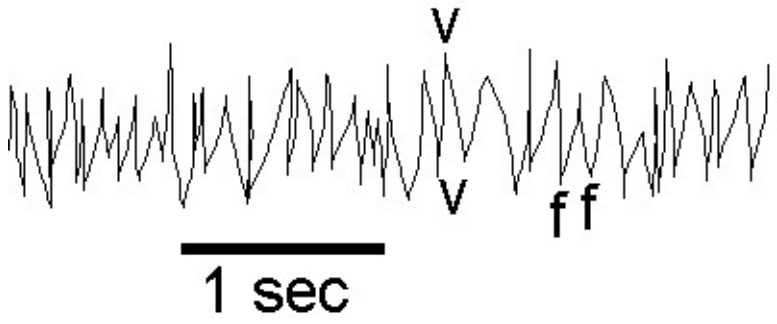


Figure 8.1: Illustration of some general characteristics of an electrogram. Up and down deflections of a pen are created by changes in the electrical signal recorded by one or more electrodes. The paper moves at a constant rate creating a squiggly line that reflects changes in electrical activity over time. Two measures used to characterize the resulting trace are spike amplitude (i.e., the height of the pen deflections, measured vertically from v to v) and the frequency (i.e., the number of cycles, defined as one up deflection followed by one down deflection, per unit time, measured horizontally from f to f).

waves is between 3 and 7 waves per second (i.e., 3–7 Hz), a pattern called theta waves. Stage II is associated with an increase in the regularity or synchronization of EEG activity, and the appearance of two intermittent wave patterns, termed sleep spindles and K-complexes. Sleep spindles are 1–3 second bursts of waves at a frequency of about 10 Hz that recur about every 10 seconds during stage II of sleep, although they are occasionally seen during other sleep stages as well. K-complexes refer to a single large (e.g., high amplitude) wave and are seen exclusively in stage II. Stage III and stage IV sleep are closely related to one another, and are collectively referred to as slow-wave sleep or deep sleep. In stage III, a low frequency (around 1 Hz) high amplitude type of activity called delta waves appears, and both EOG and EMG activity are very low. Stage IV is differentiated from stage III simply by the percentage of time spent in delta waves—with 50% being the threshold for transition from stage III to IV (and in fact much of stage IV is spent in near continuous delta wave activity).

8.1.2 The sleep cycle

There is a regular progression through the stages of sleep that is called the sleep cycle (Carskadon & Dement, 1994; Hobson, 1989). Prior to falling asleep, alpha waves, a higher amplitude lower frequency wave form, begin to appear in the awake EEG. These waves reflect a transition from being fully alert to drowsy, and seem to be required before a person can fall asleep. Alpha waves are absent from the EEG in a person that is fully alert and engaged in activity (and who, of course, would not fall asleep). When a person falls asleep, they enter stage 1 and move stepwise down through each stage

Table 8.1: EEG, EMG, and EOG characteristics during different stages of sleep.

Stage	EEG	EMG	EOG
I	Mostly moderate amplitude theta waves	Moderate activity	Low activity
II	Moderate amplitude theta waves with K complexes and sleep spindles	Moderate activity	No activity
III	Large amplitude with <50% delta waves	Low activity	No activity
IV	Large amplitude with >50% delta waves	Low activity	No activity
REM	Low amplitude irregular waves	No activity	High activity

and then back up (i.e., I → II → III → IV → III → II → I). A sleep cycle is concluded upon returning to stage 1 and is typically about 90 minutes in duration.

A typical night's sleep consists of 5 to 6 cycles, with successive cycles usually containing a larger percentage of time spent in stage I sleep (see Figure 8.3). In fact, in many cases no slow wave sleep occurs during the last 1 or 2 sleep cycles of the night. Interestingly, some evidence suggests that people feel most refreshed when they are awakened at the conclusion of a sleep cycle (Hobson, 1989). This possibility suggests you might be better off, in some cases, setting your alarm clock a bit earlier to wake yourself at the likely end of a sleep cycle (e.g., at six hours from the time you went to bed) rather than getting a bit more sleep and waking up mid-cycle (e.g., at 6.5 hours from the time you went to bed).

An additional important component of sleep is called rapid-eye-movement or REM sleep (Carskadon & Dement, 1994; W. Dement & Kleitman, 1957; Hobson, 1989). REM sleep is unique in that EEG activity returns to lower amplitude irregular activity resembling quiet waking, and there is an increase in EOG activity reflecting the rapid eye movements for which this stage is named. REM sleep occurs during stage 1 sleep at the end of sleep cycles, and thus is not typically seen during the first bout of stage 1 sleep, unless a person is sleep deprived. REM sleep is the phase of sleep most associated with dreaming, with people generally reporting dreaming about 80% of the time when awakened during this stage. The content of the dreams is typical of what most of us think of as dreams, and can often be quite elaborate.

My favorite REM dream I've had consisted of an episode in which I had been called up to play for the Philadelphia Flyers, and was on a shift standing in front of the net. The puck popped loose, but when I went to shoot I realized I had a fork instead of a hockey stick. Eventually, I managed to shovel the puck in anyway and the announcer went wild, screaming: "He shoots, he forks!". REM dreams often feel very vivid but none of the actions in the dreams get acted-out because, aside from the eye muscles, there is a near total loss of muscle tone during REM sleep, as evident in a nearly flat EMG record. This muscular inhibition is actively maintained by particular parts of the brain, and probably prevents injuries that might otherwise occur from acting out particularly physical dreams. In studies with cats, researchers have been able to block

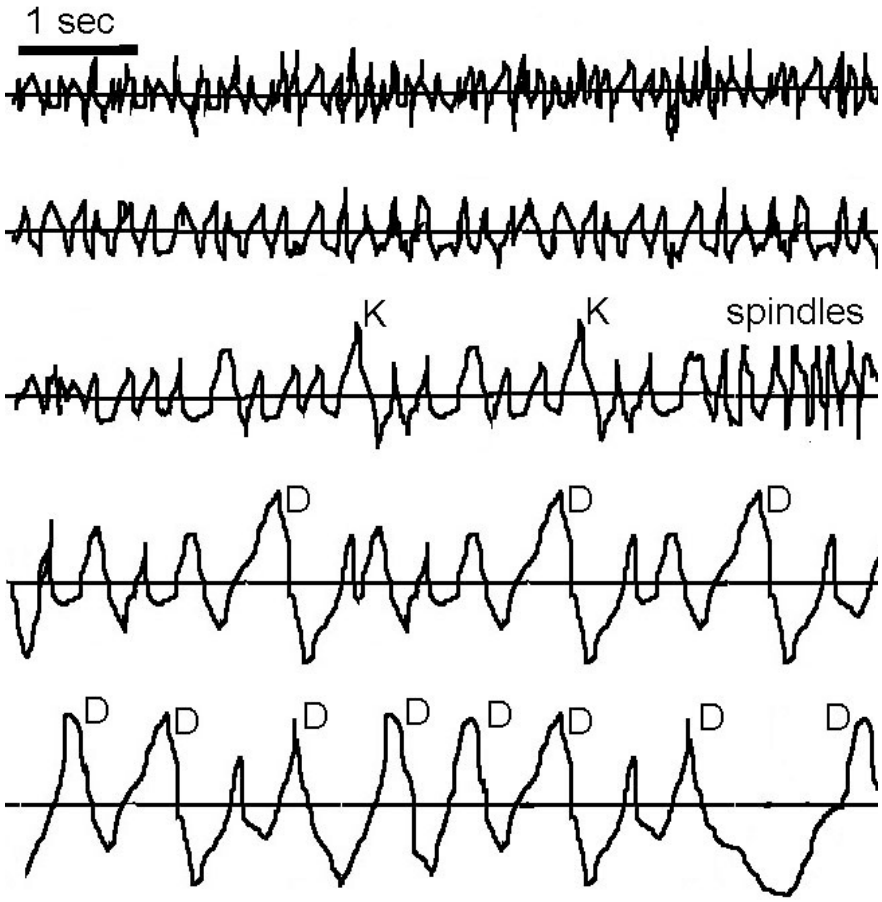


Figure 8.2: An EEG record illustrating quiet waking (the upper trace) and stages I through IV of sleep (each successive trace). K indicates a K-complex during stage II, spindles denotes sleep spindles also present during stage II, and D marks delta waves, which are present during stage III and IV. Note that REM sleep is not shown but produces an EEG trace largely similar to that of quiet waking.

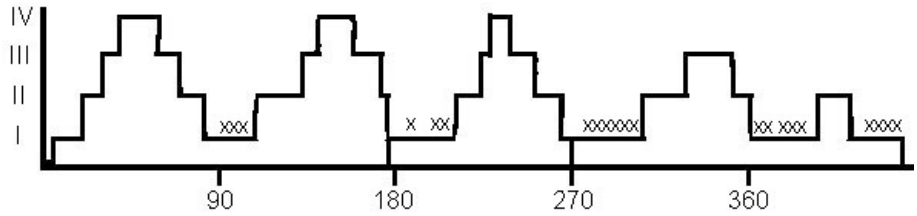


Figure 8.3: Time spent in different sleep stages during a typical human 7.5 hour night of sleep. X denotes periods of REM sleep. The vertical axis represents stage of sleep and the horizontal axis represents time in minutes from “bedtime”.

muscular inhibition during REM sleep and found that cats frequently engage in mouse hunting type behaviours, leading to speculation that mouse hunting is a likely topic of cat dreams (Morrison, Sanford, Ball, Mann, & Ross, 1995). Dreams are sometimes seen in other sleep stages (collectively termed non-REM sleep), but only about 7% of the time. Also, dreams during non-REM sleep are typically simpler, and usually mainly sensory in nature. A classic non-REM dream would simply be the sensation that one was falling.

8.2 Why Do We Sleep?

8.2.1 Theories of why we sleep

In terms of the time we commit to it, sleep might be the most important thing we do, and therefore it follows that sleep should make some easy-to-discover and substantial contribution to our overall functioning or well-being. In fact, the purpose of sleep is presently unknown, and there are a variety of opposing and highly contentious views. This ignorance has also been the case historically, and many interesting and speculative accounts have been proposed. One widely held primitive view was that sleep involved the departure of the spirit from the body, allowing it to cavort temporarily in the spirit world (Hobson, 1989). Because the spirit animated the body, this account explained both why the body seems almost dead during sleep and why we dream—dreams are a fuzzy recollection of the spirit’s adventures away from the body.

Although there are a variety of contemporary accounts for sleep, most can be grouped into two general categories: recuperative theories, and evolutionary/circadian theories. Recuperative theories hold that sleep repairs or otherwise helps maintain the body (Carskadon & Dement, 1994; Rechtschaffen, 1998; Reimund, 1994). According to these views, being awake disturbs the body, causing either damage and/or a departure from homeostasis (i.e., physiological stability of the body); being asleep repairs this damage and/or restores homeostasis. Furthermore, according to the strictest interpretation of these views, sleep should occur (or at least become “desired”) when the body reaches some threshold level of disruption, and should last until repair/homeostasis is re-achieved.

According to evolutionary/circadian theories sleep is part of a 24-hour cyclicity to

our behaviour determined by pressures of natural selection (Cohen, 1979; Kalat, 2004; Kleitman, 1963; Rechtschaffen, 1998). We, and most other species, have evolved in an environment with two distinct phases: day and night. Moreover, all species have faced the need to achieve survival and reproductive goals under conditions in which there is varying degrees of competition for adequate supplies of food and nutrients. The evolutionary/circadian view holds that sleep, then, is an adaptation produced by these selective pressures. There are likely differences in how well any given species could function during one set of conditions (e.g., day) compared with the other (e.g., night). The evolutionary/circadian view holds that it would have been adaptive to maximize activities during the phase for which a species is optimally suited, and to minimize activity during the other phase. In addition, it would have been adaptive for a species to be active each day just long enough to achieve evolutionary significant goals, such as food and mate procurement, but remain inactive otherwise, via sleep, to ensure that potentially scarce energy/nutrient reserves are not wasted. Thus, the evolutionary/circadian view predicts that a species should sleep during that part of the day-night cycle for which it is least well suited, as a means of protection from accidents and predators. Furthermore, the duration of sleep should reflect the approximate daily time required to achieve significant survival and reproductive goals without wasting energy on unimportant activity.

Evaluating recuperation versus evolutionary theories, however, is difficult. Damage/homeostatic imbalances produced by waking, and which require sleep for repair, are proving difficult to identify. For example, muscle fatigue, although an obvious example of waking-related "damage", is repaired equally well during quiet rest as it is during sleep (Rechtschaffen, 1998). At the same time, to some degree, evolutionary-based theories require speculation about ancestral conditions and associated pressures that are challenging to verify. Nonetheless, there are several important sources of data that enable evaluation of some of the predictions of these two types of sleep theories: the comparative analysis of sleep, the effects of sleep deprivation, and general circadian effects.

8.2.2 Evaluating the predictions of theories for sleep

Comparative analysis of sleep

The comparative analysis of sleep provides a number of interesting observations relevant to hypotheses on why we sleep. First, all mammal and bird species studied so far engage in sleep (Kalat, 2004; Rechtschaffen, 1998; Winson, 1993). Moreover, sleep occurs despite the fact that it may be accompanied by significant risk, as it is for prey species that are consequently less vigilant for potential predators while asleep, or require elaborate adaptations, and as it does for some marine mammals, such as dolphins, that must be able to surface regularly for air, and therefore sleep with only half the brain at a time (Mukhametov, Supin, & Polyakova, 1977). This incredible pervasiveness of sleep makes two significant points:

1. sleep must be important for something, and

2. the primary purpose of sleep cannot relate to some special, higher-order human

function (e.g., resolving emotional conflict).

A second important observation derived from the comparative analysis of sleep is that there is incredible variability in the normal duration of sleep between species (Kalat, 2004). Sloths earn their moniker through a daily average of about 20 hours, whereas humans and pigs are in the middle, at about 8 hours. At the low end, cows get about 3 hours, and horses and deer get about 2 hours. This variability demonstrates that large amounts of sleep are not necessarily required by general mammalian physiology. It also shows that sleep duration doesn't necessarily relate to daily activity levels. Sloths aren't exactly burning up the jungle in those 4 daily hours of waking! Moreover, the case of sloths stands in striking contrast with that of horses, which can easily travel upwards of 60 miles in a day, and yet require only about 2 hours of sleep?!?

Indeed, the absence of any positive correlation between daily activity and sleep duration across species clearly refutes any simply-formulated recuperation theory that argues that sleep is caused by the accumulation of waking-related damage or disturbances in the body, which must be reversed by a corresponding duration of sleep-related repair. In contrast, an evolutionary/circadian account is far more compatible with the data on the duration of sleep across species. Many of the species that sleep very little are grazers, such as zebras, which are at greatest risk to being preyed-upon while sleeping, and at the same time need long hours devoted to eating because of the relatively low nutrient density of grasses. Conversely, many species that sleep longer hours are big, successful carnivores, such as lions, which obtain and consume large nutrient dense meals in a relatively short proportion of time, and are at little risk of predation during sleep.

Sleep deprivation

The effects of sleep deprivation are particularly germane to assessing the predictions of theories as to why we sleep. Recuperative theories make the clear prediction that sleep deprivation should have obvious consequences on the body, with cumulative damage and homeostatic disruptions accompanying increased deprivation. They also predict that greater amounts of sleep should be required when greater amounts of damage or homeostatic disruption have occurred, and thus a positive correlation between amount of time awake and amount of time spent asleep. We all know that we feel lousy when we don't get our usual amount of sleep. This common observation would appear to be *prima facie* support for recuperation theory. But particular caution must be exercised in interpreting our own subjective experiences and anecdotal reports we hear from others about sleep deprivation. First, it is difficult to disentangle the effects of sleep deprivation from what has caused the sleep deprivation. Most of us dont deprive ourselves of sleep simply to investigate the effects of sleep deprivation. Usually sleep deprivation occurs because we are under stress to accomplish something, we have been celebrating with perhaps too much food and drink, or we are otherwise not feeling well. Thus, it is easy to misattribute feeling lousy the next day to sleep deprivation when the more significant culprit may be the factor that caused the sleep deprivation. Second, it can be difficult to determine the extent to which we are experiencing meaningful adverse effects from sleep deprivation. For example, although we may feel sleepy, our bodies

may be otherwise functioning at normal levels.

One of the most interesting sources of information of sleep deprivation comes from the well-publicized case of a 17-year-old, young man named Randy Gardner from California (W. C. Dement, 1978). Under the supervision of doctors and psychologists and with the assistance of a number of friends, Randy stayed awake for just under 11 days. At no time did doctors find any cause for concern about Randy's health, and at the end of his stunt, he gave an entirely coherent press conference. Overall, he reported that he felt well, although he did complain of sleepiness, irritability, and boredom occasionally. Psychological assessments during his deprivation revealed relatively few effects besides some alteration of visual perception, which he described as brief bouts in which aspects of the world seemed illusory, or dream-like. The absence of more profound physiological or psychological disturbance after such prolonged deprivation is difficult to reconcile with recuperative theories.

An additional component of Randy's experience also was not supportive of recuperative theories. When Randy finally slept, despite missing out on about 88 hours of usual sleep, he awoke after only 14 hours, and thus regained only 6 of those lost 88 hours. He didnt regain any further sleep on subsequent nights either, as he returned to his normal 8 hours a night. If the main purpose of sleep is to repair the normal damage/homeostatic disruption that occurs in about 16 hours of waking and requires about 8 hours of sleep, it is difficult to imagine how 260 hours of waking-related damage/homeostatic disruptions, almost 20 times as much as normal, would require only 14 hours of sleep, or about 1.75 times as much as normal. However, Randy is but a single case and could therefore represent an anomaly.

There have been experimental studies of sleep deprivation in humans and other animals that have had somewhat conflicting findings (Horne, 1988). Generally, the human studies have investigated a wide range of physiological and psychological functions, and have found few or no disturbances. The three consistent effects appear to be minor and include increased sleepiness, increased irritability, and reductions in some aspects of attention, although there is some question as to whether the attentional impairments are a secondary result of microsleeps (colloquially known as "the headnod") (Durmer & Dinges, 2005). Microsleeps are 2-3 second episodes of stage 1 sleep that begin to appear intermittently after several days of sleep deprivation. In contrast to the human studies, dramatic results have been obtained in many studies of sleep deprivation in non-human animals (Rechtschaffen, 1998).

For example, studies in rats have sometimes resulted in death within 2 or 3 weeks of even partial deprivation. However, it is not clear whether sleep deprivation or stress is the real culprit in such studies. In contrast to human studies, animals cannot be told that they may drop out of the study at any time if they aren't enjoying the experience. They also can't be told that the deprivation will end after a fixed period of time. Finally the techniques for sleep deprivation in animals are often very stressful in themselves. For example, one technique, called the flower pot method, involves placing the subject on a balanced pedestal surrounded by water. When the animal enters sleep and relaxes its postural muscles the pedestal tilts leading to a "rude awakening" as the animal slides into cool water. Indeed, there is considerable evidence of stress-related effects, such as stomach ulcers, in these animal studies that aren't seen in the deprivation studies with humans.

Circadian rhythms

According to the evolutionary /circadian account, sleep represents a programmed cyclicity in our activity that was selected by evolutionary pressures. Consideration of some of the features of circadian effects in organisms is generally consistent with such a view. The term circadian, roughly translated from the Latin root, means "about a day". It is used to describe behavioural or physiological cycles that are roughly one day in duration (Hastings, 1997). The circadian sleep/wake cycle is the most obvious example of such phenomena, but things such as body temperature and hunger also have a natural daily rhythm (Arendt, 1998). At one point, most psychologists assumed that such rhythms were led by environmental information or cues about time of day. Such cues, called zeitgebers, are numerous and include such obvious things as the intensity of outdoor light or the position of the sun, and more subtle things, such as ambient temperature or activity of other organisms (e.g., when the rooster crows it is morning).

However, a number of findings have now demonstrated that most circadian effects, including sleep, are largely controlled by a genetically programmed internal biological clock, although zeitgebers can have a modulatory influence (Mistlberger, 1994). The most poignant illustration of this endogenous periodicity comes from studies in which humans or other animals have been kept in conditions devoid of zeitgebers (e.g., a room with constant lighting conditions and no other information about time of day). In these cases, circadian variation in behaviours such as sleep is maintained, strongly arguing that such rhythms must be endogenously driven. Interestingly, however, for humans the sleep/wake cycle drifts to about 25 hours in these cases, which indicates that zeitgebers do have a modulating influence that normally keeps our cycle at 24 hours. Additional studies have shown that post-natal experience with the day-night cycle is not required for the expression of circadian rhythms in zeitgeber-free conditions, further illustrating the existence of a genetically programmed internal clock (Richter, 1971). The presence of such a clock is, of course, required by the evolutionary account for sleep, but its presence alone is not necessarily incompatible with recuperation theories. For example, circadian sleep rhythms could provide a regular impetus for recuperation, but still allow the duration of sleep to be shortened or lengthened depending upon how much activity-dependent, damage/homeostatic disruption an individual had experienced.

However, the data available suggest that circadian rhythms are relatively impervious to modulation in a manner predicted by recuperation theories. Numerous studies have attempted to shift experimentally the sleep rhythm in humans or other animals by exposing individuals to increases in physical, mental, or physiological stress; in all cases, little to no change in the circadian sleep cycle occurs (Mistlberger, 1994). In other words, a day of hard physical labour doesn't appear to alter when a person feels sleepy or how long they stay in bed, at least under experimental conditions (O'Connor & Youngstedt, 1995). And because increased exposure to stressors ought to increase the need for body repairs or homeostatic adjustment, recuperation theories' prediction of increased requirements for sleep under these conditions is simply not supported.

A related source of data comes from experiments in which people are forced to stay up later than usual. In these cases, recuperation and evolutionary theories make strikingly different predictions. Recuperation theories again predict that longer durations of waking should produce greater damage/homeostatic disruption and therefore result

in greater than normal sleep. Evolutionary/circadian theories predict that sleep should still be largely confined to its normal time of day, and, therefore, being awake during part of your normal sleep period should in fact lead to less than normal sleep. Such studies have found that individuals get less total sleep than normal, consistent with the evolutionary theories' prediction, although they do often sleep somewhat beyond their normal waking time (Wever, 1979).

A final illustration of the primacy of circadian factors versus recuperative factors in motivating sleep, which you can and probably will try at some time during your university career, comes from staying up all night. Recuperation theories predict that the longer you have been awake, the sleepier you should get as damage/homeostatic disruptions mount; evolutionary/circadian theories predict that you should actually become less sleepy as your normal waking time approaches. In fact, consistent with what night shift workers talk about as "getting over the hump", people begin to feel less sleepy as morning approaches.

8.2.3 Conclusions on sleep theories

Overall, evolutionary/circadian accounts for sleep are more consistent with some of the characteristics of sleep. And, if these theories are correct, there are some interesting implications for how we should approach sleep. First, if sleep needs don't relate to the events of the preceding or subsequent day, we need to change our thinking about deviations from our normal sleep habits. In other words, expressions like "Get some sleep, you had a big day today (or have a big day tomorrow)" should become obsolete. Second, if sleep evolved to solve ancestral needs of protection and energy optimization, little if any sleep may really be required in environments like modern Canada, where the selective pressures of ancestral times simply aren't present. Night-related danger, particularly in urban areas, is no longer substantial due to the ubiquity of artificial lighting and the virtual absence of dangerous predators. Energy optimization also is not particularly important under conditions where the food supply is plentiful. In fact, excess energy consumption is one the major problems faced in Canada. Perhaps reducing the amount time spent conserving energy through sleep, then, could be part of a solution to the obesity problem in Canada!

However, it is important to point out that, although a strict interpretation of recuperation theories, in which there is a direct relation between degree of waking related body disturbances and sleep needs, seems untenable, a role for sleep in repair or homeostatic processes cannot yet be fully dismissed (Rechtschaffen, 1998). For example, it may be that although sleep was "designed" to solve adaptive problems related to protection and energy conservation it may also provide an important period for body recuperation. The failure to detect any simple relation between degree of waking-related damage or homeostatic disruption and sleep needs, then, might be explained by the possibility that repair functions typically require much less sleep than we get (and therefore additional sleep isn't required when the amount of repairs needed goes up). Indeed, one idea that has emerged is that there are two types of sleep—*core sleep* and *optional sleep* (Horne, 1988). Core sleep may consist of only a few hours of sleep each night, and includes the time required for body repairs and homeostatic adjustments. Optional sleep, then, is any extra sleep, which solely serves the purposes of protection and energy conservation

proposed by evolutionary/circadian accounts. This middle ground seems reasonably to coincide with common sense, and suggests that most of us might be able to increase our productivity (or time for fun) without ill-effect by at least reducing the number of hours we sleep, if not forgoing sleep altogether.

Chapter 9

Psychopharmacology
Darren K. Hannesson

How many people in Psychology 1000 are junkies? How many are regular illicit drug users? How many frequently take mind-altering drugs of some kind? How many have used a drug that changes brain activity in the last week? The last month? The last year? Although the exact type and pattern of use varies, chances are most of you have personal experience with recreational drugs in some form. Maybe you have had a caffeine fix with a cup of java or a chocolate bar and a coke. Maybe you have been stupid and had a couple of smokes, or put on a funny hat and had a few drinks at a sporting event. Maybe you got a bit tipsy once, and drank a shooter from somebody's cleavage, or maybe you even had a couple of tokes listening to some cool music.

No big deal perhaps. But maybe you went a bit further, and experimented once, and did a line of powder or even mainlined some smack, horse, big H, doogie, white lady, junk, schoolboy (all street names for heroin). That may be a little hard core perhaps, but chances are that some of you, at least, have gone too far with drugs and suffered some adverse effects. Maybe you have worshiped at the porcelain altar, or were in an accident of some sort. Maybe you have shared some drugs with a minor, or consumed so many pills or drinks that you have found yourself passed out on the floor, or zonked-out, half-naked on a picnic table. Maybe you have even had a bad stretch of intense drug use that caused your hygiene or health to deteriorate, or even an acute medical crisis.

Drug use is extremely prevalent in our society (and around the globe) and heavily permeates pop culture. Although data vary somewhat from study to study and across different parts of the world, representative numbers for the percentages of teens and adults (13 and up) taking various kinds of drugs can be seen in Table 9.1 (Maisto et al., 2004). Approximately 40% of North American adults are likely to have tried illicit drugs, about half of which can be accounted for by experimenting with marijuana. Legal drugs such as nicotine (tobacco) and alcohol have dramatically higher lifetime

Table 9.1: Prevalence of psychoactive drug use in North America (adapted from Maisto et al., 2004). Data are expressed as a percentage of adults over age 18 that have used a particular category of drugs in the past month, year, or at all. Numbers in parentheses are data for adults in the 18–25 year old age bracket only.

Drug	Past Month	Past Year	Lifetime
Illicit drugs	7 (20)	12 (32)	40
Illicit drugs (not marijuana)	3	6	24
Psychotherapeutic drugs (misuse)	2	5	15
Nicotine	25 (38)	30 (47)	45
Alcohol	50 (60)	65 (75)	90
Caffeine	80	85	98

use rates of about 50% and 90%, respectively. The ubiquitous drug, caffeine, which can be found in coffee, tea, chocolate, and many kinds of pop, is regularly consumed by an astounding 80% of adults and is likely to have been tried by as many as 99%. In fact, caffeine is even routinely given to children (in the form of chocolate and pop). When consideration is restricted to the age range of typical university students, the data are even more striking. Illicit drug use in the last month is about 20% (vs. 7% for adults in general), cigarette use is about 38% (vs. 35% for adults in general), and alcohol use is about 60% (vs. 50% for adults in general).

So, a lot of people are using drugs. So what? A lot of people are eating pizza. A lot of people are going to movies. Why, as a society, do we care so much about drug use? The reason is simple. The drugs just mentioned alter the way that people think, act, and feel, and thus can have a profound effect on how individuals function. Because a large number of individuals engage in drug use, this behaviour can have a cumulatively large effect on society. But that is not all. Because drugs can be powerfully reinforcing, they can produces patterns of use and behaviour that can be very difficult to change. This resistance to change, then, limits the extent to which drug use can be curtailed even in the face of adverse individual or societal consequences.

Although a numerical value probably is not needed for any of us to appreciate the significance of the potential adverse effects of drug use, some quantification helps put the magnitude of the problem into perspective. However, quantifying the exact effect of drug use is incredibly complex. Many of the societal problems whose incidences are dramatically increased in association with drug use are difficult to evaluate. What are the costs of murders, assaults, rapes, accidents, and relationship problems that arise, in part, from drug use? Nevertheless, to get some sense of the costs of drug use, researchers have tried to quantify part of it by focusing on measurable things such as insurance payouts for accidents, lost productivity from absenteeism at work, treating drug-related illnesses, and crime. In the U.S. alone, the figure obtained is about $400 billion (Shreiner Institute for Health Policy, 2001), or the equivalent of the entire annual revenues of the Government of Canada!

Given the prevalence and effect of drug use, it is worthwhile to know something

about these kinds of drugs, what effects they produce, and how they produce them. The objective of this chapter, then, is to introduce briefly some concepts in the broad and fascinating field of psychopharmacology (the study of drugs and their effect on behaviour) and explore a few commonly used drugs for illustrative purposes.

9.1 The Bases for Drug Actions

There are, of course, many different kinds of drugs. But, you may have noticed that the types of drugs I have been talking about share certain similarities, and are different from things like cumidin to thin the blood or viagra to treat penile erectile dysfunction. So, we are not talking about just any drugs, but specifically those that affect our thoughts, feelings, and actions. These are called psychoactive drugs. But pizza can also affect my thoughts, feelings and actions. What is the difference between a psychoactive drug and pizza? Although there are grey areas, drugs are generally differentiated from other substances by the fact that they are not normally required for the maintenance of health, in contrast to substances such as food, vitamins, or water.

The field of psychopharmacology is broad, with many drugs of interest, and many questions to ask. As an introduction, this chapter can cover only a small slice. We will look at a few drugs used primarily for their recreational value and ask "what effects does this drug produce" and "what mechanisms mediate the drug's effects". Before we can do that, however, we need to look briefly at the bases for normal thoughts, feelings, and behaviour.

9.1.1 The neural bases of experience brain, neuron, and synapse

A central tenet of neuroscience and psychology is that every feeling, thought, and, in fact, all psychological phenomena are based on the activity of the nervous system, and primarily the brain. Actions or behaviours, then, are the product of interaction between the brain and the body. If we use a PET scanner to record the activity of the brain of an awake person, we can see that at any moment there are areas of the brain that are more active than others. When we ask the person to do something (e.g., think of the word "monkey"), we can see changes in the areas of brain that become active (Demonet, Thierry, & Cardebat, 2005). Furthermore, when we ask the person to do that same thing again, we see many of the same changes in brain activity. This consistency leads us to believe that experience (i.e., thinking of the word "monkey") is mediated by patterns of activity across the brain.

But what exactly is brain activity? Brain activity consists of the summed activity of billions of cells, called neurons. These billions of neurons are interconnected with one another via thin fibres, which enable them to send and receive signals. The primary function of a neuron, and how the brain and nervous system ultimately work, can be thought of as communication—a network of billions of neurons "talking" with one another by sending, receiving, and processing signals.

The key to understanding how most psychoactive drugs work is understanding the mechanisms of neuronal communication (and how drugs affect it). Generally, there are two steps. First, because neurons are of some length, they need a way to send signals

along that length. This signalling is achieved through unique, electrochemical properties of neurons that enable them to conduct electrical signals in a way that resembles the conduction of electricity along a wire. Second, once the signal gets to the end of a neuron, it has to be transmitted to the next neuron. This transmission occurs at specialized junctions, called synapses, through a process called synaptic transmission.

Synaptic transmission is the most complex and flexible component of neuronal communication, and is the primary process affected by most psychoactive drugs. A synapse is made up of the terminal part of the fibre from one neuron sending a signal (which we call the pre-synaptic neuron), part of the neuron that will receive the signal (which we call the post-synaptic neuron), and the space in between. When the electrochemically-conducted signal reaches the terminal, it causes the release of special chemicals that are stored in packets there into the synaptic space. These chemicals are called neurotransmitters because they transmit signals between neurons. Once in the synaptic space, neurotransmitters interact with specialized proteins found on the surface of the post-synaptic neuron, called neurotransmitter receptors, and produce changes in the activity of the post-synaptic neuron.

There are a few important points to note about synaptic transmission. First, there are many different kinds of neurotransmitter molecules (50 or more), although any given neuron generally produces and releases only a few of these. Thus, there are differences in the types of neurotransmitters seen in different parts of the brain and differences in the extent to which different neurotransmitters are involved in different brain functions. Second, for each neurotransmitter, there are specific receptors that selectively interact with that neurotransmitter, and not other molecules normally found in the brain. Furthermore, for most neurotransmitters, there are several different subtypes of receptor with which it can interact. Third, for the most part, the net effect of the signals transmitted at synapses ultimately comes down to producing either an increase or decrease in the activity of the post-synaptic neuron. Thus, we call synaptic transmission that increases the activity of the post-synaptic neuron excitation (or excitatory transmission). Transmission that decreases the activity of the post-synaptic neuron, we call inhibition (or inhibitory transmission). Different neurotransmitters tend to be involved in producing excitation or inhibition, although some neurotransmitters can produce either, depending upon which of its receptor subtypes with which it interacts.

9.1.2 Drug effects on synaptic transmission

The main way that psychoactive drugs affect brain activity, then, is by altering synaptic transmission. In general, most drugs can be grouped into one of two categories. Agonists are drugs that mimic or enhance synaptic transmission, mediated by a particular neurotransmitter or neurotransmitter receptor. Antagonists are drugs that block or reduce synaptic transmission, mediated by a particular neurotransmitter or neurotransmitter receptor. These effects are usually achieved by either: 1) changing the amount of neurotransmitter found in the synaptic cleft, or 2) interacting directly with neurotransmitter receptors.

In quick summary, then, the way drugs affect behaviour is by affecting activity at synapses, which changes the frequency of activity of affected neurons, thus altering patterns of activity across the brain. Because patterns of brain activity mediate our

thoughts, feelings, and behaviour, drug-induced changes in brain activity alter the way we think, feel, and act. Different psychoactive drugs produce different effects because they produce agonistic or antagonistic effects on different neurotransmitters or neurotransmitter receptors.

9.2 Nicotine

indexnicotine Nicotine, derived from the tobacco plant, is one of the most widely used drugs in the world. The tobacco industry generates about $45 billion a year, with as many as 2 billion users(Goldberg, 2003). Nicotine is usually consumed by smoking, which involves inhalation into the lungs, where absorption is rapid across the blood-vessel rich linings (Hukkanen, Jacob 3rd, & Benowitz, 2005). In fact, nicotine may reach the brain within 7 seconds after inhalation, which, remarkably, is about twice as fast as for a drug injected intravenously.

The tobacco plant, two species of which are commercially cultivated, is the only known natural source of nicotine. This plant was indigenous to South America, and tobacco was likely used by the indigenous populations there for perhaps thousands of years (Maisto et al., 2004). It caught on like wild fire once discovered by 15th century Europeans; by the 1600s it was everywhere around the globe. Interestingly, despite its popularity, smoking was opposed by such prominent figures as the pope, the Russian czar, and the emperor of Japan. However, laws or edicts proved ineffective in curtailing its use. As one modern researcher has noted “no society that has ever taken up tobacco use has ever abandoned it” (Brecher, 1972).

Although smoking, in the form of cigarettes, cigars or pipes, is the most common route of administration, nicotine is also consumed in the form of snuff, a processed mixture of tobacco, salt, oils, and flavorings, that can be snorted and absorbed across the nasal membranes. It can also be consumed in the form of chewing tobacco, which is absorbed across the linings of the mouth and cheeks. Nicotine, in fact, is readily absorbed across any body tissue (Hukkanen et al., 2005), a fact that is capitalized on by several products designed to make quitting smoking easier. Nicotine gum relies largely on absorption across the linings of the mouth and cheeks (called trans-mucosal absorption), and “the patch” relies on absorption across the skin (called).

The main action of nicotine is as a mild brain stimulant (Goldberg, 2003; Wonnacott, Sidhpura, & Balfour, 2005). Thus, it increases arousal, which is thought to be the reinforcing action of the drug, and enhances certain cognitive functions, such as short-term memory. Like many stimulants, nicotine also suppresses appetite, which may be part of why people often eat more when they quit smoking and gain some weight. Vomiting or nausea can be induced by nicotine, which is why novice smokers sometimes “turn green”, but tolerance to this effect develops rapidly. The main physiological effects of nicotine include increasing heart rate and blood pressure and reducing muscle tone. The chronic stimulatory effects of smoking on the cardiovascular system help account for the smoker’s dramatically increased risk of heart disease and stroke. At high doses of about 60 mg, nicotine can even kill a person. And although a typical cigarette contains between 5 and 10 mg. nicotine, overdose by smoking is practically impossible because only about 1 mg gets absorbed per cigarette, and levels don’t build

up too much because nicotine gets eliminated from the body fairly quickly.

The mechanism of nicotine actions involves its ability to activate directly one of the subtypes of receptors for the neurotransmitter acetylcholine (Wonnacott et al., 2005). This receptor is an excitatory receptor (i.e., it stimulates activity of neurons), and is widespread throughout the brain and other parts of the body, including muscle tissue and the autonomic nervous system, which is important in regulating the activity of your internal organs. Not surprisingly, acetylcholine is implicated in all of the various functions that nicotine effects, including arousal and cognition, regulating muscle activity, and regulating activity of the cardiovascular system. One of the significant effects of nicotine in the autonomic nervous system is to promote release of adrenaline—which produces part of the rush of having a smoke.

9.3 Alcohol

Behind caffeine, alcohol is probably the most widely used psychoactive drug consumed by humans (Maisto et al., 2004). In fact, alcohol pervades our cultures in ways not seen with any other drug. It is an integral part of a wide range of human social phenomena, including weddings, sporting celebrations, christenings, religious ceremonies, parties, and musical events. Alcohol use is at least 10,000 years old; in fact, it has been suggested that the first human civilizations in Sumeria may have arisen in response to the need for a stable home base to make beer! Indeed, the prevalence of written information on brewing beer in the clay tablets dating from this time and region suggest that alcohol was a topic of some importance.

So what is alcohol? There are a variety of organic compounds referred to as alcohols, but the form we drink is “ethanol”. Alcohol is consumed in a wide variety of beverages that are derived from the fermentation of sugars by yeast. Barley malt is the source of sugar for beer, honey for mead, grapes for wine, and various other fruits for fruit wines. There is a limit, however, to how much alcohol can be produced before the yeast kill themselves off (about 5% for beer, and about 12% for wine), which thus restricts that maximum alcohol content of fermented products. To create products with a higher alcohol content, then, the process of distilling evolved, leading to the creation of spirits such as whiskey or cognac.

Because alcohol is thus made in the form of “tasty” drinks, it is almost exclusively consumed via the oral route, where it is absorbed in the intestines. It can, however, be administered effectively intravenously, as well; hospitalised alcoholics will often be put on an alcohol drip to avoid severe withdrawal symptoms.

Like all drugs, the effects of alcohol are related to the dose taken. One important variable that differentiates alcohol products, then, is the concentration of alcohol in the drink, usually expressed as percantage alcohol. Thus, if you drink 1 glass (250 ml) of beer that is 5% alcohol, you are taking 12.5 ml of alcohol (250 ml x 0.05). A 30 ml (about 1 ounce) serving of whiskey that is 40% alcohol would also give you about 12 ml of alcohol (30 ml x 0.40). Thus, a glass of beer is about the equivalent of a one ounce drink of “hard stuff”.

Once consumed, alcohol is eliminated from the body relatively slowly, and thus produces effects that can last hours after consumption has ceased (Goldberg, 2003;

Maisto et al., 2004). The enzyme, alcohol dehydrogenase, found mainly in the liver, plays a prime role in inactivating alcohol by metabolizing it to acetylaldehyde. However, a fraction of the alcohol consumed is cleared directly via the lungs, kidneys, and skin—a fact that has some important practical implications. Clearance via the lungs enables breathalizers to be used to estimate blood alcohol level by measuring the amount of alcohol being secreted from the lungs (which is usually about 1/20th the amount in the blood). Clearance through the skin also contributes to the fact that someone who has been doing some heavy drinking can be detected by the "reek of booze" even for some hours after they have stopped drinking.

The effects of alcohol on a person is a function not only of the amount of alcohol consumed, but also of how quickly it is being cleared from the body—the end result being the amount of alcohol that is in the blood, known as the blood alcohol level (BAL)(Goldberg, 2003; Maisto et al., 2004). Typically, about 8–10 ml of alcohol, or almost 1 drink, can be cleared per hour. Thus, consuming one or more drinks per hour leads to elevations in BAL that are proportional to the number of drinks taken per hour and a person's blood volume. At a BAL of 0.04% (0.04 mg / 100 ml of blood), mild intoxication usually becomes evident. At 0.08%, a frequently used threshold for legally defined intoxication, more profound intoxication becomes apparent. At 0.16%, gross intoxication is evident, at 0.28% stupor, and approaching 0.48% can be fatal.

Alcohol is a brain depressant, causing reductions in overall brain activity(Deitrich, Dunwiddie, Harris, & Erwin, 1989; Maisto et al., 2004; McIntosh & Chick, 2004). The net result depends on the dose, but in general alcohol produces varying degrees of impairment of a wide variety of psychological and physiological functions (although it is important to note that the exact behavioral effects resulting from these impairments can be significantly modulated by environmental or contextual factors, especially at lower BALs). Psychological functions that are disrupted include attention, memory, speech, and reasoning/judgment. The last effect, which typically involves disinhibition (i.e., doing things you normally would stop yourself from doing), is considered one of the most problematic or dangerous effects of alcohol intoxication, both for individuals and society. Thus, lapses in judgement may play a role in drunk driving, alcohol-related violence, and dangerous sexual behaviours. Additional impaired functions include sensory processes in all modalities and motor functions.

Physiologically, alcohol inhibits secretion of antidiuretic hormone, a hormone that normally inhibits micturition ("peeing"). This inhibition explains why drinking can promote an inordinate number of trips to the bathroom and leave you dehydrated afterwards (Wiese, Shlipak, & Browner, 2000). Alcohol also has a number of other physiological effects that can lead to weight gain (the dreaded "beer belly"). Besides typically being high in otherwise non-nutritious calories, alcoholic drinks inhibit fat metabolism and increase gastric secretions thereby promoting appetite (Suter, Schutz, & Jequier, 1992).

Although it has long been known that alcohol suppresses neuronal activity, its precise mechanism of action has long been unclear (Deitrich et al., 1989; Maisto et al., 2004). Recent advances suggest that alcohol works in at least three ways. First, alcohol produces a nonspecific disruption of the conduction of neural signals by changing the membrane properties of neurons. Second, alcohol is an agonist at a subset of receptors for the neurotransmitter GABA, which is the main inhibitory neurotransmitter

in the brain. Third, alcohol is an antagonist of the neurotransmitter glutamate, the main excitatory neurotransmitter in the brain, and decreases glutamate activity both by decreasing the amount that gets released into synapses and by blocking a subset of glutamate receptor subtypes.

9.4 Marijuana

Marijuana is by far the most commonly used illicit drug (Maisto et al., 2004). It is sometimes called cannabis because it comes from plants of the cannabis genus (mainly cannabis sativa). Cannabis plants are believed to be indigenous to India, but are quite hearty and have been found to grow readily almost all over the world (Abel, 1980). Marijuana has been used medicinally for thousands of years to treat muscle spasms, nausea, and pain; more recently, the plant was grown by American settlers for its fibres, which are strong, versatile, and called "hemp". Hemp clothing, for example, is now widely available and coveted for its durability. Recreational marijuana use, in North America at least, did not really spring up until the last century, and its spread may have been largely fueled by the prohibition on alcohol.

The main means for administering marijuana involve drying the leaves and rolling them into a marijuana cigarette, also called a "joint" or "spiff". The tops and flowers of the female plant are considered the best, and are called "ganja"; the lower leaves and stem are poorer quality, and are called "bhang". In addition to joints, it is becoming increasingly common to smoke marijuana with bong pipes. These pipes have a pocket of water between the burning marijuana and the user, which is thought to filter out some of the harsher chemicals in the smoke. Marijuana is also sometimes processed to create a higher concentration of the active ingredients. Hash is created by processing the dried leaves into a resin, thereby removing most of the fibrous part of the plant. This resin can be smoked in pipes, "hot-knifed" or simply ingested (e.g., as hash brownies). Hash oil is made by boiling hash in alcohol and getting the purest resin separated from all other parts of the plant. It is usually added to tobacco cigarettes and smoked, though it too can be ingested. Smoking produces more rapid absorption than eating, and, as with cigarettes and nicotine, the active ingredients of marijuana can reach the brain in less than 10 seconds after smoking (Iversen, 2000).

Marijuana produces a wide range of psychological effects (Goldberg, 2003; Iversen, 2000; Kalat, 2004; Maisto et al., 2004). Its popularity as a recreational drug derives primarily from its ability to produce euphoria (a state of giddy happiness) and reduced anxiety. Additional effects include analgesia, increased appetite, short-term memory impairment, reduced motor coordination, and perceptual disturbances. Much of the controversy over the dangers of marijuana use relates to the effects of long-term use and something called the amotivational syndrome. Opponents of marijuana use suggest that long-term users exhibit a broad decline in motivation that adversely affects daily functioning; research does indeed show that regular marijuana users score lower on tests of motivation. However, the effect is small and skeptics point out that individuals who become regular marijuana users may simply be less motivated to begin with.

Physiologically, marijuana increases the heart rate and dilates peripheral blood ves-

sels (which can lead to blood shot eyes, which can sometimes serve to tip-off someone as being "stoned"). Marijuana also leads to dilation of the bronchial passages in the lungs, and thus has been used to treat asthma (a disorder involving constriction in the lungs). However, the utility of marijuana in this situation is limited by the fact that marijuana smoke also damages the lungs, and thus could potentially exacerbate asthma and produce other lung problems, such as cancer, in the long run. Marijuana also has well known effects on both the male and female reproductive systems. It suppresses testosterone production and leads to decreased sperm counts and reduced libido. It can also suppress ovulation in females.

Although marijuana has a wide range of very real effects on psychological and physiological functioning, and some of these are clearly undesirable, there are many myths that have been perpetuated about it (Kalat, 2004). For example, it has been suggested that marijuana causes brain damage, infertility, insanity, and violence. At present, the available research fails to support any of these claims (Goldberg, 2003; Maisto et al., 2004).

Until relatively recently, the mechanisms of action for marijuana were not understood. We now know that marijuana contains more than 500 different chemicals, 60 of which are unique to this plant, and called cannibinoids (Maisto et al., 2004). Some of these cannibinoids, then, are responsible for marijuana's actions mainly via interactions with the brain. It was recognized fairly early on that one of the most important molecules, but definitely not the only one, was tatrahydrocannabinol (THC). But, nobody knew how THC or other components of marijuana exerted effects on brain activity until THC-specific receptors were found in widespread regions of the brain (and body). This discovery, however, only posed a new question. Why were these receptors there? It took several more years before a molecule endogenous to the brain, called anandemide, was found that also could activate these receptors (termed cannibinoid receptors) (Howlett et al., 2004). Interestingly, anandemide was shown to bear remarkable structural resemblance to THC, but its normal function in the brain is still not known.

9.5 Conclusion

Psychoactive drugs are substances that alter the way we feel, think, or act, but are not needed for normal maintenance of body health. They act by altering patterns of activity across brain regions, typically through effects on synaptic transmission. Different drugs produce a range of psychological and physiological effects, which depend in large part upon which neurotransmitter (or neurotransmitter receptor) they affect and whether they produce agonistic or antagonistic effects. For any given drug, the strength and even nature of its effects vary with dose, which in turn can be significantly affected by the route of drug administration. These points were illustrated by looking at three of the most commonly used recreational drugs—nicotine, alcohol, and marijuana.

Chapter 10

Brain Plasticity
Bryan Kolb

Donna was born on June 14, 1933. Her memory of her early life is sketchy but those who saw her early on report that she did not seem to know anything for some time. She could neither talk, nor walk, nor even use a toilet. Indeed, she did not even seem to know who her father was, although her mother seemed more familiar to her. Like all children, Donna grew quickly and in no time she was using and understanding simple language and could recognize lots of people by sight almost instantly. Donna began taking dancing lessons when she was four and was a "natural". By the time she finished high school she was ready for a career as a dancer with the National Ballet. Her career as a dancer was interrupted in 1958 when she married and had two children. Donna never lost interest in dancing and kept fit in her years at home with the kids. In 1968 her children were in school she began dancing again with a local company. To her amazement, she still could do most of the movements although she was pretty rusty on the classic dances that she had once so meticulously memorized. Nonetheless, she quickly relearned. In retrospect she should not have been so surprised as she had always been known as a person with a fabulous memory.

In 1990 Donna was struck by a drunk driver as she was out on an evening bicycle ride. Although she was wearing a helmet she suffered a closed head injury (among other injuries!) and was in a coma for several weeks. As she awakened from the coma she was confused and had difficulty in talking and in understanding others, she had very poor memory, she had spatial disorientation and often got lost, she had various motor disturbances, and she had difficulty recognizing anyone but her family and closest friends. Brain scans revealed diffuse cerebral injury with some focal injury on the ventral surface of the temporal and frontal lobes where the brain presumably was banged against the skull in her fall. Over the ensuing 10 months she regained most of her motoric abilities and language skills and her spatial abilities improved significantly. However, she found herself to be short-tempered and easily frustrated with her

slow recovery. She suffered periods of depression. Two years later she was once again dancing, but she now found it very difficult to learn new steps. Her emotions were still labile, which was a strain on her family, but her episodes of frustration and temper outbursts were becoming much less common. A year later they were gone and her life was not obviously different from that of other 55 year old women. She did have some cognitive changes that persisted, however. She could not seem to remember the names or faces of people that she met and was unable to concentrate if there were distractions such as a television or radio playing in the background. She did not seem to be able to dance as she had before her injury and she retired from her life's first love.

Donna provides a typical example of one of the most intriguing and important properties of the human brain: it has a capacity for changing its structure, and ultimately its function, throughout our lifetime. This capacity to change, which is known as brain plasticity, allows the brain to respond to environmental changes or changes within the organism itself. Consider the plasticity in Donna's brain. When she was a newborn she was confronted with a world that nature could not possibly have prepared her for. She had to learn language, to distinguish different faces, to walk, to ride a bicycle, to read, to dance, and so on. Since her brain is solely responsible for her behaviour, this means that her brain somehow had to change to reflect her experiences. When Donna reached puberty her body changed but so did her thoughts. Her dreams often had sexual content and, since her dreams are a product of her brain, there must have been some change in her brain activity to change her dreams so dramatically. This change was likely induced by the estrogen surge of adolescence. When Donna returned to dancing after a 10 year break, she had retained much of her skill, even though she had not practised at all. In this case the brain somehow did not change and she could quickly relearn what she had lost. After her accident, Donna had to "relearn" how to talk and walk and so on. In actual fact she did not go through the same process she had as a baby but something in her brain had to change in order to allow her to regain her lost abilities. Whatever changed in her brain must have had some limits, however, because she never did recover her memory or her ability to learn new dances.

Thus, in the life of Donna we can see several different types of brain plasticity. First, during her early childhood the brain changed dramatically in its structure, organization, and behaviour. These changes were not accomplished quickly: her brain was fundamentally different from its adult form until at least 12 or 14 years. Indeed, the plastic changes in the developing brain are so profound that a child is effectively a different creature at different stages of its own development! The brain's plasticity reflects more than mere maturational change, however, as it includes the ability to change with experience. Indeed, the capacity to alter brain structure and function in response to experience provides the nervous system with the ability to learn and to remember information. Some experiential changes are self evident, such as the acquisition of specific bits of knowledge, whereas other changes are more subtle, such as perceptual learning or the development of different problem solving strategies. Nonetheless, regardless of the nature of experiential change, the brain has altered its form and function. Finally, after a brain injury processes are recruited to change the brain again. In this case the brain must reorganize, at least in part, in order to allow the production of behaviours that have been lost.

Although the property of brain plasticity is most obvious during development, the

brain remains malleable throughout the life span. Indeed, it is evident that we can learn and remember information long after maturation.

Furthermore, although it is not as obvious, the adult brain retains its capacity to be influenced by 'general' experience. For example, being exposed to fine wine or Pavarotti changes one's later appreciation of wine or music, even if encountered in late adulthood. The adult brain is plastic in other ways too. For example, one of the characteristics of normal aging is that neurons die and are not replaced. This process begins in adolescence, yet most of us will not suffer any significant cognitive loss for decades because the brain compensates for the slow neuron loss by changing its structure. Similarly, although complete restitution of function is not possible, the brain has the capacity to change in response to injury in order to at least partly compensate for the damage.

The brain is plastic in another way, too. Imagine the problem of learning a completely new skill, such as juggling while perched on a unicycle. Initially one is totally inept but with practice at least some people can master the task. Thus, a new behaviour, or set of behaviours, has been acquired. From what we have just discussed, it should be obvious that the brain has changed. But what has changed it? One candidate is the behaviour itself. That is, if we repeatedly engage in a particular behaviour, the behaviour itself can alter the brain, which in turn facilitates the behaviour. The idea that activity might change the heart or muscles is seldom questioned. The possibility that behaviour could change the structure and function of the brain is seldom considered! Nevertheless, it is an important aspect of brain plasticity. Indeed, there is little doubt that even thought can change the brain. Consider the now extensive research on the variables influencing eye-witness testimony. Different people's accounts of the same events are notoriously inconsistent, in part because they are altered significantly by questions or thoughts 'planted' by others. That is, the "'memory", and by inference the brain, is altered by cognitive activity. In a general sense this is the process of perceptual learning where we learn about the world by observing and thinking about sensory experience.

The property of brain plasticity confronts us with a host of fundamental questions. First, since we assume that the brain produces behaviour, then how is it that a changing brain can produce stable behaviour? Shouldn't behaviour change if the brain is changing? Indeed, how do we remember anything if the brain is changing every time we learn something? Second, it has been assumed since the time of Broca (i.e., the mid 1800s) that at least some functions are localized in the cortex. If the brain is plastic, what does this imply for the nature of cortical organization? Third, what are the constraints on plasticity? There must be factors such as hormones or other chemicals that can directly control processes fundamental to plasticity. Fourth, it seems likely that the brain has some type of limits in the extent to which it is plastic. What are the limits and what determines them? Fifth, there is the general question of establishing what the properties of the nervous system are that enable it to be plastic. And, more specifically, are all regions of the brain equally plastic? Sixth, there is a clinical, and even potential educational, question of whether we can gain control of the plasticity and turn it on or off at opportune times. Finally, there is the interesting question of what factors influence plasticity and whether individual differences in different abilities may at least partly reflect differences in the brain's capacity for plasticity.

In sum, the study of brain plasticity provides a window on some of the fundamen-

tal questions in psychology and neuroscience. In particular, it allows a way of looking at the neurological bases of fundamental psychological processes such as learning and thought, and the manner in which these develop. It also leads to consideration of important clinical issues surrounding behavioural change, whether it be related to recovery from neurological injury or disease or psychopathology related to other causes.

10.1 Assumptions and Biases

As we consider the properties of the brain that make it plastic, we need to consider several biases and assumptions that underlay thinking about plasticity.

10.1.1 Behavioural states, including mind states, correspond to brain states

Although this idea is not novel, and probably appears to be self-evident to most neuroscientists, it has been a central philosophical issue since the time of Descartes. In fact, modern day philosophers still debate this issue seriously, in large part because mind (or 'cognitive processes' in modern jargon) is the central psychological problem. I shall assume that it is the brain that thinks and controls behaviour, and try to show that an understanding of plasticity will be enlightening with respect to how it does these tricks!

10.1.2 The structural properties of the brain are important in understanding its function

It follows from my first assertion that changes in the physical structure of the brain will be reflected in changes in its functioning. Although many behavioural scientists (e.g. Skinner) have seen the hardware of the brain as virtually incidental to the study of its function, this is not my view. Rather, I assume not only that changes in hardware underlay behavioural change but also that it is possible to identify and potentially influence those changes. This does not imply that a single structural change is responsible for all behavioural change, nor that a particular behavioural change is due to only one morphological change, nor that an understanding of morphology means an understanding of the functional properties that emerge from the morphology. It does mean, however, that the cerebral hardware places significant constraints on the computations of the brain and may provide important clues for understanding the nature of those computations.

10.1.3 Plasticity is a property of the synapse

The Spanish anatomist Ramòn y Cajal was the first to postulate in the early part of this century that the process of learning might produce prolonged morphological changes in the efficiency of the connections between neurons. However, it was not until 1948 that a Polish neuroscientist, Jerzy Konorski, formally proposed a mechanism. He suggested that appropriate combinations of sensory stimuli could produce two types of changes in neurons and their connections: (1) an invariant but transitory change in the excitability

of neurons, and (2) an enduring plastic change in neurons. In other words, Konorski suggested that when neurons are active they change. This change might be transitory, much as when one looks up a phone number and then forgets it, or it might be enduring, such as the case in which a telephone number is memorized. The idea that neurons are somehow changed with use is important for it means that one could look at the neuron and try to identify the changes. The question is, however, where do you look? A Canadian psychologist, Donald Hebb, proposed in 1949 that the logical place to look is at the synapse. He suggested that when synapses are active, they change if the conditions are right. The nature of the right conditions are beyond the current essay, but Hebb's idea that the principal change in neurons with repeated activity is at the synapse has proven to be correct. This means that during development, learning, recovery from injury, and aging, there are changes at the synapse that allow the brain to be functionally plastic.

10.1.4 Behavioural plasticity results from summation of plasticity of individual neurons

It should be self evident that no single neuron will have much influence upon functional plasticity but this needs to be stated explicitly. After all, in a brain with 10^{10} neurons (or more!) it would be inconceivable to think that any one neuron would make much difference. Note, however, that in animals with very simple nervous systems, such as tiny worms that have in order of 100 neurons, changes of a single neuron may be very important. We might predict that such "brains" would be less plastic than the human brain because it would make no sense to have it changing every time one of the neurons was especially active.

10.1.5 Specific mechanisms of plasticity are likely to underlie more than one form of behavioural change

The nervous system is likely to be conservative in its construction. Thus, general mechanisms that are used for one type of behavioural change, such as in learning and memory, may also form the basis of other types of behavioural change, such as in recovery from brain injury. This preconception does not exclude the possibility of specific mechanisms for different types of plasticity, but it has the advantage in that it allows studies of one form of plasticity to provide insights into mechanisms involved in others. Indeed, it has become clear in recent years that the structural changes underlying experientially-induced plasticity such as in perceptual development are remarkably similar to those underlying recovery from some types of brain injury.

10.1.6 The cortex is the most interesting candidate for neural plasticity

Some neural structures are likely to be more plastic than others. It is reasonable to suppose that if one were designing a brain it would make little sense to make all synapses equally plastic. After all, one presumably needs some constancy in a labile system if

function is going to remain stable. For example, one could imagine that it might be advantageous for basic spinal reflexes to be less plastic than those processes involved in lexical memory (i.e., vocabulary). One difficulty, however, is that it is often difficult a priori to predict what features of neural organization are likely to be more or less labile. There are several reasons to suppose that the mammalian cortex might be a place to search for plasticity. First, it is the cortex of humans that stands out as the largest development in brain evolution. Hence, when calculations are corrected for body size, it is the cortex that has grown disproportionately in the human brain: human cortex is three times larger than one would expect for a typical mammal.

One of the functional correlations of this increased cortical volume is an increase in behavioural flexibility, so it follows that the cortex might be the most interesting place to look for a relationship between plasticity and behaviour. Cross-species comparisons of cortical structure are enlightening in this regard. For example, if one compares the thickness of cortex across mammalian species, one of the most obvious observations is that cortical thickness correlates with perceived intelligence. Monkeys have thicker cortex than carnivores, who in turn have thicker cortex than rodents.

One striking similarity, however, is that the number of neurons in a "column" of cortex is the same across species. That is, mice and humans have the same number of neurons across a slab of cortex. This means that the difference in cortical thickness is due to neural processes (dendrites and axons), which basically means synapses. Thus, we see that increasing the processing capacity across species is associated with an increase in synapses. Similarly, within a species, there is a marked difference in cortical thickness during development, even though the number of neurons declines significantly during development. Again, it is the connectivity that accounts for the increase in thickness. In sum, the cortex is a particularly interesting candidate for studying plasticity-behaviour relationships. Of course, a disadvantage is that the cortex is complex in structure and in function. Although this may be a compelling argument to consider "simpler" neural systems, such as Aplysia, it is not a substantive argument for not studying the cortex. Indeed, the complexity has an advantage in that the cortex contains a wide variety of neuron types, transmitter types, and receptors for hormones and other growth factors. This would seem like fertile ground for study, in spite of the complexity.

10.2 Building a Better Brain

An interesting story is unfolding at a nunnery in Mankato, Minnesota. A group of about 700 elderly sisters have become an experiment in brain building. There are two unusual things about the sisters of Mankato. First, they are long-lived. Of the 150 retired nuns at Mankato, 25 are older than 90. Curiously, according to David Snowden who has been studying the nuns, it is those nuns who earn college degrees, who teach, and who are constantly challenging their minds who live longer than less-educated nuns who clean rooms or work in the kitchen. Second, the nuns have agreed to donate their brains to neuroscientists to study. The idea is that there should be a fundamental difference between the brains of the long-lived well-educated nuns and the shorter-lived, less well educated nuns. What might this difference be? Given the assumptions

that we considered above, the likely difference is in the synapse.

At the same time that Hebb was proposing a synaptic explanation for learning, he did an interesting experiment. He brought several laboratory rats home from his laboratory at McGill and released them in his home. The idea was that these animals would get a more stimulating environment in his home, especially as his wife chased them around the kitchen with a broom, and this might increase their "IQ". To test his idea, he compared the later maze learning performance of his pet house rats to their littermates who lived in a drab and unstimulating laboratory cage. The results were unequivocal. The "enriched" experience improved maze performance.

Subsequent studies by many groups have shown that the reason was that the brains changed. For example, my colleagues and I have found that living in a large rat condominium with lots of toys can increase overall brain weight by up to 10% in two months (e.g., Kolb & Gibb, 1991). What causes the increase? Again, based upon Hebb's idea, we have looked at the synapse. The simplest way to do this is to examine the dendrite (see Figure 10.1). The dendrite is basically an extension of the cell body upon which synapses can be made. Consider an arboreal metaphor. Suppose you want to increase the number of leaves on a tree in order to increase the total amount of photosynthesis. There would be three basic ways. First, one can increase the density of the leaves. Second, one can lengthen the branches and keep the density constant. Third, one could increase the amount of photosynthesis by simply increasing the size of the existing leaves. If one imagines the leaves as synapses, then there are three potential synaptic changes: increased density on the dendrites, increased dendritic length and increased synapse size. In fact, all three types of changes are found. Thus, the enriched experience has built a brain with more synapses—in effect, a better brain. Thus, when the brains of the Sisters of Mankato are studied, it is likely that there shall be a clear increase in the dendrites.

10.3 Rebuilding a Broken Brain

About 60,000 Canadians will suffer a stroke this year and an equal number will suffer a closed head injury. (That's the population of Lethbridge for each!) Both stroke and closed head injury victims have 'broken brains' and thus, broken behaviour. Until recently it was believed that because we cannot grow new brain cells after birth, there was no possible way to recover from brain damage. Any improvement that was seen after brain injury was considered to be a result of a reduction in swelling or shock. It is now becoming clear, however, that the brain has a built in mechanism that can affect at least some repair. Thus, as the brain changes in response to experiences, it also changes in response to injury. That the brain has such a capacity may not be surprising when we consider that brain injury involves the loss of neurons. Since we lose neurons constantly during our lifetime, the brain must have a mechanism in place to accommodate this loss. Imagine a region of the brain with 500 neurons. Assume each neuron has 10,000 connections. Simple arithmetic tells us that the region has 500 x 10,000 = 5,000,000 connections. Now assume that 10% of the neurons die, leaving us with 450 neurons. If the number of connections was important to maintaining function, then we ought to be able to increase the number of connections in the remaining neurons.

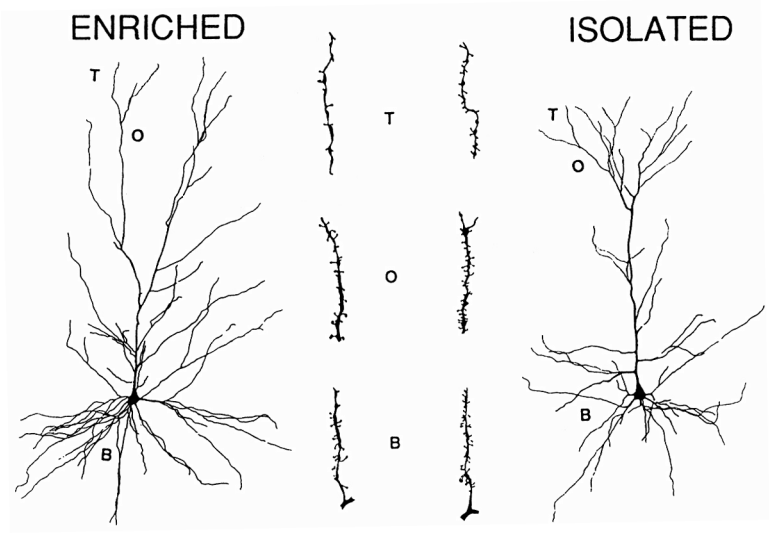


Figure 10.1: Typical neurons of rats raised in "enriched" environments or in isolated laboratory cages. Note the increased dendritic branching in the enriched cell. Short sections of dendrite in the middle panel illustrate the the density of dendritic spines in the terminal (T), oblique (O), and basilar (B) regions.

Thus, increasing the connections of each remaining neuron by about 10% leaves us with the same number of connections. This is what happens during aging. Thus, up until very old age, there is a continual increase in the number of connections on each of the remaining cortical neurons. As a result, most people suffer very little behavioural loss until well into old age. In the case of dementing diseases, such as Alzheimer, this fails to occur and there is behavioural loss much sooner. The key here is that the loss of neurons is gradual and not in a concentrated location. In a sense, we spend most of our adult life rebuilding a breaking brain!

Let us now see what happens when the brain is broken suddenly and in a localized area, such as in a stroke. In this case, millions of neurons and billions of connections are lost. As long as at least some of the original functional regions are intact, it ought to be possible to stimulate them to grow more connections, much as in aging, and this will allow some recovery of function. The complete loss of a given area would be impossible to recover from, however, because adding new connections to a region involved in vision is not going to allow language to recover.

Studies in laboratory animals have shown that whenever there is significant recovery of behaviour, there is a growth in the dendritic trees, and thus in the number of connections, in the remaining regions involved in the damaged behaviours (e.g., Kolb, Gibb, & van der Kooy, 1994). Furthermore, it appears that various factors may play an important role in stimulating the growth of new connections. These are known as "trophic factors" and they are chemicals that the brain normally produces during development in order to assist in generating connections in the first place. In addition, it

should be apparent that since experience can influence connectivity, it ought to be possible to develop therapies that would stimulate the brain to produce more connections that might assist in recovery of function.

One of the implications of the plastic changes that underlay both recovery from the processes of aging and brain injury is that the brain may be more successful in recovering in younger animals than in older ones because there are more remaining neurons to change. This pattern is unfortunately the case. Another implication is that people who have brain damage early in life may not be as successful in staving off the effects of age on the brain. This pattern, too, appears to be the case.

10.4 Conclusion

One of the fundamental, and most interesting, properties of the brain is its plasticity. This feature of the brain allows us to learn and to benefit from experience. It also allows us to live a relatively long life during which we are able to continue to learn new things. Current research is elaborating the ways that the brain is plastic and discovering ways to control the mechanisms that underlay plasticity. This offers considerable hope that sometime in the not too distant future we will be able to build better brains and to repair broken ones.

Chapter 11

Questions About The Brain's Behaviour
Bryan Kolb

It is often stated that the basis of an education is to learn the questions. My goal in this essay is to introduce you to the key questions about the best computer you will ever own—the one between your ears.

We begin by reviewing the life of our hero, Felix. Felix is a university student. He awoke at his usual time of 7:30 AM on a Friday morning. As he sat up he was puzzled as he tried to figure out if his memory of a peculiar episode, in which he tried to subdivide his bedroom, was a dream or reality. It was a dream! He got up, showered and dressed, and made his usual breakfast of coffee and toast while listening to the morning show on the radio. Felix had stubbed his toe the day before and it was still sore so he took some aspirin. He left his apartment and headed for school in his old car. He met his classmates before his biology class and they discussed the upcoming exam. He later went to the computer lab and worked on a paper. By 4:00 he was fed up with university and headed for the Student Union pub to unwind over a beer with some of his colleagues, including his girlfriend, Phyllis. He and Phyllis later went out for dinner and discussed moving in together. They went to a movie and slept at Phyllis' place.

There is little in the day in the life of Felix that is remarkable. Indeed, I can forgive readers who found his life pretty boring. However, let us consider his life in a different way by doing a postmortem on what his day entailed. When Felix awoke, it was at the same time he always awakes on school days. Why? Felix does not use an alarm so what is it that allowed Felix to wake up, or for that matter, what is it that caused him to go to sleep in the first place? Felix had a dream about subdividing his bedroom. Why would he have a dream at all, let alone one about his bedroom? (Think about his plans with Phyllis!) Why did he have to consider whether his experience was a dream or reality. What is different about his sleep when he was dreaming and his sleep when he was not

dreaming? Felix got up, and went to shower and get dressed. To do this, he had to know where the bathroom was, how to turn on the water, how to do up buttons and shoe laces and so on. We take these abilities for granted but people with certain types of brain injury can no longer perform simple operations such as showering and shaving, even though they may be quite capable of carrying on an intelligent conversation. Further, children do not find operating showers simple.

Felix had his breakfast and while doing so he consumed two drugs (aspirin and caffeine). Later he consumed another drug (alcohol). Why did the caffeine and alcohol affect his mood and energy level while the aspirin did not? Why did the aspirin make his toe feel better? Felix drove his car to school and barely gave a thought to how to work the gears, how to park, or what route to take. He also stopped at a red light and drove through a yellow one, but the whole time he was singing along to the music on the radio. How did he accomplish this? Why did he briefly stop singing when he turned left across traffic? Does he have some kind of map in his head that he followed or did he get from home to school by just following some sort of overlearned route? Why did he stop at one colour light and not at another? Is Felix capable of doing any two things at the same time or is there something special about singing and driving?

Once he met his friends he spoke to them and listened to what they said. How did he do that? We take language for granted, yet if we are in a country with a language we do not speak, we quickly appreciate the complexity of language. Felix wrote an exam by pen and later wrote a paper on a computer. How did he do this? Were the same processes at work in handwriting as in typing? Felix retrieved information about the exam in order to answer the questions. How did he do this? How did he know where in his mind the information was? Felix and Phyllis spent the weekend together. Presumably they enjoy each other's company and they both get some type of physical and emotional reinforcement from being together. Why do they find each other attractive and not another person or a person of the same sex? Why do they enjoy physical contact? When Felix and Phyllis discussed their future, both had mental images of what it would be like. How did they generate these images?

11.1 The Brain Hypothesis

We have considered but a small sample of the questions that could be asked about Felix's behaviour. Having identified some of the questions, let's now consider how we could go about answering them. To do this, we need to consider another question. What controls behaviour? Probably the most common answer to this question is the "the mind". But what is the mind? Historically, there are two views on this question. One view, which was first clearly stated by Descartes in the 1600s, is that the mind is a process that is separate from the body. For Descartes, it is some kind of spirit or energy that is the basic essence of the self. It may utilize some aspects of bodily functions, but it is essentially unconstrained by the body. This position can be called a dualistic one. A second view, which is the one held by most neuroscientists, is that the mind and the body are the same thing, the body in this case referring to the brain. In this case mental events are simply a property of the brain's activity. If the brain is altered or dies, the mind is altered or gone. This position can be called a monistic one. It is sometimes

also called the brain hypothesis.

The implications of these two views of the mind are very different. In the former case behaviour can be produced independently of the brain but in the latter case it is entirely dependent upon the brain. Felix is essentially his brain in this view. The reader can see that explanations of abnormal behaviour, whether it be in the form of a mental disorder such as schizophrenia or in the form of a pathological fear of virgins is explained very differently. I note, parenthetically, that acceptance of the dualistic and monistic views has implications for other things such as questions about the soul, the uniqueness of humans, and so on but I leave these for the reader to ponder privately. For the rest of this essay I shall accept the brain hypothesis. The question now becomes "how does the brain work?"

11.2 Principles of Brain Organization

The brain is arguably the most complex thing that humans can try to understand. It is thus with some temerity that I attempt to identify basic principles.

11.2.1 The brain is composed of neurons and glia

The human brain is a 1350 g organ found in the skull. It is composed of two types of cells: neurons and glial cells. Neurons are the functional units of the brain whereas glial cells are support cells. Estimates of the numbers of cells in the human brain usually run around 10^{10} neurons and 10^{12} glial cells, although the number could be even higher. As neurons are the basic functional unit of the brain, I shall consider only their structure and function and leave glial cells as a mystery for now. The reader might be interested to know, however, that the one distinguishing thing about Einstein's brain appears to be an unusually large number of glia.

There are numerous types of neurons but they share several features in common (see Figure 11.1). They have a cell body, which like most cells contains substances that determine the function of the cell. There are processes protruding from the cell body, which are known as dendrites. Dendrites function to increase the surface area upon which a cell can receive information from other cells. Finally, there is another process, known as an axon, originating from the cell body. Different types of neurons are morphologically distinct, just as different types of trees and bushes are morphologically distinct. The morphological differences in neurons reflect differences in function.

Neurons are connected to one another via their axons; any given neuron may have as many as 15,000 connections with other neurons. The behaviour of a given neuron is determined by the summed activity of all of the connections. The behaviour of a neuron can be conceptualized as being like a light bulb hooked to a rheostat. If we set the rheostat fairly low, the light will be on, but will appear quite dim. From this point we can either increase or decrease the brightness by increasing or decreasing the current going to the bulb. For a neuron, the amount of current going to the cell depends upon the dendrites. If the connections to the dendrites increase the current, the neuron becomes more active; if the connections decrease the current, the neuron becomes less active. Most connections with other neurons are chemical. Thus neuron A releases

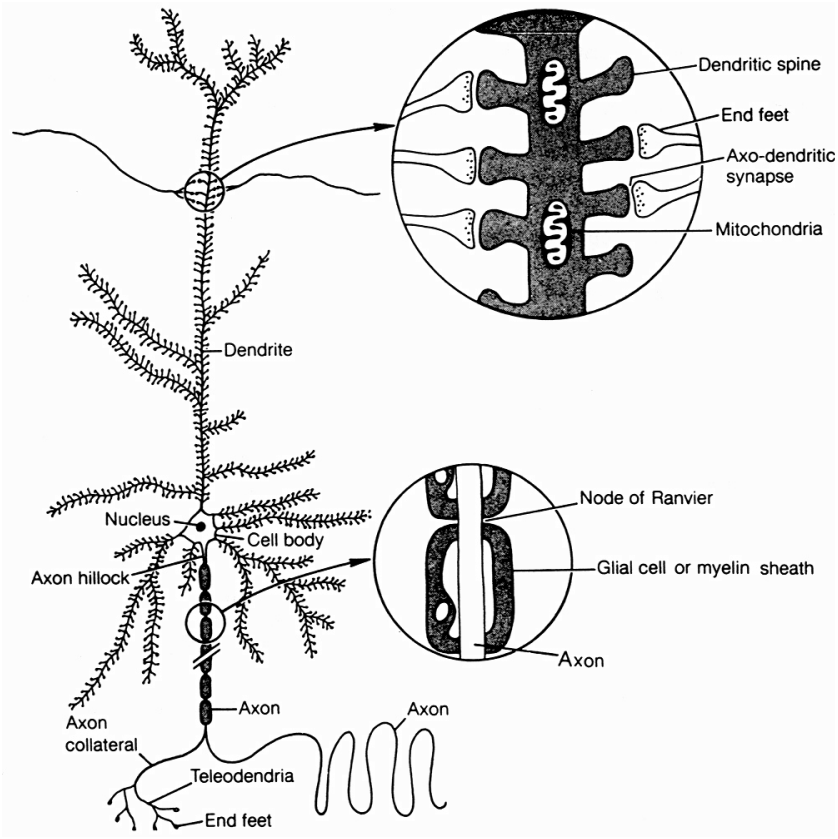


Figure 11.1: A typical neuron showing some of its physical features.

a chemical, known as a neurotransmitter, and the dendrites on neuron B are altered briefly by the chemical. Different neurons release different transmitters, the two most common ones in the brain being glutamate and GABA (gamma-amino butyric acid). Glutamate functions to increase activity in other neurons whereas GABA functions to decrease it. Thus, a drug that increases glutamate activity will serve to increase brain activity whereas another that increases GABA will decrease brain activity. Another interesting transmitter is dopamine. When dopamine activity is increased, we have a pleasant feeling. Amphetamine and cocaine increase the activity of cells releasing dopamine. We can begin to see the answer to one of our questions: the drugs Felix took all affected different transmitters and therefore had different effects on brain activity.

11.2.2 Neurons and glia are organized into discrete functional regions

Although neurons are the basic units of the brain, individual neurons are of little consequence. In fact, we lose hundreds or perhaps thousands daily without much obvious effect. Rather, it is the organization of the neurons into larger groupings that is important. The brain has thousands of such groupings but we can conceptualize it somewhat more simply. The most obvious feature of the human brain is that there are two large hemispheres, which sit on a stem, known as the brainstem. The brainstem connects with the rest of the body. Imagine a pumpkin. Now imagine your fist stuffed into the pumpkin. The pumpkin represents the cerebral hemispheres, the fist represents the brainstem, and the forearm represents the spinal cord. The brainstem is essentially all that primitive animals such as fish or frogs have for a brain. The cerebral hemispheres are a predominant characteristic of the brains of mammals and birds.

The cerebral hemispheres and brainstem are not single regions. Rather they are both made up of hundreds of smaller regions, each of which has unique connections and, in many cases, unique cell types. Consider the cerebral hemispheres. They have an outer covering, much like the rind of the pumpkin, which is known as cortex. The cortex is the dominant feature of the human brain. It can be further divided into dozens of subregions, each of which is a functional entity.

11.2.3 Functions are localized in the brain

The many different anatomical regions in the brain are functionally distinct. This is perhaps most graphically shown in the surgical theatre. Wilder Penfield at the Montreal Neurological Institute developed a procedure for studying localization of function in the awake human. Imagine the following. You are given a little Valium (which calms you down) and taken to the operating room. Your hair has been shaved off and a local anaesthetic is applied to the side of your head. The surgeon makes an incision on the side of your head, much in the shape of the top of a question mark lying horizontally. She then makes a flap of skin to expose much of the side of your skull. She then drills four 1 cm holes in your skull with a hand drill, and, using a small saw, connects the holes to remove a piece of bone about the size of a medium pancake. Your brain is visible. Because you are awake during this procedure, she can now ask you to do different things. For example, she can ask you to talk or to move your arm and so on. As you do

things for her she takes a tiny electrode and passes a small current in different regions of the cortex. Stimulation in one region stops you from speaking. Stimulation in another causes your little finger to move, and so on. She has demonstrated that functions are localized to different places in your cortex. We can now ask a reasonable question. What happens if we find a place that, when stimulated, affects our speech, and then we surgically remove it. If functions are localized, we would predict that the function will be lost. It is—provided we take all of the region. This observation is important for it speaks directly to the brain hypothesis. If we can eliminate functions surgically, they must be a property of the brain's activity. Furthermore, we can predict that if a region of brain is functioning abnormally, then behaviour will also be abnormal. I should note at this point that surgical studies of the effects of brain stimulation are not done on healthy people. They are done in preparation for surgical removal of abnormal tissue. The surgeon must locate critical functions such as speech or vision in order to avoid removing them accidentally!

11.2.4 The brain is organized into functional levels

In order to understand this we must first consider a taxonomy of behaviour. Behaviours vary in complexity. Perhaps the simplest behaviours that humans have are spinal reflexes. We are all familiar with the effect of tapping the tendons just below the knee cap when we are sitting on a counter with our legs hanging down. The lower leg "involuntarily" moves. Similarly, when something suddenly comes toward our eye, we blink. Now consider repetitive behaviours such as walking. We walk without much thought. This repetitive, rhythmical, movement is also largely controlled by the spinal cord. A more interesting example is pelvic movements during sexual intercourse. These too are controlled by the spinal cord.

The next level of behaviour includes units of movement such as the making of different body postures to maintain balance and behaviours such as swallowing. These behaviours are controlled by the first region above the spinal cord, which is known as the hindbrain. It would correspond to your wrist in the pumpkin and fist metaphor. The hindbrain not only controls the production of some kind of movements but it also controls sleep and waking. Thus injury to this area may produce a coma because the mechanism that wakens the brain is damaged. Thus, we can see one of the reasons that Felix awoke. A region in his brainstem acted like a light switch and turned on the rest of the brain.

The next level of behaviour includes behaviours that direct the body to sensory events. Thus, when we hear a noise we turn our head in the direction of the noise, which is known as an orienting reflex. This reflex is controlled by the midbrain, which corresponds to your hand in our metaphor. As we progress to the next level of behaviour, we find that the behaviours begin to get more complex quickly. Thus, we begin to add goal-directed behaviours. These are behaviours directed at things, such as food, other animals, and so on. Indeed, we have now added the behaviours seen in babies. This includes some affect, some preferences for different sensory events, and preferences for particular foods. It also includes the ability to control body temperature. This control may seem like a simple thing, but think about the problems involved. When we exercise we heat up. There must be a way of controlling this. When we sleep,

we make few movements and heat must be generated. Males are all too familiar with the way that the body controls the temperature of one of its internal organs— the testes. The extension and retraction of the testes is done in response to changes in body temperature in order to maintain the viability of sperm. These behaviours are controlled by a region known as the diencephalon. It would correspond to your fingers clenched as a fist in the pumpkin. Next, we add behaviours necessary for self-maintenance. These include sexual behaviour, maternal behaviour, feeding and drinking, and so on. Finally, we add the ability to understand the sensory world and to be flexible in response to a changing environment. These latter functions are functions of the cerebrum (the pumpkin). The names of the various brain regions are not important at this stage. The important point is that behaviour varies in complexity and different levels in the brain control different types of behaviour. Although we have not specifically discussed it, given point two (functions are localized) it must be the case that different functions are localized at each level.

11.2.5 There are multiple pathways for each general function

Imagine looking at a can of diet coke. What do you see? Now imagine moving the can to another place in front of you. What did you see when you moved the can? Presumably the answers to these questions are something like the following. First, you saw a cylindrical object that was red, white, and silver, and had writing on it. When you moved it, you saw the same thing. Intuitively we conclude that the brain must "see" the object much as we have seen it. Consider now what happens when we damage different parts of the cortex. People with damage to one area, which is known as visual area 4, can no longer perceive the colour. The red appears black, the silver is sort of grey and the white is a light grey. This condition is quite different from the usual "colour blindness" that is a result of a faulty retina in the eye. People who have retinal colour blindness still have some visible colours. People with damage to V4 have a very different sensation. Indeed, they typically lose weight because food becomes so unappealing. Think about eating a black tomato or grey potatoes. Now consider a different problem. People who have damage to another visual area, known as visual area 5, cannot see movement. What does this mean to our coke can? The can is seen normally when it is still but is invisible when it moves. When it comes to rest again it is visible again! This example is really counter intuitive because we believe that we perceive the same object when it is still and when it is moving. That is not how the brain works, however. Consider the implications of this dysfunction. You would be unable to cross the street because you would look down the road and see a car. Then you would start to cross and it would become invisible—until it stopped as it hit you. Or, consider talking to someone. You see their face but as their mouth moves it vanishes. People with damage to visual area 5 learn to avoid looking at people who are talking. This is the only way they can cope with the vanishing mouth. These two examples of visual defects illustrate the fact that the brain takes the world apart and analyses different features independently. A similar logic is used for the analysis of other senses too.

Another interesting question arises from our example of the diet coke can. How did you imagine the can? Did you use the same brain regions that you use to actually

perceive one? What would happen if you were actually looking at a coke can and then you imagined another one sitting beside it? How do you know which one is real? After all, both are in your head? Finally, if the visual system takes the coke can apart and different sensory channels analyse different bits, why do we see a coke can at all? Is there some area of the brain that puts it back together? (The answer to the last question is "no".) Finally, we can ask more fundamental questions. How do neurons do all these things anyway? These are the questions asked by cognitive scientists and neuroscientists.

11.2.6 The cerebrum is functionally asymmetrical

One of the curious features of the human brain is that it produces language. We will not consider how or why it does this in the current essay, but we will ask where it does this. In our surgical example above we noted that a surgeon can find speech zones in the cortex. Thus, we know that at least cortical areas are involved. What we did not discuss, however, was that it is only in the left side of the cerebrum (the left hemisphere) that the surgeon can find language. Thus, language processes are localized to the left hemisphere. This leads to other questions. Is language the only function that is unique to the left hemisphere? Why is language in the left hemisphere? If the left hemisphere is talking, what does the right hemisphere do? These are questions of neuropsychology.

11.2.7 There are individual differences in brain organization

It is a common observation that even siblings behave very differently from one another. If the brain is controlling behaviour, then it follows that there must be differences in the brains of different people, and these brain differences account for the behavioural differences. This account is not a big surprise as people do not look the same, they have different fingerprints, and so on. A more interesting question is whether there are any variables that are correlated with individual differences. There are at least two: sex and handedness.

I doubt that many people would be surprised at an assertion that males and females behave differently. The idea that this difference results from a difference in their brains is probably less obvious, but it follows directly from the brain hypothesis. A more interesting question, however, is why are the brains different? Studies with both laboratory animals and humans have demonstrated that the major reason is the influence of gonadal hormones. Thus, during brain development in the womb, males produce testosterone, which alters their brain.

Similarly, during development, and especially during puberty, the production of hormones by both sexes alters the brain. Take a simple example. Males have a penis and the penis has behaviours. These behaviours are controlled at the lowest level by a region in the spinal cord known as the nucleus of the bulbacavinosa. It should be no surprise that females do not have this region! It may be a surprise, however, to learn that they do have one early in brain development but that it degenerates. The reason it degenerates is because they do not produce enough testosterone. This simple example illustrates that hormones can alter the structure of the nervous system. Obviously,

the more important question is whether they can also alter the brain itself. Studies in various animal species have shown that they can.

Consider a simple example. Male song birds sing and most female song birds do not. This behavioural difference is due to a fundamental difference in brain structure, which is dependent upon gonadal hormones. Structural differences in the brain of humans are more difficult to study but there is now little doubt that there are fundamental differences in the structure of male and female brains, and that circulating hormones affect brain function throughout our lives. This idea leads to interesting questions. Can we actually measure differences in behaviour of females during menstrual cycles? Does the cognitive behaviour of females change with menopause? Does hormone replacement make a difference to these behaviours? What changes occur in males? Do hormones have any effect upon the way that the brain is able to learn from experience?

The second major contributor to individual differences is handedness. About 10% of the population is left-handed. If we examine the brain of left handed people, we find that about 1/3 have language in the right hemisphere, while most of the rest have language in the left hemisphere. A small proportion actually have language in both hemispheres. Why do some people have language on one side versus the other? What other functions are different in left-handers? Does handedness make any difference to intelligence or to anything else? Why are some people left handed? Do other animals show handedness? These too are questions of neuropsychology.

11.2.8 The brain is plastic

When we exercise muscles, they change. When we use the brain, it changes too. This property is known as brain plasticity, and with some reflection, it should be obvious that this is a necessary characteristic of the brain. When we learn things we are sometimes able to remember them. In view of the brain hypothesis, this means that the brain must have stored the information. If the brain stored the information, it must have changed in some way. Indeed, this must be true for every experience that we can recall. If we injure the brain there is often some recovery of lost functions. This means that the brain must somehow have changed to allow this to happen. Brain plasticity is the topic of another essay, so I will not consider it further here.

11.2.9 Brain activity can be influenced by knowledge

Cognitive psychologists often talk about “bottom-up” and “top-down” processes in the brain. Bottom-up processes are those processes that are driven by data affecting the brain. For example, when we analysed the colour of the coke can, we perceived red. This perception was driven by the data. Top-down processes are rather different. Many of our perceptions are affected by knowledge. Suppose there is a huge pile of pop cans and our task is to find a diet coke can. How do we sift through the visual clutter and find the diet coke can? Or, if we are looking for a face in a crowd, how do we ignore all the faces but one? Obviously, our behaviour is driven by knowledge of what a coke can, or a particular face, looks like. This knowledge must influence how the brain works. How does it do this? Are there rules that govern how top-down processes and how do they govern brain activity? These are examples of questions asked by cognitive scientists.

11.3 Conclusion

I have tried to introduce the reader to some of the questions that psychologists ask when they consider the brain's behaviour. We began by considering the behaviour of Felix. A simpler task now is for the reader to consider his or her own behaviour in the same way. Of course, we have not asked the major question of psychology: how is it that we are conscious and can even ask these questions? This is a more difficult question best saved for another time.

Chapter 12

Behavioural Genetics: The Study of Differences

Martin Lalumière

A large part of psychological science is concerned with identifying, cataloguing, and explaining individual differences. One does not need a degree in psychology to observe that there are enormous differences in how people think, feel, and act. Everyone recognizes differences in the personalities of one's family members, for example. These differences exist even though family members live in the same culture, historical era, political climate, economic circumstances, and house (at least while growing up). Think of five of your friends: Even though they have many things in common (they are probably close in age, and they all like you), I would bet that they differ in how smart they are, how outgoing they are, how emotionally intense they are, how political they are, and so on. But organizing and explaining the origins of these differences, whether cognitive, emotional, or behavioural, is more difficult.

Behavioural genetics is one of many ways to study the origins of individual differences. It is an important topic to study because, unlike many other areas of psychological research, it has produced a large body of knowledge that has survived the test of time. The results of this field of psychology stand as close to facts as any result in psychology. Many behavioural geneticists have moved on to other topics of research because they have concluded that most of their original questions have been answered.

Behavioural genetics is about partitioning the sources of individual differences in any trait that can be measured reliably. The fundamental question is how much of the observed variability in a given trait can be explained by the fact that people have different genes, and how much can be explained by the fact that people have been exposed to different environments. One could turn the question around and ask how much of the similarity among people can be accounted for by their genetic similarity, and how much can be accounted for by their exposure to similar environments. Although the

name of the discipline emphasizes genetics, the focus is on both genes and environment. In fact, as we will see, this discipline has much to say about the importance of the environment in creating individual differences.

12.1 Research Methods

Behavioural geneticists are lucky. There are known genetic differences among people, such that the size of the genetic difference can be compared to the observed difference in the trait under study. Similarly, there are known environmental differences that can be related to trait differences. When it comes to genetic differences, the level of precision of the magnitude of the difference is quite good. We know that, on average, siblings share 50% of their polymorphic genes[1], whereas cousins share 12.5%. Therefore, if differences in a trait (say, intelligence) can be explained by genetic differences, then siblings should be more similar in intelligence than cousins. Of course, siblings could be more similar because they grew up in the same house, whereas cousins did not (in many cultures cousins do grow up in the same house, but here we will assume that they do not).

The degree of environmental differences is not as easy to quantify. We know that siblings are exposed to more similar environments than cousins, but how much more similar is unclear. Siblings grow up with the same parents, are exposed to the same physical environment (the house, the number of books on the shelves, the family income, the neighbourhood) whereas cousins grow up with different parents, perhaps in different economic circumstances, perhaps in different neighbourhoods, etc. All that we know for sure is that siblings are exposed, on average, to a more similar environment than cousins. The exact degree of similarity requires precise measurement, something that is not done very often.

The two main statistics of behavioural genetics are *heritability* (h^2) and *environmentality* (e^2). The values for both statistics vary from 0 to 1, where 0 means that none of the observed variance can be accounted for by genes (for heritability) or the environment (for environmentality), and where 1 means that all the variance can be accounted for by genes (or the environment). In the classic research designs described below, h^2 and e^2 are mutually exclusive, such that the total amount of observed variance is equal to the sum of h^2 and e^2.[2] Heritability should not be confused with inheritance, the latter meaning shared genes passed on by parents. A trait can have low heritability but be completely inherited (e.g., binocular vision). Can you think of the reason for this?

There are two main research designs used to tease out the effects of genes and environment on phenotypic differences.[3] The first is the twin and sibling design and the second is the cross-fostering or adoption design. Of course, scientists do not experimentally create twins or force adoption of siblings. These events happen naturally and

[1] Polymorphic genes are those that come in different versions (e.g., genes having to do with eye colour). In contrast, monomorphic genes are those that all members of a species share.

[2] Technically, the formula also includes covariance and measurement error, but for pedagogical purposes we can ignore these terms here.

[3] The phenotype refers to the manifest (or observed) characteristics of an organism, whereas the genotype refers to the genetic complement.

scientists take advantage of them to study the influence of genes and environment on the development of individual differences.

There are two types of twins. Identical or monozygotic twins (MZ) share all of their genes (100% relatedness) and non-identical or dizygotic twins (DZ) share only 50% on average. DZ twins are just like regular siblings, except that they were in the same womb at the same time. Thus, if genetic differences account for most of the difference in a given trait, MZ twins should be about twice as similar as DZ twins. Of course, MZ twins might be more similar than DZ twins because people treat them the same way (after all, they are very similar physically) or for other reasons. Therefore, scientists have also studied a rare group of people: MZ and DZ twins who have been reared apart. In this case, any greater similarity in MZ than in DZ in an observed trait is likely due to genetic similarity and not environmental similarity.

In cross-fostering designs, unrelated siblings are reared together. Therefore, any similarity is due to the fact that these individuals are reared in the same environment, not to the fact that they share similar genes. One can also examine the degree of similarity of foster or adoptive parents and their genetically unrelated children.

12.2 Estimating Heritability and Environmentality

Table 12.1 shows the degree of similarity in intelligence as a function of degree of genetic similarity. As you can see, pairs tend to be more similar in intelligence as they are more similar genetically. The results also show that being exposed to more similar environments (being reared together rather than apart) is associated with greater similarity in intelligence. These results are very typical of behavioural genetics findings for many different traits: part of the observed individual differences can be explained by genetic difference, and part can be explained by environmental differences. This may seem like a trivial finding to you, but for a very long time psychologists and other social scientists thought that people's experiences were the sole determinants of their behavioural characteristics.

There are many different ways to estimate h^2 from twin and sibling studies. The intra-class correlation for a given trait in MZ twins raised apart directly estimates broad-sense heritability (h^2). Broad-sense heritability reflects all genetic effects that can only be shared by identical twins.[4] In the example given above, broad-sense heritability of IQ is 0.72.

Twice the difference in the correlation for MZ twins reared together and DZ twins reared together estimates narrow-sense heritability, assuming absence of non-additive effects.[5] In the example above, narrow-sense heritability is $2 * (0.86 - 0.60) = 0.52$. Twice the correlation for parent and offspring living apart from birth is an estimate of narrow-sense heritability that does not require the assumption of absence of non-additive effects (because these cannot be passed from parents to offspring). In the

[4] Broad-sense heritability = additive effects of genes at specific locations + dominance (non-additive effects of genes at a single location) + epistasis (interaction of genes at different locations, including different chromosomes). Dominance and epistasis can lead to MZ correlations that are more than twice as large as DZ correlations.

[5] Narrow-sense heritability includes only additive effects of genes and is always equal to or less than broad-sense heritability.

Pair Relations	Genetic Similarity	Similarity in IQ (r)
Identical twins raised together	100%	.86
Identical twins raised apart	100	.72
Non-identical twins raised together	50	.60
Biological siblings raised together	50	.47
Biological siblings raised apart	50	.24
Parent-offspring together	50	.42
Parent-offspring apart	50	.22
half-siblings	25	.34
cousins	12.5	.15

Table 12.1: Similarity in intelligence (based on intra-class correlation (r)) as a function of degree of genetic similarity (data from Bouchard & McGue, 1981).

example above, narrow-sense heritability is 2 * 0.22 = 0.44. As you can see, different calculations produce slightly different results. What matters is the range of values obtained across many studies using different designs. In the case of IQ, that range is centered on 0.50 for narrow-sense heritability and 0.75 for broad-sense heritability. Heritability of IQ increases with the age of the subjects in the study. Can you think of a reason for this counter-intuitive finding?

The correlation between unrelated siblings reared together directly estimates environmentality (e^2). Another way to assess e^2 is to subtract h^2 from 1.0 in studies of MZ and DZ twins reared together, assuming perfect reliability of the measure. Another way is to calculate the difference in the correlation obtained for MZ twins raised apart and MZ twins raised together. For IQ, and using data provided by Bouchard and McGue (1981), these methods produce e^2 values of 0.34, 0.48, and 0.14. The value of 0.48 is inflated because measures of IQ are in fact not perfectly reliable.

12.3 Some General Findings

It is now well accepted that most psychological characteristics show substantial heritability. These include, as we have seen, intelligence, and also more specific cognitive abilities, as well as different aspects of psychopathology (e.g., schizophrenia, depression, and anxiety), antisocial tendencies, and personality. Bouchard (1994) reported that the five major dimensions of personality (extraversion, neuroticism, conscientiousness, agreeableness, and openness) show heritability values of 0.40 to 0.50. It is also well accepted that environmentality accounts for a large portion of individual differences in these domains.

One of the most interesting findings of behavioural genetics has to do with the na-

ture of environmental effects on the development of individual differences. Behavioural geneticists distinguish between two types of environments: shared and non-shared. The shared environment refers to the part of the physical and social environment that is common to all siblings. For example, siblings live in the same house, have the same number of books in the house, live in the same neighbourhood, and have the same number of parents. The non-shared environment refers to the part of the physical and social environment that differs among siblings. Siblings have a different birth order, may be treated differently by parents, and so on. Results of behavioural genetics studies suggest that the type of environment that most influences the development of individual differences is non-shared. That is, aspects of the environment that are shared by all siblings appear much less important in influencing the development of individual differences than aspects that are not shared by siblings. In fact, siblings (other than monozygotic twins) tend to be quite different from one another in their personalities, despite sharing similar genes and living in a similar environment (Plomin & Daniels, 1987). Elsewhere we have suggested that there might be an evolved family process that accentuates sibling differences (Lalumière, Quinsey, & Craig, 1996). For some characteristics, the longer MZ twins live together, the more different they become.

Sociologists and psychologists have traditionally studied environmental causal factors that are part of the shared environment (e.g., socioeconomic status, overall parental affection). They are now turning their attention to possible causal agents that are not experienced similarly by siblings.

12.4 Some Limits and Cautions

All research methodologies have limits and require certain assumptions (statements of fact that are taken as true but that cannot be directly examined). Behavioural genetic designs have well-known limits and well-specified assumptions. Results of individual studies have to be interpreted in the context of these limits and assumptions. For example, a study of twins reared apart assumes that twins are reared in a range of environments that is similar to the natural range of environments found in human families. This may not be true, because adoptive families are screened for suitable environments for adoptee placements. This means that the range of adoptive environments is smaller that what is found generally, resulting in an overestimate of h^2 and an underestimate of e^2.

One important cautionary statement is that heritability and environmentality statistics refer to population values at a given point in time. To say that height has a heritability of 0.80, for example, says nothing about the importance of genes versus environment in explaining your own height. Your height is completely determined by both your genes and your environment—if one or the other didn't exist, you would have no height. Heritability has to do with explaining differences among a group of people. The fact that some people in this class are shorter and others taller is mostly caused by the fact that people have different genes (rather than caused by exposure to different environments). Also, the heritability of height has increased from 0.50 to 0.80 in the last 50 years, clearly showing that it is a dynamic value. Can you think of why heritability of height is higher today than it was 50 years ago?

Another caution is that heritability says nothing about the mutability (the potential to change) of a characteristic. Although IQ is highly heritable, cognitive abilities can be greatly improved or suppressed through exposure to certain environments. IQ scores have increased constantly over the last 100 years (so much so that standard IQ tests have to be re-normalized on a regular basis), and few people believe that this change is due to changes in the gene pool. Behavioural genetics speaks to what can be observed at a given point in time, not to what can be done about improving the human condition.

Chapter 13

Where Does Schizophrenia Come From?
Jennifer Mather

You might see the question in the title of this chapter as part of a larger one—where does behaviour come from? Experts might answer the question differently, depending on their individual points of view. A comparative psychologist with evolution in mind might emphasise that humans are primates, and look for the roots of our behaviour in survival-based reactions of monkeys. A neuropsychologist might see this question as an opening to discuss how the brain programs behaviour, and go on to discuss the contribution of the frontal cortex or the dopamine neurotransmitter system. A developmental psychologist would look within your lifespan and say that all the behaviour you show now, as a young adult, was built up through your childhood.

Schizophrenia is an example of adult behaviour (in this case a set of abnormal behaviours) that must have had some foundation through childhood. But it isn't always easy to see this. It's a mental illness, a set of wrong feelings, thoughts, and behaviours. It usually hits in young adulthood, often quite suddenly. So how can something like this have a background in childhood?

13.1 What Schizophrenia is

To figure out where schizophrenia comes from, and how it arises, we first need to know what it is. It's a mental illness, and a serious one, as it wrecks the lives of a lot of people. One out of 100 people has at least one schizophrenic break—that's 600 in Lethbridge alone. Mental illnesses are catalogued by the *American Psychiatric Association*, in a book called the *Diagnostic and Statistical Manual of Mental Disorders* (the latest edition is called *DSM-IV*). Psychiatrists diagnose schizophrenia by a group of symptoms.

One symptom is delusions, beliefs about things that are not true (these are often paranoid delusions, believing that "they are out to get you"). Another is hallucinations, perceiving things that aren't really there. People with schizophrenia most often have auditory hallucinations, particularly hearing voices. Maybe to rationalize that they hear things that seem to come from inside their heads, people with this disease may believe that God is speaking to them, or that someone has planted a transmitter inside their brain. Often people with schizophrenia also have what's called inappropriate affect. Their emotions don't seem to be properly expressed, or often don't match what the situation calls for, such as laughing at a sad movie.

Families of schizophrenic individuals often describe how the illness seemed to hit suddenly. Promising young men or women suddenly begin to suspect everything their family does. They get commands from Jesus or feel they're being bugged by the local police. When I worked with schizophrenics in the U.S., they often felt watched by the Communists, but in Canada we weren't quite so paranoid about the "Red Menace" and Canadian schizophrenics settled for the local police force. They stay up late, don't go to school, or drop out of work. Their delusions and strange concentrations take over their lives. My next-door neighbour in graduate school suddenly began to feel that his professors were out to get him, that they were working to discourage and fail him (you can see this delusion takes some time to be proven wrong). A promising life seems to have become derailed, and it takes everyone about two years to understand that it's a mental illness, not a phase or adolescent rebellion, and not something that can be cured by patience or "tough love".

13.2 What Happens if you Develop Schizophrenia?

What happens to people who have this illness? Often they have to leave school, quit work, be hospitalized for help, and generally leave the normal world we know. The good news is that schizophrenia can be treated. Right now the best treatment is antipsychotic drugs that seem to damp down the disruption of brain function and get people back on an even keel. Psychotherapy isn't any help, mainly because it's the brain and not the environment or personality that's going wrong. People with schizophrenia are particularly sensitive to stress and have to learn or relearn how to deal with social situations and manage stressful situations like job pressure or exam time. Thus, behaviour management and socialization therapy are also necessary.

If you follow a large group of people who have had a diagnosis of schizophrenia, you will find that the result is not gloomy for everyone. About a quarter of them recover. Maybe they have a few symptoms, maybe they have none at all, as if they have recovered from the flu. About a third do okay, but have problems. They may recover and then have relapses. They often give up their former high ambitions, and settle for a life as a clerical worker, unskilled labourer, or housewife. They take drugs to help, and they have to watch out for stress—they can't handle high demand very well. A third do badly. They are in and out of hospital and in and out of jobs. They never marry, sometimes live in old hotel rooms, and sometimes live on the streets. In Lethbridge

you can find the worst affected by schizophrenia hanging around places such as the Public Library or the malls. One man was even living in a cave in the coulees. They are "losers" in life in that they miss out on most things non-schizophrenics take for granted, and our social support system doesn't know how to handle them very well.

The most obvious losers are the 10% of schizophrenic individuals who commit suicide. Many of them had shown a lot of promise before their illness hit. Most don't commit suicide when they are hospitalized, but rather after they're getting somewhat better. After all, if you know the life described above was all you had to look forward to, mightn't you despair?

Back to the question of where the illness called schizophrenia comes from. For almost any behaviour, the answer is that it partly comes from inheritance and partly from environment. It is very difficult to get good evidence for this perspective for any behaviour, because behaviours aren't clear. If you have blue eyes everyone agrees they are blue, and everyone can count that we have five fingers. But what about paranoid behaviour? When is too much? When is paranoid behaviour justified (i.e., maybe the prof really is out to fail you), and when is it irrational? So we have to decide when something hard to measure—mental illness—is the result of something so complex as your many genes and your interactive environment. Nevertheless, we can trace it back to a combination of inherited *risk* and childhood *stress*.

13.3 Genetics and Risk

One way in which we find evidence of a genetic background for a characteristic of someone is to see if that characteristic occurs in his or her relatives. Each of us inherits one-half of our genes from our mother and one-half from our father, and we would have all our genes in common with an identical (monozygotic) twin. Of course, which ones you inherit are not known; it's all chance. Still, if inheritance matters, you would be more likely to have a schizophrenic break if others in your family have, and this is true. The risk for all of us is 1%. If you have a schizophrenic parent, your risk is much higher, 10% (it doesn't matter if it's your mother or father). If you have a schizophrenic identical twin, your risk of showing the disease is 50%. This rate is much higher, so inheritance clearly has something to do with the disease.

But families are complicated things. After all, your parents not only give you the genes that predict you, they also give you the environment that shapes you. Scarr and McCartney (1983) have called this a Genotype-Environment Correlation. They point out that your musical parents give you not only your musical ability, but a stereo and music lessons. Your shy personality means that you read a lot, stay at home and stay shy; maybe that's why shy people choose to be accountants or librarians (though accountants and librarians hotly deny this!). Your smart parents (on average) stressed learning and provided you with books; that's why you are at university and your friends aren't. So what your parents gave you is environment and heredity all mixed up together.

One way around the problem of genotype-environment correlation is to separate them, to look at adopted children. Their biological parents gave them their genes, and their adoptive parents gave them an environment (except for the important nine months

before birth). Studies of children who have a biological parent with schizophrenia, and adoptive ones without, show a risk of about 10%. Those with biological parents free of schizophrenia, and an adoptive one with the disease, show a risk of about 2%. So it's clear—the risk follows the genes, and what you inherit does affect whether you are *at risk* for schizophrenia.

Notice that just knowing the genes are there tells us nothing about how many there are, or how they act. It could be several of them, and the more you inherit the higher your risk. It could be one gene that gives you risk, and the environment could push you to mental illness.

But if there is some problem gene, where would it act? It would probably be on the brain, which is a huge and complicated structure that we are only beginning to understand a little. When scientists look at the structure of the brain of people with schizophrenia, there's nothing obviously wrong. But when they look at the metabolism they find that two areas of the cortex, the frontal and temporal, don't seem to be acting right. These areas are where I have worked with schizophrenia. I chose to study the eye movement system because the control of this system is all through the brain. When it goes wrong, the pattern of problems can be traced to dysfunction in one brain area. The pattern in people with schizophrenia looks like the one in those with damage to the frontal cortex, so the parallel suggests frontal dysfunction. I tested teenage children with one schizophrenic parent. Half of them showed this same pattern, which suggests that half of them have a vulnerable brain, likely to give them trouble if they end up in the wrong environment.

This result is also a clue to the timing of schizophrenia's onset. Our brain matures over all our childhood, and the last parts to be finished, by late adolescence, are the connections in the frontal cortex. That's when schizophrenia hits. Of course, late adolescence and early adulthood is also the time when youths leave home, get jobs and marry—all stressful events. So we have that genotype-environment correlation again.

13.4 Childhood and Stress

What could your environment do? It could do quite a lot, and it has 18 years and nine months to do it before you are considered an adult.

One part of childhood environment that people often don't consider is what happens before you are born. We used to think of the womb as total protection, cushioning the baby-to-be from everything outside. Now we know better. Everything from major stress on the mother, to her excessive drinking, to diseases she catches, can affect the unborn baby. These effects are particularly true at about the third month of the pregnancy, when the organs are forming in the foetus. Tracing backwards has shown that some people with schizophrenia were in this third foetal month when their mother went through one of the influenza epidemics that keep circling the world, usually in winter. What might happen? Influenza (flu) is a virus. It might get across the placenta from mother to baby. It could cause damage in the brain, minimal damage that doesn't show up until the brain matures fully—as much as eighteen years later.

Another part of very early childhood is the risky period near birth. Again, the statisticians find that people with schizophrenia are more likely to have had a stressful

birth than those without. It isn't anything specific. It might be lack of oxygen just after birth, prolonged and very stressful labour, some separation of the placenta before the baby is born, or something leading to an emergency Caesarean. It's hard to trace this, too. No one remembers their own birth! Hospital records and mothers' accounts of a birth may be different, and each isn't always correct, but something may have gone wrong then.

The largest part of the childhood environment is also the most complicated; we spend our whole lives in our families. We think of childhood as a time of innocence and support, but it often isn't. So much can go wrong, from sexual abuse to poverty to divorce of parents, from bullying and shyness to brain damage from car accidents and culture shock. Some of these factors, such as demanding and inconsistent parenting, seem more likely to push children toward schizophrenia.

So genes and childhood set the stage. What causes onset? It could be the brain. Maybe the brain defects that were caused before birth show up at maturity. Maybe it's the hormone surge at puberty, though that's a bit early. It could be the social pressures. Adulthood is stressful, and this may tip a vulnerable person over the edge to schizophrenia. One study in Israel showed that it wasn't short-term, extreme stress that did it, but longer-term, repeated hassles. There were few instances of schizophrenic breaks during the violent, but short Six-day War, but lots more during the demeaning and hassling six-month period in Boot Camp. Once the vulnerability is there, many of different stresses can do you in. The onset is later and less for women than men. Why? It may be because men, although physically stronger, are physiologically and psychologically weaker than women—more vulnerable. It may be that women's emphasis on friendship, support, and cooperation buffers them against stress and gives them more coping skills. We don't know for sure.

13.5 Adults and Schizophrenia

This chapter has discussed schizophrenia as if it were always the same symptoms, easy to see as the same disease. It definitely isn't. Part of that variability in persons with schizophrenia is laid down during childhood, too. These 600 people in Lethbridge coping with their disease are just that, variable people, normal except for the illness. With a more outgoing personality, with a happier childhood, and with an intact family to help them, they have a better chance to cope with schizophrenia when it hits. Of course, with an intact and loving family to help, anyone with an illness will also do better in the long run. Some will crumple under the impact of schizophrenia, some will live with the weakness they have newly uncovered, and some will just go back to their competent, former selves. But their past follows them—the better your childhood, the better you will cope with your schizophrenia.

Wordsworth said "The child is father of the man", and pop psychology now talks of "the child within" the troubled adult. It's not at all that clear; there is lots of potential for recovery, and we develop and change as adults too, learning all our lives. Research is helping; there are better drugs with fewer side effects, and we're learning what works in family therapy. Still, the schizophrenia that mysteriously appears in adulthood has its roots in the genes and environment of our childhood. It is an example that all we

are and *will be* comes from what we *were*.

Chapter 14

Illusory Visual Motion Perception
Jennifer Mather

When we talk about dealing with incoming information, we use the twin titles of "sensation" and "perception". There is a good reason for this division, and it has a lot to do with what will be presented in this chapter. Sensation is the reception of information. Light falls on the retina, sound moves hair cells in the ears and chemicals stimulate taste bud cells on the tongue. Then information is passed to the brain that then sorts and combines it. Perception is different. Perception is the evaluation and construction of the world out there based on but not limited to what comes in. It involves expectation, generalization and memory. If you think about movement for a bit, you'll realize that movement must be based on perception, not sensation. We don't have any specific movement sense, we have to construct movement from what we know and what we receive.

14.1 Why is Movement Important?

Movement is important to humans because there is a lot of information out there. In addition to constructing our world from memory and action and expectation, we have to select what matters from this dazzling array of information. Attend for a minute to the things you filter out—the sound of the clock, the feel of the chair on your bottom, the look of your jeans and t-shirt. You filter them out because their continued presence doesn't matter. Movement matters. It gives us cues for what's around us and where it is located. It helps us to see what is important. And it guides our own actions so we don't bump into chairs, coffee tables, and rocks. After all, we don't need to keep processing that the wall over there is the same wall it always was. But something moving may be opportunity—or trouble.

The third important thing about perception of motion is that we humans are visually dominant. Vision is a finely-tuned sense, with all the early processing in the retinas of our eyes. Visual information then goes to the brain, used in several brain regions, and also takes up much of the cortex with its combining, sorting and assessment. When there's a conflict between visual and other information, we believe our eyes. Movie makers use this dominance. Think of watching the characters on screen. We hear the voices coming from where the actors move their lips, though of course the sound comes from speakers on the wall. Watching a foreign film or a poorly dubbed one shows us how much we depend on this coordination and how much we believe the sound that matches the lip actions.

Because movement is so important, we often assume that we automatically and accurately know when we are moving and when we are not. This assumption is not necessarily true. I usually ask my Perception class to sit very still in their seats, then I ask them if they are moving. Their answer is usually something to the effect that of course they are breathing, but apart from that they are still. Then I point out that the earth is rotating at quite a speed (1000 mph), with which they agree. In addition, the earth is revolving around the sun (at 11,000 mph), and they agree with that too. So when we think we are still, we really aren't. The reasons we don't perceive all this as movement are two-fold. First, just as when you are on an aeroplane, we become habituated to constant movement. Second, if we and our surroundings move at the same speed and in the same direction, we don't see and don't compute movement. A plane could go fast enough to counter the earth's rotation and be "stationary". If you were in it, would you feel stopped? Probably not. So, given we don't always get "the truth", how do we know we are or aren't moving? Some of the ways we make mistakes will help us understand how this motion perception works.

14.2 Five Ways to Make Things Appear to Move That Don't

14.2.1 Movement after-effect

Without moving your eyes, stare at a pattern or relatively large part of the field of view that is moving in the same direction for a relatively long time. If, after about a minute, you look away or the object stops moving, you'll see a strange thing. Something that moved leftward will seem to back off to the right. Something that was expanding will now seem to shrink back. If you get a good dose of this biased exposure, you should be able to look at something else and see it move or change.

Why does this happen? First we have to look at what happens in the retina when an image moves across it. As the image hits a rod or cone receptor cell, it fires and passes information toward the brain. All the receptor cells pass information to the next step in the relay, the bipolar or amacrine cells. By the way they are interconnected, these cells change the information that passes to the brain. Some of them are hooked-up in such a way that they function as motion detectors, responding to motion in the left-to-right direction, say, but not to its opposite.

We have a lot of different motion detectors feeding into the brain, ones that signal

left-right, up-down, and toward-away, and report on optic flow. The toward detector is very important; perception specialists call it the "looming detector". The looming detectors inform us whether something is coming at us, maybe on a collision course with us, and suggest, duck! Even babies seem to have a primitive ability to detect looming. Anyway, all these detectors are firing at a low rate. This firing stays in balance and so we know there is no motion when nothing is moving.

When an object moves steadily in one direction across your retina, it stimulates the motion receptors in only that direction. After a while, they fatigue. They begin to fire at a lower level, still enough to tell you it's moving, but at less than the original rate. When the object stops moving, however, the balance is upset; the detectors in that direction are still fatigued and fire less, and the ones in the opposite direction fire more. What does you brain make of it? It perceives the imbalance in firing as movement in the opposite direction.

14.2.2 Autokinetic movement

Look at an isolated spot of light in a darkened room, or at a star in an area of the sky that has no surrounding stars. Soon it will appear to wander, not necessarily very far, and not always in the same direction. If you were to turn the room lights on and watch that same small light, you'd realize that it wasn't moving at all.

What causes this wandering movement? Surely the retinal motion detectors told the brain exactly what kind of movement across the retina was actually happening? Yes, they did. But the brain itself actually needs more than one local signal to know what's moving and what's not. The surrounding darkness means that there's no *frame of reference* in which to put that stationary light, and without it the brain is fooled. This effect is probably complicated by the eyes moving around, so the dot is moving on the retina too, and all in all it's confusing to the motion detector cells.

Because this is a situation of too little information, you can fool around with people's perceptions about whether the light is moving or not. "Watch, it's going to start to move right!" This effect was the basis for an old social psychology experiment by Sherif. He brought people into the lab first separately, and then in threes, told them the light was going to move right, and observed their reports on distance. When viewing alone, they varied from less than an inch to nearly seven. After they all discussed how far it went, they ended up agreeing on the distance as four inches. Even more interesting was a study that told participants that they would be able to see the light trace written words. Some of their subjects reported the names of these words, mostly simple ones. However, one subject stomped out of the experiment because he said the experimenters were reporting such personal information about him. Of course, it was all in his head.

14.2.3 Induced movement

To produce this kind of illusory movement, simply find or project a small dot of light onto a screen surrounded by a large frame. If the dot moves, or course you know it moved. But if the frame is moved, the dot seems to move in the opposite direction. This

only works if the dot and frame are all that you see. If they are in a larger background or there are other objects around, the effect disappears.

What causes this inaccurate assessment of movement? It's again perceptual, and sort of a cousin to autokinetic motion. All we have as a reference is the dot and frame, so even though the eyes correctly record what moved, it all goes to a *Frame of Reference* system in the parietal cortex of the brain. This system operates according the rule of thumb, little things move, large ones don't. People move, walls don't (if they do, run! it's probably an earthquake). Researchers have even recorded neurons in monkey brains that make this "decision".

It is easy to find real-world demonstrations of this rule. On a night with a moon and broken clouds, you can go outside and stare at a demonstration of induced movement. The moon seems to be racing through the clouds, and we can't stop seeing it that way, even though we know the clouds are actually moving much faster than the moon. Closer to home, just put a sticky dot on your television screen. Run a basketball game behind it. Get close to the screen if you can, as a bigger frame around the frame itself makes the effect much smaller. The result is that the dot seems to do a lot of moving, opposite the direction of play.

14.2.4 Afterimage movement

If you watch a sudden and intense light such as a flashbulb going off, you will probably still "see" the image of the flashbulb after the flash has stopped. If you look around, you will find that the image moves where you look. This effect can last quite a while, at least a minute. If you get a good afterimage from a long distance and then you focus closer, such as on your outstretched hand, you'll discover that the image also gets bigger.

What happens in this case is partly on the retina and partly in the brain. The flashed image is so intense that the retinal receptors do the opposite of what happened in section 14.2.1: they keep firing. This persistent firing is called after-discharge. So why does it appear to move? That part comes from the brain computation of eye movement. Our eyes move around when we look at the world. Of course the brain receives information of how we moved them, probably in the superior colliculus and the parietal cortex. So the brain knows that the eyes moved, the retinal information says that the image was stable on the retina, and the brain does simple subtraction and decides the image moved. You can get the opposite result simply by carefully pushing on your eyeball. "Image moved", reports the input from the eyes. "I didn't tell the eyes to move", says the brain. "OK", the brain decides, "the world must have jumped."

When I was a graduate student working on eye movement, I did an interesting experiment with afterimage movement. I wanted people to eye-track their unseen hand in the darkness. I attached a flashbulb so that it went off when the arm was in the same position during each cycle of back-and-forth movement. That built up an image of the hand on the retina, and people "attached" the unseen hand to the afterimage and tracked it fairly well. They also reported that their arm felt as if it were where the eyes saw it to be—visual dominance again.

14.2.5 Stroboscopic movement

If you have a series of lights set in an array in front of you and they go on one after the other, with just the right timing and spacing, you see instead one light, moving. This is a pretty powerful effect; you can see it at the *Fresh Express*, and there's a television commercial that does almost the same thing with changing faces. What is most interesting about this effect is that it is a transformation. Notice that the lights don't move, nothing moves. Instead we interpret the sequence of several stationary lights being illuminated in turn as the movement of one light. What the eyes see as steady the brain reports as a moving object.

This effect was one of the main phenomena that caused the Gestalt psychologists in the early part of the century to push the perceptual aspect of the sensation-perception combination. They argued that what is out there and what is reported by the brain aren't always the same; that we don't just receive information, but also construct images. They argued further that the whole is different from the sum of its parts. To this day, people studying perception still refer to the Gestalt approach. Of course we process what we see and build on it, but we also expect and construct the figures that fragmentary information suggests.

Notice that the whole of the television and movie industries is based on this quirk of our processing. A screen is just a very large number of pixels, changing colour across time. As we watch it, we construct in our brain the images of moving people and objects that these dot patterns suggest. This effect is even more true for movies, which are just a series of still shots. In the old days, technicians painstakingly hand-crafted frame after frame of just-a-little-different pictures that, when run one after another, were interpreted by watchers as flowing movement sequences.

14.3 Real Movement That Doesn't Move

This last example is the opposite of all the ones I have given about "unreal" movement. Simply stick your thumb out at arm's length and slowly make a wide arc in front of your face with your arm. Watch the thumbnail closely with both eyes, and you will be tracking the thumb. Keep watching it. You know the thumb moved. Yet think of it, the image of the thumb was stationary on your retina, and you didn't stimulate retinal motion detectors.

The explanation for this "constructed" movement is simple if you followed what happened to the illusory movement of the afterimage of the flashbulb. The brain computed both what happened to the image and what happened to the eyes. The image was stationary, the eyes moved, *ergo* the object moved. We use this effect a lot, and we probably don't need much of a glance to compute where the object whose image we pick up is going. Think of the ball being passed to you in a soccer game, the clay pigeon moving before you try trap-shooting at it, the sailboat crossing in front of you. We are pretty good at using even fragments of motion to track where things are and where they are going to be in space.

14.4 Why Aren't We Confused About Motion All the Time?

If I have given you a convincing argument that it's easy to fool you about motion, you may next ask why we aren't lost and baffled most of the time. How come we know what we are doing and where we are going? Part of the answer is that yes, our perception is limited. We don't know all that goes on in the world. Look at how much better visual acuity the eagle over the coulees has than us. Look at how well your cat gets around in the dark when you can't see anything. Or remember Dr. Kolb telling you how well your dog navigates and follows tracks by smell when you don't even know any animal was there.

Still, we do well in our niche. We use binoculars if we want to see anything far away. We aren't nocturnal, but we don't try to stumble around in the dark. More important, look back at the situations that produce illusory movement. We don't usually stare ahead as something moves across the retina to get the motion aftereffect. We usually have an excess of visual information so we don't have the chance to get either autokinesis or induced movement. The flashbulbs that give you afterimage movement are a relatively recent invention. When the image of an object appears near where a similar one was a few seconds ago, as in stroboscopic movement, it usually is the same object.

Illusions don't happen often because they only occur under special conditions. They are important to researchers in perception, though. There is a research rule that fooling the person or the system shows you how they work—and that's what we want to know. Illusions tell us about how we construct and receive our world, and remind us of the important lesson that's part of the description of the area, sensation and perception. Our sensory world, and thus all our world, is the result of top-down and bottom-up processes, what we receive and what we surmise, and what we take in and what we already know.

Chapter 15

Multiple Combinations of Co-factors Produce Variants of Age-related Cognitive Decline

Robert J. McDonald

Dementia is defined as cognitive impairments that are serious enough to disrupt normal daily living. The dementias include over 50 different disorders including vascular dementia, human prion diseases, Parkinson disease, Pick disease, frontal disorders, dementia with Lewy bodies, Huntington disease, Alzheimer disease, and others (Haase, 1977). Alzheimer disease is the most common cause of dementia accounting for over 65% of all forms of dementia if combined with vascular pathology. The prevalence of the disease in the general population is correlated with chronological age, with rare reports in the fourth decade and then logarithmic increases over subsequent decades (Katzman & Kawas, 1994).

The onset of Alzheimer disease (AD) is often first detected by the patients themselves or by family members, and is usually manifested by changes in personality. Following these subtle changes, anterograde memory impairments without any obvious sensory, motor, or motivational deficits are the defining feature of the early stages of this disorder (Hyman, Van Hoesen, Damasio, & Barnes, 1984). These memory deficits are usually seen on neuropsychological tests that are sensitive to medial temporal lobe damage (Grady et al., 1988). The medial temporal lobe includes various structures implicated in mammalian learning and memory processes including the hippocampus, amygdala, and parahippocampal cortex.

The hippocampus is a learning and memory system that is critical for the acquisition, encoding, and retrieval of representations of the various elements that define a specific event and the context in which it occurred (Sutherland & Rudy, 1989). The hippocampus in humans has been described as an episodic memory system essential

for conscious awareness of personal experiences. The amygdala has been implicated in tracking the biological significance of stimuli in the environment (Cador, Robbins, & Everitt, 1989; Kapp, Wilson, Pascoe, Supple, & Whalen, 1990). A significant portion of the parahippocampal cortex called the perirhinal cortex has been implicated in object memory (Mumby & Pinel, 1994). There is evidence to suggest that the brain damage and pathology associated with AD begins in the parahippocampal cortex and then to the hippocampus and amygdala (Hyman et al., 1984) and these changes correlate strongly with the specific memory impairments found in the early and mid-stages of AD. The transition between specific medial temporal lobe memory deficits to more global deficits correlates with the location and extent of brain degeneration. In the late stages of AD, brain damage and pathology spread beyond the medial temporal lobes into extensive portions of the cerebrum. As is seen in these stages of AD, memory impairments become more pronounced and more global deficits appear, including language, movement, and executive functions.

The neuropathologies associated with AD are the features that differentiate normal neural tissue from tissue of patients diagnosed with the disorder. These defining features are neurofibrillary tangles and senile plaques. These structures can be found in both cortical and subcortical areas although structures in the medial temporal lobes show the most dramatic accumulation of these abnormalities.

In Alzheimer's (1907/1987) original description of the brain of a patient with what would become recognized as Alzheimer disease, he noted that the internal structure of neurons was abnormal and called them neurofibrillary tangles. These abnormalities are called paired helical filaments and they consist of paired filaments that are wound around each other in a helical arrangement. The neurofibrillary tangles' primary constituent is tau protein which has 6 isoforms in the human brain (Heutink et al., 1997). The tau protein is an axon specific microtubule associated protein and seems to play a critical role in the construction of microtubules (Goedert, 1996). Neurofibrillary tangles found in the brains of AD patients have hyperphosphorylated tau protein, and it is thought that this process leads to a dissociation of portions of microtubules that then form into paired helical filaments in neurons (Clark, Trojanowski, & Lee, 1997). Several types of senile plaques also have been found in the brains of Alzheimer disease patients. The neuritic plaque consists of an amyloid core surrounded by a ribbon containing neuronal processes and glial cells that co-exist with paired helical filaments associated with neurofibrillary tangles.

15.1 The Model

This theory was formulated in response to the continued promotion of the idea that Alzheimer disease is caused predominantly by a single dominant factor. In general, there is no good evidence to suggest that this idea is true for the common non-familial version of the disorder, suggesting that a new and comprehensive approach is needed if we are truly committed to solving this complex, neuroscientific and social problem.

This theory proposes that Alzheimer disease and all forms of age-related cognitive decline that are ultimately linked to hippocampal damage and related structures are the consequence of interactions between various combinations of co-factors. These

co-factors include genes, neurotransmitter changes, vascular abnormalities, stress hormones, circadian rhythms, head trauma, seizures, and others not covered in this lecture.

The various co-factors can be roughly divided into two major categories. One set of factors, if present, can create a neural environment that makes neurons more susceptible to damaging agents or forces. I will call this group the *passive factors*. This group of factors includes senile plaques, neurofibrillary tangles, head trauma, cholinergic depletion, and prolonged glucocorticoid exposure. The other set of factors are damaging agents or forces that can produce neuronal death. I will call this group the *active factors*. These factors include seizure activity, vascular abnormalities, circadian dysfunction, and other unknown substances humans are exposed to in their particular environment.

The basic prediction of the model is that a combination of at least one factor from each category would be required to produce significant brain damage in the hippocampus, and associated memory impairments. A patient with more than one factor from the passive category would increase the speed and amount of cell death in the hippocampus simply because the cells are more susceptible to the damaging effects of the agents in the second category. A patient with multiple factors from both categories would have increased amounts of cell death in the hippocampus, at a faster rate, and severe memory impairments. Examples of different combinations of these factors and the outcome of these interactions will be discussed in the reiteration of the model in the final section of this lecture.

15.2 Co-factors

15.2.1 Genes

Alzheimer disease is the most common cause of dementia and is an etiologically complex and heterogeneous disorder with gene mutations and gene isoforms that are considered to have roles in this disease (Pericak-Vance & Haines, 1995). Research over the last decade suggests that 3 gene mutations on different chromosomes cause a rare early onset form of Alzheimer disease that occurs when patients are 65 years of age or younger. These gene mutations include the amyloid precursor protein (APP) gene found on chromosome 21, the presenilin 1 (PS1) gene found on chromosome 14, and the presenilin 2 (PS2) gene found on chromosome 1.

15.2.2 Vascular abnormalities

Abnormalities in the cerebrovascular system also appear to be a co-factor in the onset and intensity of AD (Snowdon et al., 1997). The discovery of a relationship between the APOE gene found on chromosome 19 and late-onset AD provides further support for the link between this factor and the disorder because the APOE protein is found in the major neuropathological hallmarks of AD (plaques and tangles) and in the vasculature and cerebral vasculature of these patients (Premkumar, Cohen, Hedera, Friedland, & Kalaria, 1996). The amyloid deposits found in cerebral blood vessels act to obstruct

blood flow and can interfere with normal blood circulation to the brain (de la Torre, 1994).

15.2.3 Neurotransmitters

The discovery that patients with Alzheimer disease show abnormalities in the cholinergic neurotransmitter system gave rise to a popular theory suggesting that this type of dementia is caused by cholinergic dysfunction (Bartus, Dean, Beer, & Lippa, 1982; Coyle, Price, & DeLong, 1983). Alzheimer patients show significant neuron loss in the nucleus basalis of Meynert (NBM), neurons that project to widespread cortical areas (Perry, Perry, Blessed, & Tomlinson, 1977). The targets of these neurons are not equally affected in AD with the temporal lobes showing the most loss of cholinergic markers including various precursors and enzymes.

15.2.4 Stress hormones

It is generally agreed that there is significant variability in the incidence of age-related cognitive decline within a group of chronologically identical individuals. Usually, a group of aged individuals will separate into two sub-groups with one group showing normal cognitive abilities and the other showing obvious impairments (Gallagher & Pellymounter, 1988). This discrepancy in age-related memory dysfunction suggests that some factor or co-factors are responsible for good versus poor cognitive aging. One factor associated with poor cognitive aging is elevated glucocorticoid levels. For example, experimentally induced elevations of glucocorticoid levels in rats have been associated with neuropathology and cell death in the hippocampus (Landfield, Baskin, & Pitler, 1981) and impairments in hippocampal-based memory processes (Landfield, Waymire, & Lynch, 1978).

There also seems to be a relationship between elevated glucocorticoid levels and AD found in humans. For example, patients diagnosed with AD show higher blood cortisol levels (Davis et al., 1986), which has been correlated with hippocampal atrophy (Deleon et al., 1993), and also show alterations in the HPA axis in the early stages of the disorder (Raskind, Peskind, Rivard, Veith, & Barnes, 1982). Further support for the idea that there is a strong relationship between AD and HPA axis integrity comes from a longitudinal study of healthy, aged subjects in which yearly assessment of cortisol levels were correlated with cognitive function (Lupien et al., 1997). Another study showed a strong relationship between increases in cortisol during aging and hippocampal shrinkage and related hippocampal-based memory deficits (Lupien et al., 1998).

15.2.5 Circadian rhythms

Temporal organization of physiology and behaviour in mammals is provided by circadian rhythms that are generated in the suprachiasmatic nucleus (SCN) of the anterior hypothalamus. A subsection of aging research has shown an interesting relationship between chronological age and circadian rhythm deterioration in human and non-human animals. This age-related deterioration of circadian rhythms is expressed in fragmented and reduced patterns of activity.

The effect of aging on circadian function and sleep patterns is intriguing because the disruption of paradoxical sleep or slow wave sleep (C. Smith, 1985) disrupts hippocampal-based memory processes in human and non-human animals. Patients with various forms of dementia, including AD, show increased incidences of circadian dysfunction (Swaab, Fliers, & Partiman, 1985; Van Someren et al., 1996), sleep disorders, and related hippocampal-based memory impairments.

15.2.6 Head trauma

Various studies have shown a significant relationship between the occurrence of head trauma and incidence rates of AD(Mortimer, French, Hutton, & Schuman, 1985). A review of 11 case studies suggests a significant contribution of head injury experience with the likelihood of getting AD (Mortimer et al., 1991). Further support for a relationship between head trauma and AD comes from research investigating boxers, who, following significant amounts of head trauma over their careers, also show neurofibrillary tangles and senile plaques (Corsellis, Bruton, & Freeman-Browne, 1983), which is the neuropathology similar to that found in the brains of AD patients.

15.2.7 Seizure activity

The effects of seizure activity on the brain and memory are similar in human and non-human animals. In humans, it is believed that mild epilepsy is not associated with memory impairments unless it is accompanied by another brain disorder, such as AD (Esiri, Wilcock, & Morris, 1997). However, severe epilepsy can lead to cognitive deficits that are related to hippocampal damage (Bruton, 1988). Extended inducement of seizure activity leads to neuropathology and cell loss in the hippocampus of other mammals as well.

15.3 The Model Revisted

In this section, I will revisit the model presented in the introduction of this lecture and restate that there are various forms of AD and age-related cognitive decline caused by different combinations of co-factors. The common thread linking all of these different disorders is that memory systems in the medial temporal lobe, particularly the hippocampus, are damaged in all cases. One reason why the hippocampus and related structures are more susceptible to these factors compared to other brain regions is that some of the pathologies associated with AD seem to preferentially accumulate there (plaques and/or tangles), the foci of seizure activity are located within medial temporal lobe structures, and specific cell groups in the hippocampus are particularly sensitive to changes in neuronal environmental conditions.

One of the general assumptions of the current model is that an individual with only one factor would show low to moderate memory impairments. Individuals with two or more of these co-factors would get a severe form of dementia, depending on the combinations. A combination of factors from both the passive and active categories should

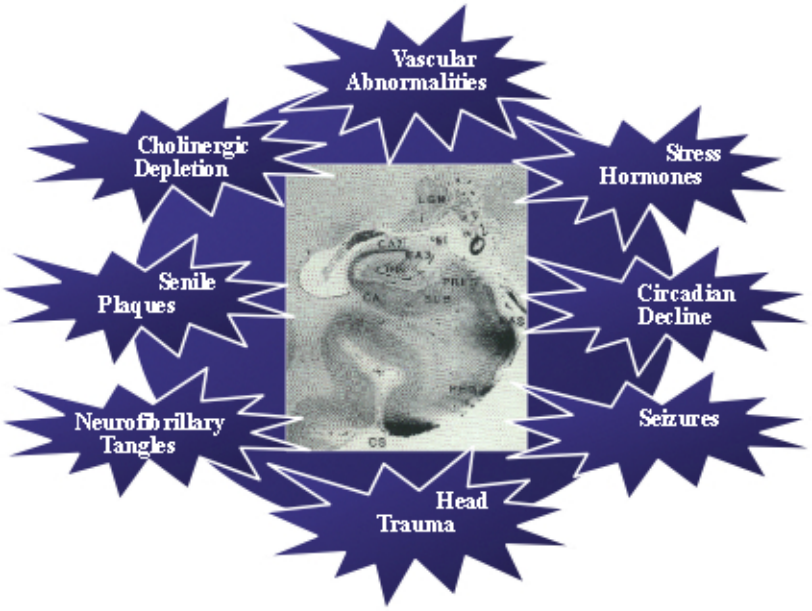


Figure 15.1: A hypothetical example of the etiology of a case of age-related cognitive decline. This is the worst case scenario, according to the current model, for an aged individual in which the cause of the disorder is hippocampal exposure to a neural environment containing all of the co-factors that can cause damage to this memory system. This patient would show relatively early and severe memory impairments.

produce the most severe cases of medial temporal lobe damage and associated memory impairments. For illustrative purposes, various hypothetical cases of age-related cognitive decline will be described and discussed in lecture. This is not an exhaustive description of all of the possible permutations but rather represents a selection of some of the more interesting combinations and their outcomes. After introducing each case, a brief indication of the onset (relative to all cases of late-onset AD) and severity of memory impairments will be provided.

Figure 15.1 shows all of the co-factors reviewed in the present paper, and represents the worst case scenario for an aged individual. This subject would have all of the co-factors that could act on the integrity of the hippocampus. The passive factors produce an environment that make neurons more susceptible to damaging agents or forces (plaques, tangles, head trauma, cholinergic depletion, prolonged glucocorticoid exposure) and the active factors could produce neuronal death, particularly when these cells are more vulnerable (seizures, vascular abnormalities, circadian dysfunction). This patient would show early onset and severe memory impairments associated with medial temporal lobe structures.

15.4 Implications

The main implications of this model are threefold. First, patients with severe memory impairments that are virtually identical to AD would be characterized as having the disorder even when an autopsy reveals no plaques or tangles in their brain. According to the present model, this would reduce a significant number of false negatives. Second, researchers, clinicians, and doctors interested in the reduction and prevention of age-related cognitive decline should be open to the development and use of prevention and treatment regimes directed at various combinations of co-factors instead of hoping for a magic bullet in the form of a pill that may never materialize. Finally, future research on the etiology of age-related cognitive decline should continue to be directed at providing a deep understanding of each co-factor implicated in these disorders, but should also make a concerted effort towards understanding interactions between these factors.

Chapter 16

Interactive Memory Systems
Robert J. McDonald

> Material things are there by means of their images: knowledge is there of itself; emotions are there in the form of ideas or impressions of some kind, for the memory retains them even while the mind does not experience them, although whatever is in the memory must also be in the mind. My mind has the freedom of them all. I can glide from one to the other. I can probe deep into them and never find the end of them. This is the power of memory! This is the great force of life in living man, mortal though he is!
>
> – St. Augustine

In his book, *Confessions*, St. Augustine captures many critical aspects of our memory at a time in which little or nothing was known about this complex brain process. The opening epigraph suggests that our memory is:

1. multifaceted and not unitary
2. the repository for your own past and identity
3. the most important biological force in the human experience

St. Augustine goes even further and suggests that our memory allows us to change our behaviour because memory contains a record of our past history and can be replayed and analysed; it is the only force through which we can grow and change as individuals.

The latter point is critical for this lecture because it is my assertion that the organization of memory in the mammalian brain and the neural systems that mediate multiple kinds of memory must play a pivotal role in our thoughts, emotions, choices, actions and even our personalities. Furthermore, these complex neural circuits in our brain not only contain remnants of our past that are the basis of personal identity but also exert an enormous influence on individual behaviour.

The first section of this lecture will introduce a simple but powerful theory about the organization of learning and memory processes called Interactive Memory Systems Theory (IMST). The second section will briefly discuss a theory suggesting that normal and abnormal manifestations of behaviour are determined, to a large extent, by some complex set of interactions between an individuals genetic make-up, developmental events during pre and post-natal development, and accumulated experience through life. The relationship between the two theories will also be introduced and then, in the final section, I will provide evidence that the etiology of many of the major psychiatric disorders may be linked to alterations in the integrity of various memory systems using drug addiction as an example.

16.1 Interactive Memory Systems Theory (IMST)

The foundation of modern views of the organization of memory in the mammal was built on the influential work of Pavlov, Hull, and Tolman. Briefly, each of these scientists formulated a general theory of learning and memory in which these functions were mediated by a basic, underlying mechanism. The proposed mechanisms included classical conditioning for Pavlov, reinforced stimulus-response (S-R) learning in Hull's theory, and the flexible cognitive mapping view for Tolman. Despite significant acrimony between supporters of these different positions, it now appears that all of these theorists were correct in that the mammalian brain uses all of them, as well as other types of learning mechanisms that appear to be mediated by different brain circuits.

The first direct evidence for the idea that there were multiple memory systems in the mammalian brain came from Scoville and Milner's (1957) discovery that patients with damage to the medial temporal lobe showed impairments in some types of learning and memory function, but were normal in other aspects. Milner concluded from this data set that structures in the medial temporal lobe, most likely the hippocampus, were involved in complex memory processes, and that brain structures anatomically and functionally independent of the medial temporal lobe mediated other learning and memory function.

Most of the influential multiple memory theories of mammalian brain function were formulated during the 1970's and were inspired by Milner's findings. Many of these theories are dual memory formulations in which the hippocampus is the central module, while some other brain area(s), independent of the hippocampus, mediates non-cognitive S-R habit learning and memory function (Gaffan, 1974; Hirsh, 1974; O'Keefe & Nadel, 1978; Tulving, 1972). Hirsh and Krajden (1982) were the first to provide considerable detail on the different kinds of interactions that could theoretically occur between cognitive- and habit-based memory systems. A summary of this view is captured in the following quotation: "When two different systems appearing to address the same substantive matters are present, it is worthwhile to ponder how they might interact. We think that on some occasions the two systems compete; on others they cooperate. Once the fundamental differences between the two systems are understood, their differing capacities become clear. Each has capabilities that the other does not. There are certain features of knowledge that cannot be attained without using the capacities of both."

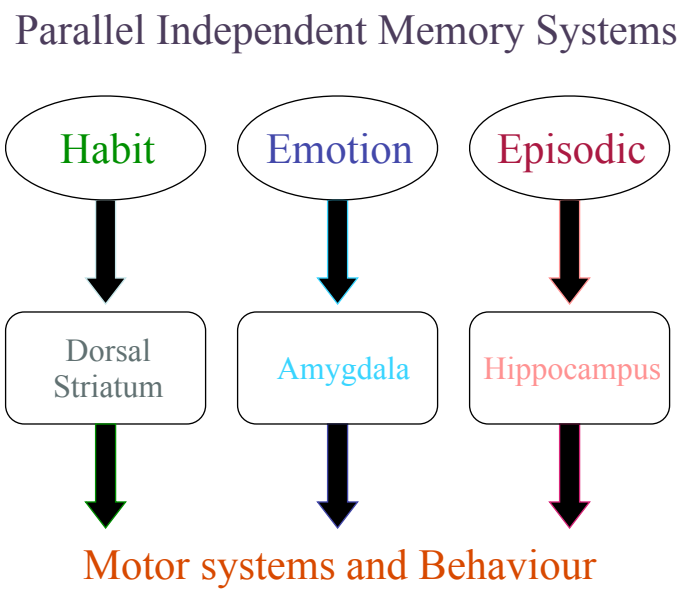


Figure 16.1: Parallel learning and memory systems in the mammalian brain.

Thus, in the majority of situations both systems are processing information in parallel and it is the circumstances or details of a particular situation (e.g., the performance requirements of a task) that determine whether systems interact competitively or cooperatively.

While this work was ongoing, a parallel line of research was accumulating a significant body of evidence suggesting that the dorsal striatum, cerebellum, and the amygdala were also learning and memory systems. The combination of innovative dual memory theories and evidence of anatomically distinct learning and memory systems provided a fertile research context in which various pairs of double dissociations were demonstrated.

This theory of multiple, parallel, memory systems suggests that the mammalian brain has at least 3 major learning and memory systems. Each system consists of a "central structure" and a set of interconnected neural structures. The "central structures" of these different circuits include the hippocampus, amygdala, and dorsal striatum (see Figure 16.1).

These memory systems:

- acquire information simultaneously and in parallel and are always on-line.
- have access to similar information during events, but each system is specifically designed to represent different relationships among the elements of a learning situation. These elements include stimuli, internal and external responses, and reinforcers.
- processing style is determined by the intrinsic organization of the system and the

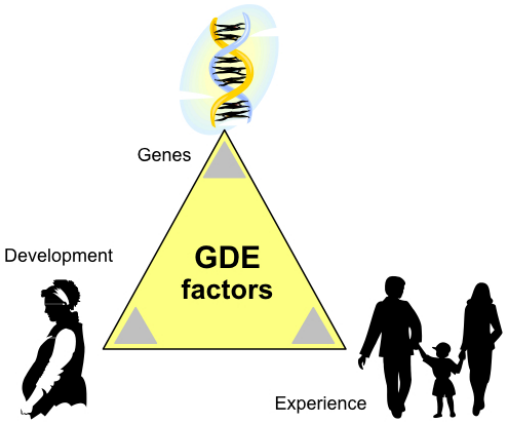


Figure 16.2: Normal and abnormal manifestations of behaviour are determined, to a large extent, by some complex set of interactions between and individual's: genetic make-up, pre- and post-natal developmental events, and accumulated experience through life.

input/output relations to the rest of the brain.

- process information independently, however, these systems can interact cooperatively or competitively to produce or influence ongoing or future behaviour.

Among the three central memory system structures for which the hippocampus is thought to be critical, is the formation of episodic memories, in which a complex representation consisting of the various elements of a situation or event is constructed (Sutherland & Rudy, 1989; Tulving, 1972). The amygdala has been implicated in the formation and storage of emotional memories (Bagshaw & Benzies, 1968). These emotional memories uniquely encode the subjective valence of the experience (positive or negative). The dorsal striatum has been implicated in stimulus-response habit learning and memory processes (Packard, Hirsh, & White, 1989). This kind of learning occurs when the subject engages in repetitive behaviours. For example, the voluntary behaviours by the driver elicited while driving a car on a repeatedly travelled route are thought to come under the control of the habit system.

16.2 Who am I?

The second brain theory that will be explored in this lecture suggests that normal and abnormal manifestations of behaviour are determined, to a large extent, by some complex set of interactions between an individual's genetic make-up, developmental events during pre and post-natal time periods, and accumulated experience throughout the lifespan (see Figure 16.2). All of these factors can have major effects on the organization of the brain. Alterations in the organization of the brain could affect overall relationships between each learning and memory system, as well as the relationships

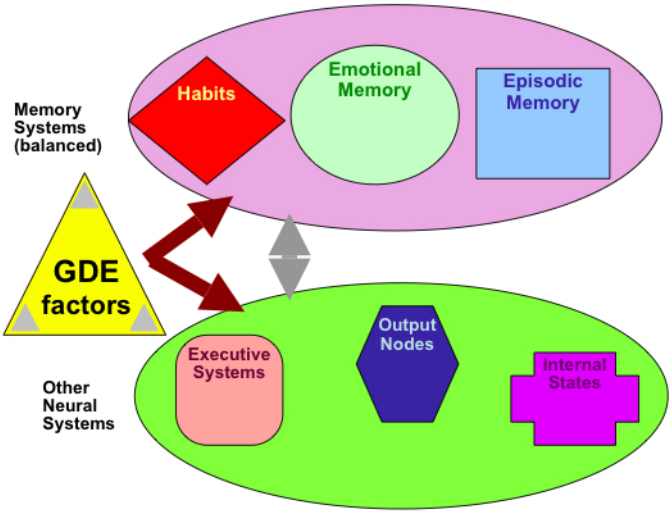


Figure 16.3: A hypothetical outcome of complex interactions between genes, development, and experience (GDE) factors on the organization of learning and memory systems in the mammalian brain.

of these systems with the rest of the brain. For the purpose of the present discussion the combination of factors will be referred to as the GDE (genes, development, experience).

Within the normal range of variability, alterations in the balance between these memory/behavioural systems can lead to individual personality, affective style, choices, actions and certain strengths and weaknesses associated with different tasks or situations (e.g., mathematics, athletics, music, social interactions, etc.). Figure 16.3 shows a hypothetical outcome of complex interactions between GDE factors. One important effect of these factors is on the intrinsic organization of various memory/behavioural systems with each other and other neural systems. This example represents a normal individual in which there is a balance between these systems that when activated result in relatively normal patterns of behaviour in a wide range of situations.

A logical extension of this view is that there can also be changes in the balance of these systems that lead to abnormal manifestations of behaviour including major psychiatric disorders such as schizophrenia, drug abuse, and mood disorders. Also, the way memory systems interact can produce “abnormal” talents, as in the case of savants and others who display unusual abilities (see Luria, 1968; Sacks, 1970) that may lead to great accomplishments.

16.3 Psychiatric Disorders: the Central Role of Interacting Memory Systems

In the scientific literature, there is an emerging focus on the idea that the etiology of almost all major psychiatric disorders are linked to abnormalities in brain areas implicated in learning and memory processes. The evidence suggests that dramatic changes in the relationship of these systems to one another, and with other brain systems, lead to abnormal manifestations of behaviour. These cognitive and behavioural abnormalities include schizophrenia, anxiety, depression, and drug abuse.

Historically, ideas about the etiology of these major psychiatric disorders have been dominated by single factor theories. The idea was that these complex brain disorders were caused by alterations in a neurotransmitter system, gene or some other single factor. There are many single factor theories of brain disorders that continue to dominate important areas of research. Examples of such theories include the idea that schizophrenia is caused by an over-activation of the neurotransmitter dopamine, or that Alzheimer disease, in which alterations in a small number of genes leads to overexpression of beta-amyloid peptide or neurofibrillary tangles, is associated with neuronal damage and cognitive deficits. Although single factor theories have generated a significant amount of important research, they do not accurately account for the complex etiologies of these disorders.

In the final portion of this lecture I will discuss the etiology of drug abuse and the possible role of different learning and memory systems.

16.4 Drug Abuse

Various theories of the neural mechanisms of drug abuse have been put forward. One popular view is that addicts administer drugs for their reward or pleasure-inducing effects. This is sometimes referred to as the hedonic theory of drug addiction. Another theory focuses on various learning and memory systems in which the normal functions of these complex neural circuits become subverted leading to compulsive drug seeking behaviours (Everitt, Dickinson, & Robbins, 2001). In this model, drugs of abuse initiate plasticity mechanisms in different learning and memory systems that come to control behaviour of the individual over other, pre-existing memories. In this model, the amygdala acquires information that promotes approach and interaction with drug associated stimuli. The dorsal striatum promotes the acquisition of stimulus-response (S-R) habits and the hippocampus acquires information about the context in which drug stimuli are obtained.

16.5 Drug Abuse and Interacting Memory Systems Theory (IMST)

The IMST view of the organization of memory/behavioural systems in the mammalian brain might be a powerful way of understanding the neural basis of drug addiction. The

various theories mentioned earlier indicate the critical role different learning and memory systems play in drug addiction in which powerful plasticity processes and associated memories are formed during drug experiences that come to dominate behavioural control.

Another window on the mechanisms of drug addiction that the IMST might open is an explanation for individual differences in susceptibility to drug addiction (Glantz & Pickens, 1992). We have previously argued that although these learning and memory systems affect behaviour, there are GDE factors that alter the relationships between these systems and the relationship of these systems to other brain areas. It is believed that normal and abnormal manifestations of behaviour, like drug addiction, are determined, to a large extent, by some complex set of interactions between these factors that can have a major effect on the organization of the brain. Thus, interactions between GDE factors can affect neurobiological integrity and impart an organizational change in the relationship of these memory/behavioural systems to one another, and with other brain systems that could make an individual more susceptible to drug addiction.

One possible neural change that could mediate addictive behaviours is via enhanced behavioural control exhibited by one memory/behavioural system. For example, various GDE factors interact to enhance the dominance of the S-R habit system that would lead to an increased tendency towards habitual control over behaviour. This could be influenced by the ease of access to common output sites or via enhanced plasticity processes associated with cognitive processes linked to addiction.

A second possibility is that, in particular individuals, alterations in brain organization caused by GDE factors can lead to a bigger reward signal occurring when drugs of abuse are administered. This increased signal could result in an acceleration of specific types of learning and memory processes associated with compulsive drug seeking, which, in turn, could result in an acceleration of learning and memory processes dependent on these signals and ultimately behavioural control by these systems.

A final example is based on the idea that various GDE factors could lead to an organizational change in the brain resulting in a reduction of inhibitory control via prefrontal cortical mechanisms (Kolb, 1990). This reduction in perfrontal inhibitory control could result in increased behavioural control by memory/behavioural systems that require the contribution of executive systems for appropriate choice behaviours (Fuster, 1989; Moscovitch, 1994).

16.6 Summary

This current review and analysis puts forth the idea that the organization of memory in the mammalian brain and the neural systems that mediate them must play a pivotal role in our thoughts, emotions, choices, actions and even our personalities. According to this view, complex interactions between neural circuits that contain remnants of an individual's past experience not only provide the basis of your identity but also exert an enormous influence over ongoing behaviour. Interactions between these systems and related brain areas are thought to determine who we are and how we behave in particular situations. An extension of this idea is that abnormal manifestations of behaviour are caused, to a large extent, by alterations in the relationships among different

memory/behavioural systems and other brain areas.

Symptoms associated with various psychiatric disorders are hypothesized to be caused by complex interactions between patients' genetic background, their pre- and post-natal development, and their life experiences. All of these factors can have major effects on the organization of the brain and even subtle alterations could affect the overall relationship between memory/behavioural systems (balance) as well as the relationships among memory/behavioural systems and the rest of the brain. In many instances, it is this relationship between GDE factors and interactive memory systems that ultimately determines manifestations of normal and abnormal behaviour.

Chapter 17

Repairing the Injured Spinal Cord

Gerlinde A. Metz

When Christopher Reeve played his famous role as Superman in the 1980's, he would not have expected that years later his career would change from an actor to an activist who would fight for the chance to walk again. In May 1995, Christopher Reeve participated in an equestrian competition in which his horse stopped suddenly. Reeve was thrown forward, head-first, over the horse. As he hit the ground, the top two vertebrae of his neck shattered and damaged his spinal cord. This accident left him seriously injured and unable to move or feel anything below his neck. When rushed to a hospital, assessments revealed that the injury caused neurons to die and left a fluid-filled cyst in the center of the spinal cord. "When Mr. Reeve was brought to the hospital, we predicted that he would never be able to move his arms or legs again", one of his doctors said. Reeve was not only paralyzed from the neck down and confined to a wheelchair, but he was unable to breathe without a respirator that steadily pumped air into his lungs.

Soon after his accident, Christopher Reeve began participating in various rehabilitation programs. He regularly performed exercise procedures that included training on a Functional Electrical Stimulation bicycle, treadmill therapy, and spontaneous breathing training. About four years later, Reeve found he could voluntarily move his index finger and continued to recover some other movements, which came back in small stages and strengthened over time.

In an attempt to improve Reeve's weak voice and breathing, doctors implanted small electrodes in his chest in an experimental procedure that they hoped would enable the paralyzed movie star to breathe without a respirator. In 2002, the respirator was turned off and "all you could hear was me breathing through my nose", Reeve said. After that, Reeve was able to wake up one morning and sense the smell of fresh coffee for the first time in eight years. He became able to breathe independently for half an

hour without the respirator while hoping to also be able to strengthen his voice.

Reeve founded the Christopher Reeve Paralysis Foundation, a non-profit organization that supports research to develop effective treatments and a cure for paralysis caused by spinal cord injury and other central nervous system disorders. Reeve dedicated a portion of the funds toward stem cell research that might one day help in replacing spinal neurons lost by an injury.

Reeve's case illustrates the sudden and disturbing impact that a spinal cord injury has on an individual's life. But it also demonstrates that with continuous therapy and training recovery might become possible even years after the injury. The belief had been that most recovery occurs in the first six months after an injury, and that there is no possibility of further change by two years after the damage. Christopher Reeve showed that there is new hope for patients with spinal cord injury by using new state-of-the-art therapies.

17.1 The Spinal Cord and Its Connections

The central nervous system includes the brain and spinal cord. The spinal cord is the major bundle of neurons that carry information to and from the brain to the rest of the body. It allows the brain to send information to the muscles of the body that result in the ability to move, walk, or breath. In turn, the spinal cord gathers information from the arms, legs, chest and back and sends it to the brain. The spinal cord is divided into segments each accessing a specific section of the body.

With regard to the extensive connections between the spinal cord and the brain, a local injury to the spinal cord can lead to extensive functional impairments. The initial reason for seeking help after an accident is usually the inability to move the legs (paralysis) and numbness in the limbs. The loss of function (due to the loss of the brain's descending access to the limbs) or loss of sensation (due to the interruption of the ascending connections from parts of the body to the brain) are often permanent. The loss of motor and sensory function is usually accompanied by many other symptoms depending on the site of injury. A lesion interrupts the communication between parts of the body below the level of injury and the brain. Lesions at a lower spinal level affect leg movement and sensation and can also cause autonomic dysfunction of the inner organs including a reduced ability to control the bladder and bowel, sexual dysfunction, and exaggerated reflexes. Lesions at a higher level of the spinal cord can additionally cause decline of voluntary breathing and speech, and disturbed arm movement.

17.2 Causes of Spinal Cord Injury

There are many causes of spinal cord injury, including demyelinating disorders, tumors, infections, vitamin deficiencies, or displacement of soft cartilage plates that are located between the vertebrae, the intervertebral discs. The most prominent cause of spinal cord damage, however, is traumatic injuries induced by accidents. Although the spinal cord is embedded in the cerebrospinal fluid and protected by the surrounding solid bone structures formed by the vertebrae, a strong impact might crush the ver-

tebrae and compress the delicate spinal cord tissue (contusion injury) or penetrate it with pieces of broken bone (transection). It is noteworthy that more than half of the individuals suffering from such traumatic spinal cord injury are under 30 years of age. The highest incidence of accidents occurs during the summer months due to outdoor activities. In the following, we will mainly focus on traumatic spinal cord injury, and we will explore some of the latest findings in the treatment of spinal cord damage.

17.3 The History of Spinal Cord Repair

The characteristic and devastating symptoms of spinal cord injury were first recognized by the ancient Egyptians. The discovery of a papyrus depicting the main features of traumatic spinal cord injury dates to 4500 years ago. This surgical papyrus, named after its discoverer, Edwin Smith, is considered the first medical document in the history of humankind: it contains the description of 48 cases as written by an Egyptian surgeon. One of those cases is that of a spinal cord injured patient. Notably, the description of spinal cord injury concludes with the words "... a disease one cannot treat".

The first suggestions on how to treat spinal cord injury were made by the Greek physician Hippocrates. Hippocrates was a pioneer of modern medical methods and believed that, with nutritional support, the body is able to cure itself. According to his statement "let the food be your medicine and medicine be your food", he recommended a high fluid intake and a special diet to reduce bladder and digestive dysfunctions. Nevertheless, although these therapies provided some symptomatic relief, they still were not able to extend the life span of the patients. It was common that most patients died within the first two weeks after injury, and a patient rarely was able to survive up to two years.

It was 400 years ago that more invasive treatment approaches were investigated, including the first surgical approaches. Significant progress, however, in treating spinal cord injury was made not earlier than during World War II when thousands of spinal cord injured patients were brought to the hospitals. This large influx of patients led to systematic assessment of general medical care options and surgical approaches, and physiotherapeutic and rehabilitation programs were developed. Over the following years, combinations of different therapies have been explored in order to prolong the life expectancy of patients and improve their quality of life.

17.4 Recovery from Spinal Cord Trauma

Critical factors in the recuperation period after spinal cord injury are the level and extent of injury. The level of injury determines which extremities and organs will be affected. Furthermore, the quantity of the local spinal tissue damage is important. If the lesion was incomplete, there is still some functioning below the primary level of the injury and the remaining tissue creates a bridge for the exchange of information between the brain and lower spine. Even a small tissue bridge can provide an important substrate for the nervous system to compensate for the functional loss to at least some degree. A person with an incomplete injury may be able to move one limb more than

another, may be able to feel parts of the body that cannot be moved, or may have more functioning on one side of the body than the other. In this case, rehabilitative training might play a role to help the central nervous system compensate in a functionally meaningful way.

In the worst-case scenario, the spinal cord lesion is complete and leaves no intact tissue bridge to mediate functional recovery. The number of therapies available to improve the quality of life in patients with this condition is very limited, however, standard methods such as physiotherapy can be used to maintain the viability of muscle tissue and strength. In general, only a few treatments have been internationally approved for the treatment of spinal cord injury, but the latest research provides promising new findings.

17.5 Treatment Strategies for Spinal Cord Injury

Despite the long history of knowledge about spinal cord injury, the inability to recover from the consequences have been accepted as a law of nature. So far, the three major goals of spinal cord injury treatment have been (i) to minimize further injury to the spinal cord, (ii) to prescribe rehabilitative therapy so the patient can function to the best possible ability, and (iii) to deal with possible long-term complications. Nevertheless, the surprising recovery process of Christopher Reeve and other patients has set new standards for this traditional view of spinal cord treatment, and the recent progress in spinal cord research has provided some intriguing new technologies for therapeutic strategies.

17.5.1 Protecting the spinal cord

A mechanical lesion of nervous tissue leads to destruction of neurons and immediately triggers an immune response that leads to infiltration of the damaged tissue by immune cells. Although the immune response supports the process of local wound healing, it also has detrimental consequences by causing swelling of the tissue and exaggerating the tissue loss. At this early stage after injury anti-inflammatory treatments come into play. Anti-inflammatory substances suppress the immune response and therefore reduce the tissue damage caused by the inflammatory response. The number of neurons dying early after an injury can thus be reduced (Tator, 1972).

The anti-inflammatory drug Methylprednisolone is currently the only standard drug treatment in North America and Europe. The drawback of anti-inflammatory treatment is that the therapy needs to be initiated as early as possible after injury. Mehtylprednisolone, for instance, has been proven beneficial only if given within the first three to eight hours after injury (Bracken, 2002). Such time intervals for treatment often are impracticable as spinal cord injuries might happen in locations that are difficult to access by paramedics; for example, when climbing in the mountains. Therefore, alternative treatments that are effective at a chronic stage after the injury need to be investigated.

17.5.2 Teaching the spinal cord to move the body

Until about 20 years ago, it was believed that the adult central nervous system was incapable of plastic rearrangements. As a consequence, rehabilitation programs mainly focused on the immediate post-injury phase to maximize existing function. More recently, scientists started to believe that by encouraging patients to use their injured body parts, functions would be able to improve even years after the injury occurred. This has become the focus of rehabilitation and physiotherapy programs performed at specialized centers (Dietz & Colombo, 2004).

The aim of physiotherapy is to assist and promote the recovery process by physical means, such as massage and exercise. Repetitive exercises of arm or leg movements are usually part of comprehensive programs that consider physiological, psychological, and economic aspects of spinal cord injury rehabilitation and fitness. These programs teach the patients to use alternative movement strategies to overcome the disability. For instance, patients often need to re-learn daily activities, such as dressing themselves, buttoning a shirt, or brushing their teeth. Other common training programs involve walking on a treadmill. The treadmill allows adjustments to speed, so that patients can be trained according to their individual capacity to perform stepping movements. In patients who don't have sufficient muscle strength to support their own body weight when performing steps on a treadmill, the person can be suspended in a harness. Even after complete spinal cord injuries that restrict the capabilities to acquire new motor strategies, regular exercise helps to maintain and enhance mobility and muscle strength.

A revolutionary technology has been introduced with the Locomat, a robotic device that supports walking movements on a treadmill (Wirz et al., 2005). The patient's legs are strapped to a machine that delivers power to the hip and knee joints. A groundbreaking new computer system controls the speed and extension of the movements and adapts them to a patient's own ability to generate stepping movements. This novel technology is expected to help patients redevelop functional walking patterns.

As we learned from Christopher Reeve, it is crucial for the success of any rehabilitation program that the program is continued over long periods of time. Although the training process is often quite demanding for patients, recent studies performed in Canada indicated that actively participating in training programs only twice a week can significantly improve both long-term physical and psychological well-being of patients (Hicks et al., 2003).

17.5.3 Rewiring the injured spinal cord

Although a spinal cord injury interrupts communication between the brain and the muscles, the nerves and muscles below the site of injury remain alive and intact. A new cutting-edge technology involves using a small electrical current to stimulate the muscles that no longer receive signals from the brain, causing an otherwise paralyzed muscle to contract. This technique, so-called Functional Electrical Stimulation (FES), sends electrical currents by means of a computer system to the peripheral nerves that control the individual muscles (Scott, Peckham, & Keith, 1995). The currents are delivered by small electrodes that are implanted under the skin. The electrodes can stimulate muscles to perform even complex movements such as grasping an object

with the hand or executing a walking pattern. In turn, feedback about the performed movement is provided by a variety of sensors so that any movement can be precisely adjusted. Thereby, the FES system remodels the lost function of the spinal cord by relaying motor commands and sensory feedback.

At the moment, FES is still a research project and not integrated in a larger clinical program. FES has been successfully used in a small number of highly motivated patients, including Christopher Reeve, who have volunteered to participate in the program.

17.5.4 Regrowing the spinal cord

There is one central question that still remains to be discussed: Why does the spinal cord fail to regrow and repair itself after an injury? Why do functional disturbances after spinal lesions show only minor improvements and usually remain for the rest of the patient's life? The major problem is the fact that the capacity of the nervous tissue to regrow after injury is very limited. There are a number of reasons that account for the failure of regrowth of nervous tissue in adults. One of the key issues is the presence of growth-inhibitory proteins in the environment of the neurons(Schwab & Bartholdi, 1996). The function of the nervous system depends on a highly specific pattern of connectivity formed between neurons during development. To stabilize the neuronal network in the mature central nervous system, the environment of the neurons produces inhibitory proteins that maintain the connections once established. This benefit becomes a drawback once there is damage, and the trade-off is the loss of the ability to regenerate. The presence of growth-inhibitory molecules then make it difficult to restore the damaged connections. An exciting treatment approach that currently is still in the experimental stage is to deliver neutralizing proteins directly into the damaged area to block the inhibitory property of the neuronal environment. In animal experiments, this treatment allowed remarkable neuronal regrowth in the injured spinal cord. Detailed investigations also revealed that damaged neurons were able to regenerate over long distances, and that the neuronal growth contributed to substantial functional recovery.

Another approach that aims to rewire the injured spinal cord is stem cell therapy. Stem cells are undifferentiated cells able to renew themselves and can divide into many types of specialized cells, including neurons and support cells. The basic idea is that implantation of stem cells into the damaged area of the spinal cord might help to replace the lost tissue (Brodhun, Bauer, & Patt, 2004). While stem cells can be cultured under laboratory conditions, the use of human stem cells is one of the most controversial methods of treatment. It was only in 1998 that the first human stem cells were identified, and fundamental research is necessary before such cells can be used to treat spinal cord injury. Scientists have not yet discovered the techniques needed to grow large numbers of human stem cells and to manipulate them so they divide into the particularly desired cell types. Once convincing evidence for successful regeneration and improved functional outcome is obtained in animal models, clinical studies in spinal cord injured patients could be initiated using stem cell therapy in combination with other treatment strategies.

17.6 New Frontiers in Spinal Cord Injury Research

There are more than 1,000 new cases of spinal cord injury in Canada each year, and most patients are young individuals. The increasing numbers of spinal cord and brain injured patients strengthens the need for new therapies to repair the damaged spinal cord and other parts of the central nervous system.

The recent achievements in basic and clinical research outlined above overcame the dogma that patients with spinal cord injury are unable to recover to some degree. Despite Christopher Reeve's sudden death in 2004 at the age of 52 due to an infection, his case still defies many scientific and medical expectations. His story shows that new discoveries and future endeavors in research might help provide a cure for spinal cord injury even for those with the most severe neuronal damage.

Chapter 18

What's the Matter with Stress?
Gerlinde A. Metz

The opening of Robert Sapolsky's book *Why Zebras Don't Get Ulcers* (1994) describes the chasing of a zebra by a hungry lion in the African savanna. Sapolsky uses this example to illustrate a stressful condition affecting both the zebra and the lion under different circumstances. The case of the zebra reflects the classic mammalian crisis of being chased by a predator, and the stress experienced is short and intense. The mammalian body copes with such an acute emergency situation by a stress response, consisting of a complex cascade of rapid physiological changes. For a zebra sprinting away from a predator, this stress response is ideal for dealing with that kind of situation. However, it is anything but helpful for humans when they are feeling stressed over exams or worried about the future. While the human body is still built to deal with short-term stress such as outrunning a lion, the constant stressors of modern life might place the body in a critical condition.

18.1 The Concept of Stress

Although some ideas about stress and its implications in disease were developed millennia ago, the term stress has been used only recently. Hans Selye (1950) was the first to introduce the term stress and to systematically investigate the effects of stress on bodily functions. This line of research actually began by accident when Selye was investigating physiological effects of hormones as a young scientist. His experiments required that he inject rats on a daily basis with tissue extracts, a procedure that was quite bothersome for the animals. He noticed that the animals treated with the tissue extract showed enlarged adrenal glands and developed peptic ulcers and other pathological symptoms. His excitement about these findings ceased, however, as he discovered that control rats that received only saline injections showed similar symptoms. Selye assumed that not the tissue extract itself but the unpleasant procedure of daily injections led to the adverse symptoms. To test whether pathological changes in fact derived from

his manipulations, Selye subjected rats to other unpleasant stimuli, such as cold or hot temperatures, and illness. Surprisingly, these animals again developed peptic ulcers and enlarged adrenal glands. Selye hypothesized that different stressful manipulations result in similar consequences and he termed this effect the nonspecific stress response. Selye's findings initiated a flood of scientific investigations that led to the concept of stress. Stress can be defined as the sum of all nonspecific factors that can act upon the body to increase energy consumption above the basal (resting) level. Stress threatens the steady state of biological and physiological processes and thus adaptive responses take place in order to counteract the stressor and reestablish a steady state. A stressor is the stimulus eliciting such a stress response.

Selye's definition of a stress response is characterized by the onset of a suite of physiological and behavioural reactions in order to cope with an emergency situation. The zebra trying to outrun the lion has to deal with the acute need of an unpredictable amount of metabolic fuel. The adaptive stress response helps switch the body's processes into an emergency state to perform the "fight-or-flight" response. Selye proposed three phases of a stress response. Phase 1 is the alarm reaction under acute stress once the stressor becomes detected. In this phase, anabolic processes are interrupted in favour of mobilizing glucose (sugar) to deal with the acute need for energy. In other words, the zebra can digest its lunch *after* it has managed to escape the lion. Phase 2 of the stress response characterizes the body's resistance to a continued exposure to the stressor by adaptive coping mechanisms. The signs of the body that characterize an acute alarm reaction have virtually disappeared. Depending on the type of the stressor, the body is able to maintain the equilibrium of energy production and consumption for a limited period of time. Phase 3 describes the exhaustion stage in which the body's adaptive response to the stressor wears out. The signs of the alarm reaction might reappear and are irreversible, and the individual might eventually die. Work by Robert Sapolsky and others showed that the pathological effects of such chronic stress are due to prolonged activation of the stress response (Sapolsky, 1992) instead of a depletion of chemicals mediating the stress response as originally proposed by Selye.

It is important to note that the physiological response to stress is an evolutionary adaptation essential to survival of mammals, including zebras and humans. The stress response is designed to cope with an emergency situation via different physiological mechanisms that will be outlined in the following.

18.2 How to Survive an Acute Emergency

Let us assume the zebra was able to escape from the lion. How did it manage to immediately mobilize the energy necessary to outrun the lion? Obviously, the zebra's stress response has been adaptive and saved its life under these circumstances. Acute stress such as that experienced by the zebra results in an orchestrated cascade of physiological responses. These processes lead to an immediate availability of energy to sustain the sprint away from the lion, increased respiration and cardiovascular rates, a shift in blood flow away from inactive areas of the body towards muscles and brain, and inhibition of processes not critical for coping with the emergency situation (digestion, growth, reproduction, and immune function). Moreover, stress modulates pain percep-

tion so that responses to injuries become blunted. These responses are mainly mediated by two endocrine systems that become activated under acute stress. One system leads to a fast stress response within seconds, the other one leads to a slow stress response within minutes after the appearance of a stressor.

The immediate increase in available levels of glucose and oxygen is mediated by the fast stress response via release of epinephrine (also called adrenaline) from the adrenal glands. This response is produced by the autonomic nervous system. The autonomic nervous system mediates involuntary control of bodily functions and consists of two components that take opposite roles. The parasympathetic nervous system controls anabolic functions associated with growth, digestion, and reproduction. The sympathetic nervous system, in contrast, is stimulated by arousal or stress. Both these systems innervate essentially every organ in the body. Acute stress leads to inhibition of the parasympathetic nervous system while activating the sympathetic system. An increase in sympathetic tone results in the release of epinephrine and a close relative, norepinephrine, within a few seconds. These two chemicals stimulate the mobilization of glucose, increase cardiac activity, contract peripheral blood vessels, and affect muscle activity.

The complex action of the sympathetic nervous system is accompanied by the slow stress response mediated by steroid hormones. Like epinephrine and norepinephrine, the steroid hormones of the slow stress response, also called glucocorticoids, are produced by the adrenal glands as well. Within minutes after the onset of a stressor, the secretion of glucocorticoids is upregulated. This response is regarded as the essential component for the adaptation to a physical stressor. Glucocorticoids, mainly cortisol in humans, can exert a variety of effects, reflected in a inverted U-shape function (Sapolsky, 1997). The inverted U-shape function illustrates that, depending on the duration and severity of stress, the response to glucocorticoids can have beneficial or detrimental consequences. The name glucocorticoids indicates that these chemicals have the function to inhibit glucose and fat storage while at the same time stimulating glucose mobilization to provide muscles with metabolic fuel. In addition, glucocorticoids increase energy availability via the breakdown of fatty acids from adipose tissue. Glucocorticoids also have a variety of other effects, including anti-inflammatory actions. The suppression of immune function by steroid hormones became of clinical importance when they showed beneficial effects in the treatment of inflammatory and autoimmune disorders such as allergies.

The release of glucocorticoids from the cortex of the adrenal glands is stimulated through releasing factors produced by the brain (see Figure 18.1). It was quite tedious to determine the individual factors involved in this cascade while only minuscule amounts of each chemical are present in specific areas of the brain. It took thousands of brains derived from slaughterhouses and many years of research to determine the key hormones of this cascade. The two competing scientists who pioneered this field, Roger Guillemin and Andrew Schally, finally received the Nobel Prize for their discoveries. Their main finding was that the actual master gland initiating the release of glucocorticoids from the adrenal glands is the brain. Upon the experience of a stressful stimulus, the hypothalamus receives the signal to produce corticotropin releasing factor (CFR), which is transported to the anterior pituitary to initiate the release of corticotropin (ACTH). The anterior pituitary releases ACTH into the blood stream, which

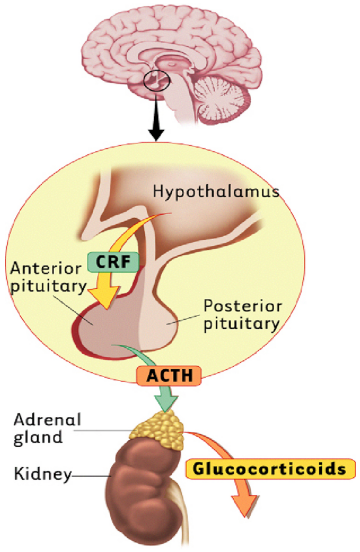


Figure 18.1: Overview of the control of glucocorticoid release through the hypothalamic-pituitary-adrenal axis (CRF is corticotropin releasing factor and ACTH is corticotropin).

then reaches the adrenal gland where it stimulates the production of glucocorticoids. They in turn act on various sites to divert energy to muscles, enhance cardiovascular tone, suppress nonessentials such as digestion, growth, reproduction, suppress immune function, and sharpen cognition. The rate of circulating glucocorticoid levels is regulated by a complex feedback mechanism through neurons possessing receptors for these chemicals.

18.3 Staying in the Equilibrium

Another important model for understanding the consequences of the stress response still needs to be introduced. We have outlined above the features of the adaptive stress response, yet it is crucial for the body's processes to maintain a steady state even during an acute emergency. Walter Cannon, a professor of physiology at Harvard University, introduced a unifying framework of mammalian physiology that is known as the concept of homeostasis (Cannon, 1932). By 1929 he recognized the emergency function of the adrenal glands in meeting vital threats to the body and in maintaining the equilibrium of the many physiological processes of the organism. Cannon's concept of homeostasis refers to the coordinated physiological processes of the body that adjust to current needs in order to achieve an optimum performance in a given situation. Under regular circumstances, this results in homeostasis, a physiological steady state. Homeostasis can be challenged by various factors, including temperature extremes, food or water deprivation, or an argument with the neighbour. To deal with such a stressful

situation, the organism might require more energy while the steady state of bodily functions has to adjust to these needs.

The concept of homeostasis is important for the understanding of the pathophysiology of chronic stress. While the mammalian body is adapted to cope with acute stress, prolonged exposure to a stressor might lead to permanent disruption of homeostasis. How long does a zebra's sprint in the Savanna take? It is not likely that it lasts for more than a minute or two, while today's life style requires the body to cope with stress for much longer time periods.

18.4 Stress and Disease

In the latter part of this chapter we return to our original question. Why are zebras less susceptible to develop ulcers than humans? One reason, as you might already expect, are differences in the duration of the stressor. A zebra that outruns a lion has overcome the stressful situation and might finally rest and continue digesting its lunch. In contrast, the human life style might provide more frequent and prolonged threats to homeostasis. Moreover, an aging human population is likely to contain individuals who have been exposed to higher amount of glucocorticoids across their lifetime than a zebra—which has a much shorter life expectancy.

Hans Selye was the first to explicitly note that chronic stress can cause disease (Selye, 1950). His interpretation was that a stress response becomes terminated in the long term because the body eventually runs out of supplies of the required hormones and neurotransmitters. Thus, he concluded that the individual becomes unable to cope with the chronic stress. This original view was revised by later research indicating that disease begins when the stress response actually continues and sooner or later becomes harmful (Sapolsky, 1987; McEwen, 2000). The excessive glucocorticoid exposure during a stress response, well-suited to save the zebra's life, might turn this adaptive reaction into a devastating process under chronic conditions. For instance, chronic stress or exogenous treatment with glucocortioids can cause or worsen various disorders, including myopathy, diabetes, hypertension, impotency, and infertility; and can persistently suppress immune function. The potential damage by elevated glucocorticoid levels can accumulate during the course of a lifetime and might accelerate senescence. In addition, recent research has indicated that the stress response in the elderly becomes impaired thus further increasing the risk of stress-induced disorders. It takes more time for an older individual to develop a full-blown stress response, and it takes longer to return to resting levels and reestablish homeostasis afterward. Furthermore, it was hypothesized that the exposure to glucocorticoids over a lifetime might determine the rate of cognitive decline during aging. It has been found that there is a high density of receptors for glucocorticoids in particular areas of the brain. A glucocorticoid molecule binds to one of these receptors inside a nerve cell and so interacts with physiological processes inside the cell. Neurons possessing these receptors therefore are very susceptible to the effects of glucorticoids. Interestingly, there is a particularly high density of glucocorticoid receptors in the hippocampus, a structure underlying learning and memory. Because of its high density in glucocorticoid receptors, this structure serves as a feedback station to control the amount of glucocorticoids in the system and to reg-

ulate the secretion of these hormones. Its sensitivity to glucocorticoids, however, might place the hippocampus in a vulnerable position under chronic stress conditions. Studies in laboratory rodents and field studies in primates have shown that hippocampal neurons might degenerate under chronically elevated levels of glucocorticoids, leading to loss of memory functions. It has been argued that cognitive decline during aging is at least partially due to exposure to glucocorticoids and loss of hippocampal neurons.

Chronic stress not only affects brain functions, but also keeps bodily processes activated for the fight-or-flight response for longer than necessary. Stress-related disease occurs when the stress response is activated for too long or too frequently. The symptoms of ulcers, cardiovascular disease, diabetes, reproductive inhibition, psychosocial dwarfism, and dysfunction of the immune system really are the tip of the iceberg of stress-induced pathologies. Nevertheless, it usually is not the stress itself that makes one sick, but stress makes one more likely to become sick. For instance, the suppression of immune system function seems appropriate for a zebra running for its life, but if this effect persists into chronic conditions it might make the organism more vulnerable to infectious disease or autoimmune dysfunctions.

18.5 Why is Stress Stressful?

The term stress is often used as a hypothetical concept in a mainly negative context. It is noteworthy that the organism is able to habituate to many of the physical stressors so they are no longer experienced as stressful. Nevertheless, many stressors in today's life style have a psychosocial component and thus the magnitude of a stress response depends on the individual's perception, personality, and perinatal experience. It is well accepted that psychological variables, for instance an individual's perception of a stressor, can modulate the physiological response. A number of studies have generated ideas about psychological factors that can modulate the stress response. For instance, the stress response can be diminished with gaining control over the stressful situation such as the option to avoid or reduce the magnitude of a stressor. Rats placed in a chamber that delivers mild electric shocks showed a significantly reduced stress response when they were trained to press a lever to decrease the rate of shocks (Plaut & Friedman, 1982).

Another variable that can modulate the stress response is the predictability of the occurrence of a stressor. If rats are given a signal before the onset of a mild electric shock, they show a lower stress response than rats that received an identical shock without the warning. Finally, an important component of stress coping behaviour is to find an outlet for frustration. Taking the example of a rat receiving mild electric shocks, a major reduction in the physiological response can be found when the animal receives a piece of wood to gnaw on or access to a running wheel. Such physical activities can serve as protective outlets to reduce the stress response. Furthermore, social support also plays a major role as a mechanism to cope with stress. The simple presence of a conspecific might reduce the response to a stressful procedure. Similar mechanisms can be found in primates. Imagine preparing for a difficult exam or undergoing a medical procedure. The presence of a close friend or family member might take away much of the stress associated these events. These examples illustrate that although the

stress response is a relatively rigid hormonal cascade, there are numerous factors that modulate the magnitude of the physiological response.

18.6 Conclusion

In today's life style, we rarely encounter the physical stressors that our ancestors experienced. Nevertheless, the original mechanisms mediating the fight-or-flight response are preserved. The complex influences of stress on the health of an aging individual require more research, however, the lesson that we can learn from previous findings is that the impact of external stressors can be considerably modulated by their perception and by psychosocial factors.

Chapter 19

Standing Up Straight, and Staying There

Sergio M. Pellis

Stand up and gently lean your back against a wall. Just a light touch is necessary. Now raise both arms so that they are straight out in front of you. Feel anything? Try again, and this time pay close attention to your back. You will notice that as you raise your arms, your back presses against the wall. This result may seem obvious and trivial, but actually it is quite a profound illustration of a general principle about movement and behaviour.

As you are reading this chapter, you reach over to pick up a can of Coke[1], and we see the same principle is at work. Quite simply, when you make a voluntary movement, such as reaching out to grab a can of Coke, your body automatically adjusts your posture so that you remain stable and balanced. If someone with an advanced case of Parkinson's disease[2] is asked to do the above described experiment, as that person

[1]The main reason you need the Coke is that it helps keep you alert while studying. Actually, what achieves this effect is a drug contained in the Coke, called caffeine. A major means by which nerve cells (neurons) communicate with one another involves chemical substances called neurotransmitters. The caffeine influences certain kinds of neurotransmitters that act in part of the brain that regulates attentiveness. Thus a little bit of caffeine stimulates this part of the brain and you can keep alert while reading this chapter or driving a car for a long distance at night. Too much, however, results in stimulation of lots of other brain areas, and you end up with the "coffee shakes".

[2]Parkinson's disease involves a particular constellation of symptoms, including a low amplitude tremor of the extremities and difficulty in initiating movements. The tremor is actually what is called a resting tremor. That is, the hands shake while they are resting, such as when folded in one's lap, but as soon as the person reaches to grab something, the tremor stops. The difficulty in initiating voluntary movements arises primarily from the absence of the automatic responses that enable postural adjustment. The disease arises from degeneration of a particular group of cells in the midbrain called the substantia nigra (black body). This cell degeneration results in a reduced input of a neurotransmitter (dopamine) into another brain area in the forebrain (the caudate nucleus), which is involved in working out how movements should be combined together to achieve a task.

raises their hands upward they will fall forward. By raising your arms you shift your centre of mass forward, in front of your centre of gravity (somewhere between your pubic area and navel); by leaning backwards, you shift your centre of mass backwards to compensate for the added weight out front caused by your raised arms. This compensation is an automatic, unconscious adjustment. We only become aware of it when we are robbed of this ability by a disease, such as Parkinson's (J. P. Martin, 1967).

19.1 Integrating Posture and Movement

There are a large number of such postural responses, usually automatic, which make voluntary movement possible, of which we are ordinarily not aware. In fact, these postural responses are so well blended into the execution of movements that we do not notice them, until something comes along to disrupt their normal function. Let me give you two concrete examples. Rats, like most other four legged mammals, walk with a characteristic sequence of steps, so that if the rat steps with the left front foot (LF), the next step will be with the right hind (RH), then the right front (RF) and finally the left hind (LH), before the cycle starts again[3]. Following some forms of brain damage, rats switch to a different gait sequence (LF–LH–RF–RH)[4]. Because four-legged animals switch to different gaits at different speeds of locomotion (e.g. walk, trot, canter, run, gallop), one possibility is that the brain damage has scrambled the organizing system that selects appropriate gaits (Hruska & Silbergeld, 1979). This did not end up being the correct explanation. Rather, we found that the gait shifts because of changes in the rat's ability to adjust its posture during walking. The brain-damaged rat will begin by walking with the normal gait, but after a few steps it will switch to the abnormal gait! At first, the rat's posture is normal, so that when the hind foot moves forward, it lands close to the forepaw onto the same side, and this unweights that paw (try it). However, as the steps of the forelimbs tend to be longer than normal, and the steps of the hind-limbs tend to be shorter than normal, after a few steps, the body is elongated. Once elongated, a step by a hind foot only moves a short distance forward, landing far from the forepaw on the same side. Now, instead of that forepaw being unweighted, it bears more of the body's weight, so that it is actually the forepaw on the other side of the body that supports the least of the body's weight, and hence the one that is easiest to step with (again, try it). These are the postural changes that lead to a change in gait (Pellis, Pellis, Chesire, Rowland, & Teitelbaum, 1987).[5] Confirmation for our explanation came recently from one of my colleagues in Israel. Studying the gaits of a

[3] To get this straight, try it. Get on all fours and try walking in this sequence. Given that your legs are so much longer than your arms, try this by standing on your hands and knees.

[4] This abnormal gait sequence for a rat is actually the normal gait sequence for some four-legged animals, such as baboons (Gambaryan, 1974), which suggests that varied types of gaits are represented in the brain for different species. For as yet unknown reasons, some of these are used, but not others. Brain damage may therefore reveal gaits not normally seen.

[5] This example also illustrates the difficulty in determining the processes that give rise to observed behaviour. In this case, a change in gait could be explained by either a direct change, such as in the brain system which regulates gait sequence, or indirectly, via changes in posture. To be able to differentiate between these two possibilities, a detailed analysis of the movements of the animals was necessary. To achieve a plausible diagnosis of such a pathology requires a lot of sleuthing, where all kinds of clues about behaviour are used to detect a pattern of symptoms that support a particular hypothesis over another.

variety of rodents, David Eilam (personal communication) found that a species of mole rat, which has a long sausage-like body and small legs, walks with the same gait as our brain-damaged rats!

In another example, brain-damaged rats were required to reach through a slot onto a little platform to grasp and retrieve a small, highly tasty (to rats) food pellet. Normal rats learn this task very quickly and mould together a rat-typical sequence of body and arm movements to achieve this task (I. Whishaw & Pellis, 1990). The surgically treated rats, who had a form of brain damage that was a variation of that of the rats in the walking experiment,[6] had enormous difficulty in retrieving any food pellets successfully (I. Whishaw et al., 1994). The rat would position itself at the slot, then dither while making little forward and backward movements, and when it finally did reach forward, it would fall over. The rat's problem turned out to be similar to that of the Parkinson's patient described above. That is, the rat was unable to adjust its posture. Typically, as a rat reaches forward, it shifts its body backwards simultaneously, thus keeping its centre of gravity in about the same place. The brain damaged rat cannot do this. After the initial difficulty in raising its paw (hence, all the dithering—how to unweight the forepaw?),[7] it does so, but then it topples over.

In these examples I have focused on what is called postural fixation, that is, the ability to shift one's posture so as to enable movement (e.g. walking, reaching) without falling over. There are, in fact, a whole range of such 'postural reflexes', including ones which are used to initiate walking, to stop yourself from falling over if pushed, and to right yourself if you do fall. Most of these were identified and characterized in the early part of this century by the neurologist, Rudolf Magnus, and his colleagues (Magnus, 1924). However, to facilitate this analysis, Magnus used various kinds of brain-damaged animals. As a consequence, many have been studied in clinical settings as signs of pathology. It was a Japanese scholar, who having taught himself German so as to be able to read Magnus, realized that these varied reflexes actually operate daily during normal behaviour (Fukuda, 1984). For example, a reflex called the crossed-extensor reflex occurs when turning one's head to the side. This results in contraction of the extensor muscles in the arm to which you turned your head[8]. Fukuda noticed

[6]The rats in the walking experiment were made Parkinsonian by destroying the cells carrying dopamine to the striatum (the rodent equivalent to the human caudate—see the footnote on Parkinson's disease, 147). When this operation is done to an adult rat, it becomes completely akinetic (a = absence, kinetic = movement). In order to make these rats walk, they were injected with a drug that blocked the action of another neurotransmitter, acetylcholine, that occurs in the striatum. However, once the drug wears off, the animals again cease to move. They will also not spontaneously feed themselves, and have to be fed, by pumping a liquid diet directly into their stomachs, every 3-4 hours, day and night. The rats in the reaching experiment were given the same surgery, but as newborn infants, not adults. These rats grow up seemingly normally, and can feed themselves, but as explained in the text, they do have deficits.

[7]After sufficient practice, the brain-damaged rats do succeed in retrieving some pellets. In this sense, they seem to recover. However, 'recover' usually means the reinstatement of that which is lost. In this sense, these rats never really recovered; rather, they learned how to compensate for their problem: they raised the other forepaw and then caught themselves with that paw on the front wall, to one side of the slot. They thus held up their body with this paw, which freed the other paw so they were able to reach for the pellet! With such an incredible capacity to compensate for deficits, differentiating between true recovery and compensation can be quite difficult.

[8]This behaviour would be most clearly seen in a baby (not a toddler), as they have fewer complete neural systems than adults to interfere with this reflex. A medical practitioner would be able to show you that if you rotate the baby's head to face the right, the right arm would extend, and if you rotate the baby's head to the

that this reflex is the one used to strengthen and lengthen one's reach such as when catching a ball. Similarly, stretching the ligaments on the front of your foot and ankle causes the muscles at the back of your lower leg to contract, which makes you stand straighter. This effect is why models usually wear high heels to "lengthen their line", and why cowboys who wear high-heeled boots "walk tall". It is now well recognized that posture and movement are integrally interconnected processes, which develop at different rates. This difference in developmental rates is one of the main reasons why babies are clumsy, and fall over. The neural systems controlling movement mature before those that control posture. Similarly, our clumsiness when we learn some new physical task arises from an initial lack of coordination between posture and movement.

19.2 Righting Reflexes

So postural adjustments ensure that we can move without falling, but what if we do fall? Without being aware of your actions, you would first thrust out your arms towards the ground and raise your head. This response would reduce the likelihood of injury, especially to the head. Once on the ground, you have to regain your upright stance. Again, by studying Parkinsonian patients, it can be demonstrated how protecting yourself when falling and then getting to upright involves a number of reflexive responses[9], of which we are normally unaware. As the disease increases in severity, Parkinsonian patients have greater and greater difficulty in turning over in bed (righting), and have to adopt all kinds of bizarre strategies in order to do so (recall our reaching rats).

I first became interested in righting reflexes when studying Parkinsonian rats. The puzzle was that while Parkinsonian patients have difficulty righting, Parkinsonian rats do not[10]. Trying to solve this puzzle led into a 12 year journey to try to make sense of the righting response.

---

left, the left arm would extend.

[9] There is a long history of debate about how to define a reflex. The simplest reflex is illustrated by the quick withdrawal of your hand when you touch something very hot. Here the sensation of heat on your fingers travels along sensory nerve cells (neurons) to the spinal cord, then back to the fingers along motor neurons, which causes the muscles in your hand and arm to contract, so you thus withdraw your hand. This pathway is the classic reflex arc—sensory nerve to spinal cord, spinal cord to motor neuron—defining a reflex in terms of the structural connections in the nervous system. However, many reflexive responses such as righting, involve more complex relationships of connections in the brain as well as in the spinal cord. The main feature distinguishing these reflexes is that they appear to be automatic, that is, they are performed without our awareness. In this case, a reflex is labelled by its functional qualities, not by its specific structural arrangements. In this chapter, the term reflex is used to signify a response which we normally perform without being aware of its execution.

[10] Animal models of human pathologies are essential to be able to characterize the damage producing the pathology and to experimentally develop avenues of therapeutic intervention to ameliorate the symptoms, if not cure the disease. However, while animal models are never 100% replicas of the disease, they are usually close enough to be useful. For example, animal models of Parkinsonism simulate akinesia very well, and have provided a valuable means of developing new drugs which provide relief to patients. One of the reasons why there is a lack of 100% similarity in animal models of Parkinsonism is that in the animals, the damage is confined to the substantia nigra, the key neural system affected by this disease. In the real disease, while this structure is hit first, the pathology may spread to other neural systems, and so may result in additional deficits. For example, problems with righting tend to occur after difficulties with initiating movements arise. Treatment with dopamine ameliorates the difficulties with initiating movements, but not with righting (Lakke, 1985).

When a rat is placed on its back on the ground, it will quickly right itself by rotating the body, head first, to a prone, standing position. In one of the forms of brain damage that we were studying, the rats made all kinds of bizarre movements—when their head turned to the left, their pelvis would turn to the right, and their legs would make 'pumping-like' movements. It seemed that the rat took forever to right itself (Pellis, Pellis, Chen, Barzci, & Teitelbaum, 1989). At the same time, my wife was studying the righting of infant rats: they also showed the same bizarre patterns of righting (Pellis, Pellis, & Teitelbaum, 1991). What were they doing? Why did the infants resemble the brain-damaged adults? And what, if anything, did this have to do with Parkinson's disease?

The clue came from Magnus' classic work on righting (1924). He showed that righting can be triggered independently by several sensory systems. The organs of balance, situated in a bony labyrinth in the inner ears[11], can trigger the head and body to right. Similarly, the sense of touch on the body can also do so. Magnus divided this tactile (touch) based righting into two distinct types. Touch on the body can trigger the head to right, but if the head is physically restrained, such tactile information triggers righting by the body. More recently, a third type of tactile based righting was discovered. Tactile input on the face can trigger a righting by the head (Troiani, Petrosini, & Passani, 1981; Pellis et al., 1989, 1991; Pellis, Pellis, & Nelson, 1992). The final type of triggering involves vision. A scheme of what triggers a righting movement and then what it is that moves can be constructed. There is, therefore, vestibular-on-head, vision-on-head, head (tactile)-on-head, and body-on-head. The fifth differs in that it is body-on-body. The first four could be thought of as four different ways to trigger the same response[12]. This aspect of Magnus' model had us confused for a long time. These sensory systems were not triggering the same response. Rather, each form of righting was distinct. The clues were the tell-tale differences in the details of the movements used during righting when each of these types of righting were engaged.

Following brain damage (Pellis et al., 1989) and in early development (Pellis et al., 1991), some of these righting systems are missing, those that are present are incomplete in their execution, and several of the systems present occur simultaneously. The net effect is that movements counterproductive to righting are performed. For example, body-on-body righting involves flexing the hind-leg closest to the ground, and tucking it beneath the body and placing it on the ground, and then rotating the lower body on top to an upright position. In the early stages of recovery from brain damage or in early infancy, the hind-leg is placed, but no rotation of the body occurs. In order to right, the leg extends pushing the body up and over to prone. Unfortunately, more

[11] Collectively, this collection is referred to as the vestibular apparatus, and contains two major components. The sacculus and utriculus are two little chambers sitting askew from one another. Cells with little sensory hairs (cilia) project up from the floor of each chamber. The cilia are coated with a layer of mucous, on top of which sits little stones (otoliths). As you move your head, the stones pull on the cilia. This information is then passed on to groups of cells in the brain. The second component involves three fluid filled canals (one horizontal, and two vertical at 90° from one another). The inside of the canal has cells projecting cilia into the tube, so that when you turn your head the fluid moves, moving the cilia. Again, this information is passed on to cells in the brain. The first system lets you know the position of your head in space, and the second informs you about the movement of your head.

[12] In an ingenious set of experiments, Magnus showed that each of these sensory systems could, independently of each other, trigger the righting response.

often than not, the foot is not far enough beneath the body, so that when it extends it flexes out into the air. Repeated attempts to get a foothold leads to the 'hind-leg pumping' mentioned earlier. In both recovery and development, each righting system goes through the same stages of transformation. The interaction of these incomplete systems leads to the bizarre patterns of movement that puzzled us at the beginning of our work.

This new model of righting, where each type represents not only a distinct type of sensory system, but also a distinct type of movement, finally made sense as to why brain damaged adults resembled normal infants. But what has this to do with Parkinson's disease? Lakke (1985) showed that the difficulty that Parkinsonian patients have in righting results from an inability to rotate their body axis[13]. He described various idiosyncratic ways in which patients "learned" to overcome this problem. Some used their legs and hands to push or pull themselves over. Others first sat up and then turned to the side, which required minimal axial rotation. Our work suggests a different interpretation for why different patients use different strategies to right.

As Parkinson's disease progresses, the patient's abilities diminish, whereas both our brain damaged rats and our developing infants improved over time. We can view these two trends as mirror images of one another. Many of the "idiosyncratic" strategies of the patients actually looked similar to our rats when only certain portions of particular righting systems were available to them. This suggests that as patients deteriorate, they gradually regress, showing the opposite changes to the rats; that is, they are losing, rather than gaining, components of righting systems. If so, then a diagnostic-schema could be chartered to identify the stages of deterioration which particular patients have reached[14].

19.3 Conclusion

We have seen that maintaining yourself in an upright position involves a complex interweaving of postural changes which adjusts your balance with your movements. Once you do fall, there are a myriad of types of responses you can use in order to regain your upright stance. In fact, we recently demonstrated that righting from a lying down position involves a different class of righting than when you right while falling over (Pellis & Pellis, 1994). This distinction is true across sensory modalities—therefore, there is righting when stationary and righting when falling—one which utilizes information from the vestibular apparatus, and the other which utilizes tactile input.

Which type of righting, or which types should be blended together in a particular instance, is all decided without us being aware of the process. Indeed, most of our more complex behaviours are performed without any thought. If you think about stepping as

[13] Actually, when lying down they cannot rotate the body axis, but if they fall they can. Therefore, they are not paralysed (i.e. unable to move those muscles), but rather are apraxic (i.e., the ability [praxis] needed to move these muscles is absent [a]). The muscles are able to move, but not when you wish them to. For this reason, Lakke (1985) called this inability in righting axial apraxia.

[14] Another major problem is that these deficits in righting do not result from damage to the substantia nigra/caudate system that is responsible for the typical Parkinsonian symptoms. Clearly, additional brain systems must undergo pathological change. What are they?

you are walking downstairs, chances are you will trip. One of the tasks of psychology is to figure out how the brain is able to work out these complexities.

Chapter 20

What is the Point of Play?
Sergio M. Pellis

That young animals play is well known. Most of us are familiar with the antics of kittens and puppies, or even monkeys at the zoo. But these are captive and pampered animals who do not have to hunt or forage for themselves. What is more surprising is that in the wild, young animals spend an extraordinarily large amount of time playing. For example, for my Ph.D. research, I studied the development and play of Australian magpies, a crow-sized bird, native to Australia. By four weeks, after fledging, they spend up to 20% of their daytime activity engaged in play. This is at a time when about 50% of their food comes from their own foraging, which has a low rate of success[1]. Clearly, young animals are spending time and consuming energy that could otherwise be devoted to growth. In addition, studies of free-living juvenile animals have shown that play also exposes them to the risk of predation and injury.

So why do they play? Any television nature program will provide the answeryoung animals learn all kinds of things from play—how to be good hunters, how to avoid being prey, how to fight, how to have sex, etc. There is a problem, however; the actual evidence for this is weak, or non-existent. However, for humans, the value of play appears self-evident. Again, the evidence is weak or even contradictory.

A marvellous experiment by Corinne Hutt illustrates the problem nicely. She presented children with a novel and complex object, a box with a lever that operated a bell, a buzzer, and four counters. Once the object was fully explored, it was incorporated into play. During play however, the children were less likely to discover unknown properties of the object. Indeed, Hutt concluded that "...by being repetitive, play is by definition a highly redundant activity, and can actually prevent learning" (Hutt, 1966, p. 76). Another supposed benefit of play in humans is illustrated by the use of play therapy to solve all kinds of psychological problems in children. A review

[1]Magpies live in open woodland and mainly feed on ground living (soil and grass) insects and other invertebrates.

by Phillips (1985) on its effectiveness revealed that most reports involved uncontrolled case studies, or assumed folk wisdom. The few experimental studies do not support the supposed power of play to ameliorate psychological difficulties. For example, in one of the studies reviewed, retarded children were exposed to one of three conditions, with each group being exposed to 29 sessions over 16 weeks. One group received play treatment, another was given a movie to watch (placebo), and the third received no special treatment. Personality and behaviour were found to improve in both the play and the placebo groups[2], with no statistical difference between them. The researchers concluded that the beneficial effects were due to meeting in groups and to positive affect (psychological jargon for feeling good).

The basic problem has recently been well put by Nigel Barber (1991), who points out that "(p)lay research suffers from the inheritance of empirically unsupported assumptions that are frequently accorded the status of fact—apparently because they are reasonable, satisfy a need for explanation, and have been repeated by authorities" (p. 129). The bottom line is we do not know why animals, including humans, play. I will illustrate the lack of empirical support for some of the common assumptions about play and then suggest a different way of looking at the problem.

20.1 What Do Animals Do During Play Fighting?

As the name implies, play fighting is an activity which looks like fighting but does not lead to injury, or other consequences typically associated with aggression[3]. Until recently it was believed that "(t)he behavioural components of play fighting vary from species to species but are generally predictable from a knowledge of those components involved in the fighting of adults" (Meaney, Stewart, & Beatty, 1985, p. 5). But this is not true. Rather, while some species use species-typical components of fighting in their play, others do not. The key concept that was essential for distinguishing between these two types of play fighting was that of the "target". In a dog fight, dogs do not bite each other just anywhere, but aim their bites at specific body parts—the back of the neck, the throat and the muzzle. They will generally avoid biting other body areas. Therefore, a dog fight involves the attack and defence of these particular body targets. Similarly, play fighting by puppies involves the playful attack and defence of these same body targets. From this we can conclude that play fighting for dogs is indeed predictable from the components of serious fighting.[4] This was also true for my magpies.

[2]This experiment shows us the importance of a placebo group. That is, some factor, other than the one you think is important, may be producing the experimental effect. For example, at a party, most of us will behave in a jolly manner after drinking beer, even if unbeknownst to us, there is no alcohol in the beer! This does not mean that alcohol has no effect, rather, it means that when we assess these effects we have to make sure to factor out other contributing influences. That is the role of the placebo group.

[3]Social play, or play involving two or more members of the same species, is one of the most commonly observed forms of play. The most commonly observed form of social play is play fighting or rough-and-tumble play. Other forms of play include object play, which is frequently seen in carnivores, such as cats, and incorporates predatory behaviour patterns; solitary play which usually involves locomotor activity, such as the gamboling of lambs; and humans also have fantasy play, which involves make-believe scenarios.

[4]You may have noticed that I have avoided defining play. The reason is simple—we all know play when we see it, or think we do, but trying to pin it down by a set of criteria that unambiguously distinguishes play from other behaviour is notoriously difficult. Therefore, for our purposes, think of kittens, puppies or young

Like everyone else, I believed this was true for all species that exhibit play fighting—that is, until I met the rat. When engaged in serious fighting, rats target their bites at the lower back and rump of the opponent. However, when engaged in play fighting, rats target the back of the neck. Adding insult to injury, when they contact the neck playfully, they do not bite it, but instead rub their noses into the fur! After five years of work, my colleagues and I confirmed that for 12 different species of muroid rodents (mouse-like rodents including voles, deermice, gerbils and hamsters), the play fighting targets are not the same as the body targets bitten during serious fighting, which for all species are the lower back and rump (Pellis, 1993). For example, in the Syrian golden hamster, the cheeks are the targets of attack and defence during play fighting. If contacted, the cheeks are gently nibbled. In the Djungarian hamster[5], play fighting involves trying to get at the mouth, which is licked and nuzzled if contacted. Moreover, the body areas contacted during play fighting are the same ones that are contacted during sex (during foreplay, to be exact). Therefore, for some species, play fighting has only a superficial resemblance to serious fighting, and its form can be predicted from the species-typical components of sex, not fighting.

Nature is never so kind as to provide us with simple categories that neatly divide the world into black and white—shades of grey always intrude. So too with play fighting. Had our finding been as straightforward as it appeared at first, then there would be two forms of play fighting—true play fighting, which involves components of serious fighting, and pseudo-play fighting, which involves components of sex, and hence should really be called sexual play or play sex. The Djungarian hamster shatters this picture. In fact, during play fighting, it attacks and defends both the mouth (a sex target) and the rump (a fighting target). This may also be the case for other species. For example, one of my students at Lethbridge found that for kittens, playful attacks can be directed at the neck and throat (a fighting target) and the anogenital area (a sex target). These are not just randomly directed bites, rather, the cats orient themselves so as to gain access to these body areas, even if it means moving the mouth past some other part of the body. Therefore, play fighting may be seen as involving a spectrum: at one end, all the behavioural components are predictable from fighting, and at the other, all are predictable from sex. In between it is a mixture. A necessary precursor for any general theory that accounts for why animals play fight is to know what it is that animals do during play fighting. Clearly, the what is not as simple as was previously believed (Pellis, 1993).

20.2 Big Brains Mean Lots of Play

Think about the animals you have seen playing. Dogs, cats and monkeys come to mind. Not sparrows or starlings. What about a playful lizard? Not likely. More species of mammals (warm blooded, hairy or furry, milk producing animals like us) have been reported to play, than have species of birds. As for other types of animals, no convinc-

children engaged in what you believe play to be. That is the phenomenon I am dealing with here.

[5] So named because it comes from Djungaria, which is near the Russian part of Mongolia.

ing reports have yet come to light. Birds are warm blooded and feathery (not furry)[6], so why don't as many bird species play as do mammalian species? Think about what it means to be called a 'bird brain'? In general, birds have smaller brains than mammals, even when differences in body size are taken into account[7]. Even within birds, those that are most playful are the larger brained species, such as crows. After all, you don't get much play out of chickens! Amongst mammals, predators are larger brained than prey species, and hence dogs and cats are more playful than sheep and cattle. The most playful mammals are primates—the playful antics of monkeys and chimpanzees are legendary. Not surprisingly, primates are the largest brained mammals. The most playful primate is our own species, and we have the largest brain amongst primates. Get the picture? The brainier the species, the more it plays. No less an astute observer of the universe than James T. Kirk, Captain of the USS Enterprise, has made this very clear: "(t)he greater the complexity of the mind, the greater the need for the simplicity of play" (Star Trek; Episode – Shore Leave).

Is it really true that we need a large brain so as to play? First of all, it is necessary to consider the brain in a little more detail. When comparing two mammals, if one has a relatively larger brain than the other, it means that the forebrain area (the cortex) is larger. More specifically, it is the part of the cortex called the neocortex, which is a new (i.e., neo) development in the mammalian brain. So the above conclusion can be restated: mammals with a larger neocortex play more. But is a neocortex needed in order to play? My colleagues and I recently did an experiment to test this conclusion. Under anaesthetic, the cortex of infant rats was removed by suction. Their play fighting as juveniles was then compared to that of the play of siblings whose cortices (plural of cortex) were left intact. The decorticate rats played just as frequently as intact rats, and used all the same behaviour patterns. The decorticates differed a little in how frequently they used some of the behaviour patterns, but in sum, it could be concluded that frequent, complex play did not require an intact cortex for its occurrence. That is, more primitive parts of the brain are sufficient to generate play (Pellis, Pellis, & Whishaw, 1992). So it is not true that you need a big brain in order to play, but it may still be true that you need the neocortex in order to learn whatever lessons are to be learned from play.

To understand what may be learned, a digression is in order. Adult male rats that have been decorticated as infants are incompetent when it comes to sex. They will approach a receptive female, and mount her head or her flank, with a consequent lack of success regarding copulation. That is, adult male rats that were deprived of their cortex as infants are inept at foreplay (I. Q. Whishaw & Kolb, 1985). Intact adult male rats that are deprived of the opportunity to play show the same incompetence at foreplay as do decorticate rats that have had the opportunity to play as juveniles. From this we can draw two conclusions. First, play is necessary for the development of male sexual competence, at least in rats and some other species, such as rhesus monkeys (C. Moore, 1985). Second, the cortex is necessary for learning this skill. So play does seem

[6]Fur and feathers both serve the same function, that of preventing heat loss. Both birds and mammals produce energy so as to generate heat and hence are called "warm-blooded". The energy loss would be greater without insulation, and hence more energy would need to be generated so as to keep warm.

[7]The brain of a blue whale is much larger than the brain of a human. However, the whale weighs several tons, a human weighs around 100-200 lbs. Therefore, relative to our size we have larger brains than whales.

to be important for learning something, and having a big brain does mean something when it comes to play. Maybe.

20.3 Immediate Costs/Delayed Benefits

"Most theories about the functions of play assert that play has immediate costs and delayed benefits. The young animal, it is argued, plays at the immediate expense of time and energy expenditure but subsequently benefits from, say, improvements in predatory or social skills later in ontogeny" (P. Martin & Caro, 1985, p. 85). That is, play now, profit later. Several studies have attempted to demonstrate that improved predatory skills in carnivores (e.g. cats and coyotes) result from playful experiences as juveniles. The results have been inconclusive. As noted above, however, for rats and monkeys, juvenile play experience has been demonstrably linked with adult sexual competence. Therefore, juvenile play may function to practice foreplay. You play, then, in order to get good at sex. However, there are some problems with this conclusion.

The first difficulty with this conclusion is that not all species that exhibit juvenile play show this relationship to adult sexual competence (e.g., guinea pigs). The second problem is with the timing of the critical age at which sexual competence is acquired. Play typically follows an inverted U-shaped curve during development, so that it is relatively infrequent in infancy, peaks in the juvenile period, and then is reduced to a low level again after sexual maturity. Thus in rats, for example, the peak in play occurs 30-40 days after birth. Yet rats fail to acquire sexual competence if deprived of sibling contact at 14-18 days (Gruendal & Arnold, 1974), which is before the peak in play! In a study of play and sexual competence in sheep, Orgeur and Signoret (1984) showed that the peak in play occurs at four weeks after birth, yet the critical time for the acquisition of sexual competence is between three to six months. In other words, the critical experience for sheep is gained after the peak period. The period of maximum play in the juvenile period cannot be explained by the function of acquiring sexual competence. The third problem is that for both rats and sheep, one successful copulation by these play deprived individuals leads to a complete elimination of the incompetence in foreplay. So why all that play if one correct result is all that is needed? Practice implies the more you do it, the better you get. On the other hand, once monkeys miss out on the juvenile experience, they will never gain competence again, no matter how much sexual experience they have later. So while there may be some truth to play being present so as to practice foreplay for some species, this is not a valid conclusion for all.

What about the possibility that the benefits are gained at the time the play is performed? That is, juveniles play in order to make them better juveniles, not better adults. There is not much evidence for this either. One study suggests that juvenile rats play so as to regulate their energy and nutrient intake (Barber, 1991). The idea is quite simple. Imagine that a juvenile needs 10 g of protein a day[8], and has access to an unlimited supply of food that has a concentration of 0.01 g of protein per 1 g of food. That means they would have to eat 1000 g of the food per day. Given that a 30 day old rat only

[8]Proteins form the basic building blocks of our cells, and are critically needed during development when our bodies grow in size, which means increasing the number of cells present. Other major nutrients in food (carbohydrates and fats) are used as a source of energy to maintain and repair cells, and for cells to reproduce.

weighs about 35 g, this means that that rat will need to eat nearly 29 times its own body weight! In order to get the needed proteins, the rat would gain far more calories than needed, and unless eliminated, this would lead to a very fat juvenile rat. This would be a great liability. Think of what chance an obese rat would have in escaping from a cat! Therefore, Barber's solution is that by playing, the juvenile rat eliminates the excess calories, while accumulating the needed proteins. Ingenious, and it may well be true for rats. Unfortunately, this cannot explain play in all species. For example, Galapagos fur seal pups gorge themselves suckling so as to gain an enormous amount of weight and then once the mothers have left them for good, uses these energy stores in order to sustain themselves and to complete their growth. Yet when the mothers leave them, the young seals persist in playing with one another, wasting energy that would otherwise go toward growth and that cannot be replaced (Arnold & Trillmich, 1985). Therefore, energy regulation cannot be the universal explanation for the occurrence of juvenile play.

Recently, my colleagues and I have found a possible immediate benefit for the play of rats. Rats live in colonies, with the males forming dominance relationships, where one male is dominant and the others are his subordinates. Male rats, furthermore, keep playing together well into adulthood, but only at a rate of about 5% compared to the juvenile peak. It turns out that the subordinates preferentially initiate playful contact with the dominant, not with each other. Also, when the dominant responds to their playful overtures, the subordinates respond as they did as juveniles, which is by rolling over onto their backs. Basically, the subordinates are using play as a means of 'sucking up' to the dominant (Pellis, Pellis, & McKenna, 1993). Syrian golden hamsters are solitary as adults, and if forced to live together after the juvenile period, the subordinates will try to avoid playful contact with the dominant (Pellis & Pellis, 1993). This shows that in this species, sucking up has not evolved as a strategy for peaceful coexistence. The example of rats shows that play can have immediate benefits, but this does not explain the juvenile peak in play. Also, the example of the hamsters shows that a function for play in one species does not predict a similar function in other species.

20.4 Conclusion

Statements beginning with 'play is' or 'play does' should be viewed with great scepticism. As I have indicated, there is very little evidence of play having any beneficial consequences, either at the time it occurs, or later in life. Some of the evidence presented that does show beneficial consequences may only be true for one, or a few species. There are no benefits of play that are true for all species that play. It seems common sense that if juveniles, of whatever species, engage in play it must be good for them. Science, however, is usually the opposite to common sense.[9] If something

[9]This is well illustrated by our understanding of motion. "Imagine being in the centre of a large flat field. If one bullet is dropped from your hand and another is fired horizontally from a gun at exactly the same time, which will hit the ground first?" The answer is neither (or both). "That the bullet which is fired is travelling horizontally has no effect on how fast it falls under the action of gravity" (Wolpert, 1992, p. 3). Another example: imagine an aircraft dropping a bomb. Where will the bomb hit the ground, in front, beneath or

makes common sense, then in science that rarely will be the correct answer.

With respect to play, perhaps the question we started with is all wrong. Although playful behaviour occurs, especially in the juvenile period, it may not be the phenomenon that needs explaining. Let's turn this problem on its head. Perhaps what needs explaining is the occurrence of a juvenile phase. Broadly, this period can be defined as that phase in life where the animal no longer depends on its parent(s) for food, but is not yet sexually mature[10]. There are two general theoretical perspectives on this issue. First, a juvenile phase exists so as to provide the young organism the opportunity to learn many things, with one of the means of learning being play. Second, the juvenile phase may be a by-product of growth rules. That is, it takes a certain amount of time for the organism to achieve its adult size. From this perspective, play may be just something juveniles do to fill up all the spare time they have.

The first view cannot be correct, because as I have indicated there is virtually no evidence that animals learn much at all during play. But then, the second view cannot be completely correct either, because I also showed some evidence that some animals do gain some benefits from play. This leads to a third perspective. That is, the length of the juvenile period is set by growth laws, but once present, the juvenile can make good use of its time. Put simply, the juvenile period did not evolve so that youngsters can play, but since they have to pass through a juvenile period, those juveniles that do play may be better off than those who just pass their time doing nothing (Pagel & Harvey, 1992). An advantage of this perspective is that it discourages universal pronouncements about what play does. Rather, it recognizes that in some cases, play may have no benefits, in others, it may have benefit A, while in others again, it may have benefits B, C, or D. Thus, we are forced to look more closely at each individual instance of play, rather than assuming we know why it is there.

behind the plane? Unlike common sense, "science often explains the familiar in terms of the unfamiliar" (p. 3). In this case, the relationship between gravity and horizontal motion is not as we would intuitively think.

[10]The juvenile period is a phase of delayed maturation that occurs in invertebrates (e.g. insects) as well as vertebrates (e.g. fish, mammals). However, it is most pronounced in mammals, especially primates. This takes us back to brain size. Mammals with the largest brains also have the longest juvenile period. Therefore, the correlation between brain size and play may result from a common cause—length of the juvenile period. Correlation does not equal causation!

Chapter 21

Uncommon Sense
Glen Prusky

"Look at me!
Look at me!
Look at me now!
It is fun to have fun
But you have to know how.
I can hold up the cup
And the milk and the cake!
I can hold up these books!
And the fish on a rake!
I can hold the toy ship
And a little toy man!
And look! With my tail
I can hold a red fan!
As I hop on the ball!
But that is not all
Oh, no.
That is not all. . . "

That is what the cat said . . .
Then he fell on his head!
He came down with a bump
From there on the ball.
And Sally and I,
We saw all the things fall!

—*The Cat In The Hat*, Dr. Seuss

This passage, from one of the best known children's stories in the English language, is a feast for our senses. It is not difficult to imagine the Cat In The Hat perched precariously on a ball, juggling a menagerie of coloured articles along with the reluctant fish, not to mention the noise accompanying his fall from grace. Presumably, the reason the words on the page alone can provide us with such a vivid percept is that we have had many related sensory experiences in our life. We know that balancing and juggling simultaneously would be difficult, that red is different from blue, what milk tastes like, the smell of a cake and the noise a rake and a fishbowl would make as they met the floor, let alone the bemused look the cat would wear after the disaster! However, it is unlikely that you have ever given serious consideration to how it is that you distinguish the sounds, smells and sights of the world, because it happens instantaneously and it takes no effort at all. It may come as a surprise, though, that sensory processing is not a simple task. Your senses do not just

reflect the external world but they first filter most of it out and then "create" a percept for you that has every appearance of clarity, precision and completeness. The problem is that everything you have ever experienced or will ever experience, must have or will pass through your sensory systems! There is no other way to acquire information. Like it or not, this is your Reality with a capital R.

One can view our sensory systems as being exquisitely sensitive to their respective environmental stimuli. For example, it is often stated that your olfactory system can detect a single drop of perfume in a very large room, or that you can hear a pin drop in a quiet room, or that you can see a few photons of light in a dark room. While in some ways these are impressive statistics, it is my contention that it is a wonder that we see or hear at all, given how bad our external receptors are.

In the first place, our receptor systems are anything but sensitive. In the auditory system, for example, adults can only hear things in a frequency range between 50 Hz and 20,000 Hz, if they are lucky. Rodents and other small animals can produce and detect frequencies up to 100,000 Hz! The next time that you are in a forest at dawn or dusk and notice that the animal chatter is loud, imagine how loud it would be if you could hear all of it!

Your visual system is even worse. It is true that you can probably detect visible electromagnetic energy from about 350 nm (violet) to 700 nm (red). This detection seems impressive until you realize that it represents only a tiny fraction of all the electromagnetic information that is out there in the world. If you turned on an AM radio and tuned it to a local station it is detecting electromagnetic energy 10 orders of magnitude higher in frequency than you can! In addition to the AM signals out there, AC, TV, FM, radar, infrared, ultraviolet, X-rays, Gamma-rays and other debris from the Big Bang are all around you all the time and you are completely unaware of it because you do not have the sensory equipment to detect it.

Why do we lack sensitivity? The reason is that our senses have not evolved to give us all information, but just give us the information that has been useful to our lineage through evolution. The ability to hear ultra-high frequency vibrations, or broadband electromagnetic energy has not been advantageous. Hominids are generalists and, with the exception of colour vision, none of our sensory systems are really good or really bad. You can find many animals that have more sensitivity than humans in one or more of their sensory modalities. That sensitivity has been advantageous to their evolution, but they have paid a price for it; the sensitivity they have gained in one modality has usually been offset by a reduced sensitivity in another. Rats, for example, have excellent auditory, somatosensory and olfactory capabilities, but their visual systems have lower acuity and no capacity to process colour. Eagles, on the other hand, have far better visual acuity than humans, but their somatosensory capabilities are less impressive. Rodents don't need high visual acuity and colour vision because they occupy dim niches in a nocturnal environment. However, birds of prey need excellent vision in order to localize their dinner while flying. We do not predominantly occupy either the night or the sky and the sensitivity of our sensory systems reflect this.

Second, in many ways your sensory receptors are simple, crude and convoluted in their function. Your auditory system seems to have been cobbled together from the junk heap of evolution. Sound waves enter your ear and cause sympathetic vibrations in your ear drum. Your ear drum is connected to a series of small bones, which believe

it or not, your reptilian ancestors used to articulate their jaws. The last bone in the series moves like a piston within a small hole in the cochlea, creating waves in a fluid that move the hair cells of the basilar membrane. In other words information is initially air conducted, then it is translated into bone movement, then it is translated into fluid conducted vibrations, where it finally mechanically-stimulates hair cells that translate the movement into neural activity. Hardly a pillar of engineering efficiency!

Again, in many ways the visual system is designed even worse! Imagine if I told you that you had to look at this page upside down and backwards, through a thick, yellow liquid and a layer of tissue that had a big, black spot in the middle! It doesn't sound easy, but that is the problem that your visual system faces. Light enters the eye through the cornea where it travels through some fluid and is then inverted by a lens and projected on the retina at the back of the eye. However, between the lens and the retina there is a large chamber filled with a jelly-like substance. Once the light gets to the retina, the real problem starts. The retina is backwards! The photoreceptors are not at the front surface of the retina, but are hidden behind a number of layers of opaque cells. Finally, there is a hole in your retina at the point where the optic nerve leaves the eye.

At this point it is reasonable to ask the question, if sensory detector systems are so bad, then how do I see and hear as well as I do? The simple answer to this question is that your brain does it.

Your external sensory receptors account for only a small part of your sensory systems. Each sensory system has specific brain regions that process their information. In fact, more than half of your cerebral cortex is involved in sensory processing, and the visual system occupies the lion's share of that. Why is this the case? If sounds, sights, surfaces, and smells exist, why does it take so much brain to detect them? The problem is that they don't exist; your brain makes them up for you. Let me use an example to illustrate this point. A notorious pastime of vision researchers is to ask someone how far they can see. Often, when people realize that they can see the sun and the stars, their answer is on the order of light years. The fact is that you don't see any great distance at all: light travels to you, so you "see" no further than your retina! Nothing seems more obvious to us than the observation of "things" out there in the world. But when you think about it, nothing could be further from the truth. All that your retina "sees" is light falling on its two dimensional surface, all your skin "feels" is mechanical stimulation, all your nose "smells" or your tongue "tastes" is the concentration of chemicals, all your ears "hear" is vibrations in air. No colour, no touch, no flavour, no sound, no pleasure, no pain. It is all manufactured for you by your brain.

If that doesn't convince you that the brain is creating your world, try this. Remember that I told you that the retina has a hole in it. This hole is actually called the optic disk and it is an area near the centre of the retina where there are no photoreceptors, so there is a "hole" in your visual field. It is unlikely, however, that you see a black spot as you read these words. Why not? The short answer is that your brain hides the hole so that you don't notice it. You can demonstrate this point to yourself by visualizing your blind spot. First, close your left eye and stare at a point in front of you on a surface at arms length from your eye. Then move a pen from left to right with your left hand across the middle of your visual field. At first you will be able to see the entire pen, but as you move the tip of the pen into your blind spot, the tip will disappear and then

reappear again as it exits the other side of the "hole". You will notice that the "hole" is not black, but rather blends into the surroundings of your visual field. It is as if your brain airbrushes this hole out so that it looks like the surrounding picture: a cunning forgery!

The fact of the matter is that sensation is not a simple task. It requires a huge amount of your brain to produce meaningful sensations out of nothing more than energy. The brain does such a good job of this that we are convinced that sounds, sights, smells, tastes and objects actually exist.

Maybe Dr. Seuss knew what the brain was up to when he created the character of The Cat in the Hat. At the end of the story the children are faced with the messy consequences of playing with the cat:

Then our fish said, "Look!
Look!"
And our fish shook with fear.
"Your mother is on her way
home!
Do you hear?
Oh, what will she do to us?
What will she say?
Oh, she will not like it to find
us this way!"
(. . .)
And THEN!
Who was back in the house?
Why, the cat!
"Have no fear of this mess."
Said the Cat in the Hat.
"I always pick up all my play-
things
And so. . .
I will show you another
Good trick that I know!"
Then we saw him pick up
All the things that were down.
He picked up the cake,
and the rake, and the gown,
And the milk, and the strings,
And the books and the dish,
And the fan, and the cup,
And the ship and the fish.
And he put them away.
Then he said, "That is that."
And he was gone
With a tip of his hat.

In many ways, the cat does the same job for the children in the story as your brain does for you. It cleans things up and leaves you (and your mother!) with the impression that everything is nice and organized no matter how messy things were before.

Chapter 22

Nurture's Little Secret: Use it or Lose It

Glen Prusky

> At birth our death is sealed, and our end consequent upon our beginning.
>
> — Marcus Manilius, 1st Century A.D.

It is unlikely that Marcus Manilius was commenting on brain development when he made this statement nearly 2000 years ago. However, nervous systems basically follow the rule that "their end is consequent on their beginning". As a student, you rely on your brain every day to get you to the university and back home again. Yet as mundane as this seems, not even that simple task would be possible if your brain did not have appropriate experience during the first few years of your life. This chapter explores the issue of why this experience is so important to normal function and how this process works.

Before these questions can be answered, it is important to understand some of the functional properties of the nervous system. The human brain is composed of many individual cells called neurons. Neurons contain the same structures as other animal cells, including a nucleus and a membrane, but they differ in their form and function. For example, the shape of neurons is very distinctive, due to the many small branch-like fibres that can extend for great distances from a central cell body. These fibres are divided into two functionally-distinct types.

One type of fibre, called a dendrite, is specialized for receiving chemical information from other neurons and converting it into electrical signals, and another type called an axon, is capable of generating electrical signals and sending chemical information from one neuron to another. Although the size and shape of neurons vary widely, most neurons have many dendrites and only one axon. However, because both dendrites and axons can branch widely, one axon can send the same message to many different

neurons, and a single dendrite can receive information from many different neurons. Neurons communicate with each other at points of close proximity between axons and dendrites, called synapses.

At a synapse, electrical activity in the axon of one neuron releases a chemical called a neurotransmitter that interacts with receptive sites on the adjacent dendrite of another neuron, and either stimulates the cell or inhibits it. Individual neurons or groups of neurons with specialized functions form precise networks of specific synaptic interconnections, and the combined activity of these networks forms the basis of our behavioural interactions with the world through our sensations, thoughts and actions. This elaborate organization in the nervous system comes about by a series of highly ordered steps, with a precise temporal sequence that is characteristic of each individual neural network. Moreover, each neuron in each network forms synapses only with a limited subset of potential partners and these connections are formed only at specific regions of the partner's dendritic surface.

At first glance, it would appear that the organization of the brain is so intricate and predictable in its development that it must be predetermined in a genetic plan. The problem is that it is not possible for the genome to prescribe this complexity because, as voluminous as it is, it does not contain enough information. It is estimated that there are approximately one million genes in the human genome, but, there may be as many as one billion times that many synapses in the mature brain! If the genome alone can't produce this complexity, then what does? The answer is that the brain works very hard to organize itself through a combination of active destruction and construction.

First, it is obvious that your body started off in life small and got larger as you developed. You are taller today than you were when you were born because you have added many new cells to your bones, skin and other tissue. Brains also start off small and get bigger as they age, but they actually lose neurons during development. In addition, many of the connections between the neurons that you are born with are also lost during development. Why would a brain that needs to develop complexity in its function dismantle itself? The fact of the matter is that it does not dismantle itself; it just removes the neurons and the connections that it does not need. It is not that different from creating a statue. There are at least two ways to create a statue. In one instance, a statue can be built out of clay by adding one piece at a time in all the right places to assemble the form and features of your subject. However, it could also be built by starting with a block of marble, much as Michelangelo did with the statue of David, and removing all of the stone that is not part of the form you desire. The brain does the latter; it begins with many inappropriate neurons and connections and unmasks its mature form by removing them.

Second, brains build stronger connections between neurons during development to enhance function. The mature nervous system has fewer neurons and connections than it did in infancy, but it often has more efficient synapses between the remaining ones. For example, the neuronal networks that are responsible for activating your muscles in infancy must become much more effective as the organism matures in order to control the increased size and complexity of the musculature. If not, as an adult you would have no more ability to move than you did as a child. It is not different from managing your money more efficiently as the size of your family increases.

The main question, though, is what rule does the brain use to determine which

neurons live or die and which connections are lost or strengthened in this process? For the most part, the rule is that neurons and connections between neurons, whose activity is associated with uncommon stimuli are lost, and neuronal connections associated with common stimuli are retained and enhanced. In other words, the functional properties of neurons are, to a large degree, molded by environmental experience. You start off life with more neurons and connections than you need, and then you maintain and strengthen the ones that are useful based upon your experience, and discard the rest.

Examples of this process have been well documented in sensory systems, the interface between you and your environment that provides the brain with most of its activity. If these systems are sufficiently active during development, your brain acquires the ability to use them effectively at maturity. If, on the other hand, your sensory systems are deprived of their normal environmental stimulation, the activity of the brain will be diminished and your sensory abilities will be permanently curtailed. The brain simply reflects the environment that it matures in.

An interesting feature of this process is that it is not equally active throughout development, but primarily occurs during a restricted period of time early in development called a "sensitive" or "critical period". As a result, at these critical stages of development, the precise organization of the brain is dependent upon specific interactions between an organism and its environment, and this influence of the environment on the brain, and therefore on behaviour, changes with age. This process means that abnormal environmental experiences usually have more profound effects on the maturation of the brain and on behaviour during early stages of postnatal development than in adulthood. One notable exception to the principle of restricted critical periods, are the neural networks responsible for many forms of learning. These networks seem to have the ability to change throughout life to facilitate the acquisition of new information. It is this exception that you are probably counting on to get you through university!

There are probably many of reasons why critical periods of neural flexibility exist. First, it allows the nervous system and its behaviour to reflect the environment that it will live in. There is no point in having a constantly changing nervous system throughout the course of life. Basic features of nervous system function probably need to be laid-down first, so that more complicated functions can be built on them. For example, complex language function, which requires accurate sensory information, could probably only come about if reliable sensory function is established first. Second, neural flexibility is probably biologically adaptive. Species that can adapt to changing environments have probably been more successful during evolution. Humans occupy more niches than any other species on the planet and it is probably the incredible flexibility of our nervous system that has played a large part in our ability to live in many different environments. Third, neural flexibility is a way to increase the complexity of the nervous system without using more genes to do it. The competition for resources has increased through, making it a more complex undertaking to survive as a species. Undoubtedly, increased complexity in brain function and behaviour is an advantage under these circumstances.

Finally, imagine for a moment that a terrible accident occurred and your brain was damaged beyond repair. However, a new brain could be cloned from your cells, and this brain could be placed back in your skull and you could "survive". (This is a common occurrence according to the pages of the *National Inquirer*!) Would you be

the same person? Of course not! The neural connections in your new brain would not be the same as your old brain because they would have to be modified in order for your nervous system to work efficiently. But because the environment that you are in now is different from the environment that your old brain developed in, you would sense things differently, perceive things differently, think differently, and act differently and your memories would be all gone. Genetically, you would have the same brain, in the same body, but you would be a different person, and there would be no way to turn you into the old you.

Take care of your brain.

Chapter 23

Language, Evolution and Human Uniqueness: On How to Explain Really Complex Things

Drew Rendall

Depending on your orientation, psychology is either the science of human behaviour or the science of the human mind (or better yet, both, since the two are inter-related). Either way, language is central, because language mediates much of human behaviour and because language recruits a variety of cognitive processes. Now there are at least two broadly distinct components to a science of anything: [1] sorting out exactly what the thing is, and [2] sorting out how the thing came to be. Both components of any science involve the use of general principles and processes that are applicable to a wide array of things (phenomena). In the life sciences, the sciences of all things biological (of which psychology is a part), component [1] reduces to sorting out how the thing is structured, how it develops, and how it functions, while component [2] reduces to sorting out how the thing evolved, in both cases using general principles and processes of evolutionary theory, the reigning theoretical paradigm for the explanation of biological phenomena.

Right away contemporary psychological science risks running aground because in some quarters there persists the notion that humans are unique, somehow qualitatively distinct from the rest of life on Earth . . . and therefore not subject to general principles and processes that govern biological phenomena. The obvious proof of our uniqueness lies in language, the pinnacle of human complexity, unrivalled by anything observed among animals. The implications for psychological science are clear: language is central to human psychology; a comprehensive account of human psychology requires an understanding of its evolution; but humans and language are unique. Thus, a comprehensive science of psychology is unattainable because one fundamental component

of that science—understanding how human behaviours and minds have evolved—is precluded. How can we avoid this pitfall?

The first step is to review evolution and our place in nature.

23.1 Evolution and Primate Phylogeny

Evolution is about *descent with modification.* *Descent* simply refers to the line of historical relatedness between ancestors and descendants that arises as a simple by-product of sexual reproduction that occurs in all biological organisms. Parents beget offspring, and one generation gives rise to the next, which gives rise to the next, which gives rise to the next . . . in an endless and connected sequence of reproduction. *Modification* simply refers to the fact that small changes in form and behaviour can occur over time such that descendants begin to diverge slightly in form and behaviour from their ancestors.

Evolutionary science allows us to use the patterns of similarity and accumulated change between organisms to chart their phylogenetic (i.e., evolutionary) relatedness. As we all know, humans are primates, members of a taxonomic grouping (the order Primates) that includes more than 200 living species of prosimians, monkeys and apes that share a common biological ancestry (Figure 23.1). Primates first appeared at the end of the Cretaceous (ca. 65 million years ago) when the extinction of the dinosaurs left many ecological niches vacant. The subsequent radiation of the mammals included the first primate species who successfully penetrated the arboreal (tree-living) niche.

As primates, we all share a variety of traits. That is, we can trace continuity between ourselves and all the species that share our taxonomic order. All primates, by virtue of our legacy of common ancestry, share a basic suite of traits such as forward facing (rather than laterally facing) eyes, pentadactyly (five fingers on hands and feet), opposable thumbs, and nails (rather than claws) on all digits. This suite of traits reflects our heritage of a life in the trees—it is an adaptive complex that allowed primates to exploit the arboreal niche. The five-fingered condition, combined with an opposable thumb and digits supported by nails rather than claws, provided improved grasping ability allowing primates to exploit arboreal food resources using a wide range of substrates (trunks, branches, twigs and vines). Forward facing eyes provided overlapping fields of vision (i.e., stereoscopic vision), adding depth perception crucial for moving and leaping in a complicated 3-dimensional environment.

These are traits that humans possess and that we share with all other primates. Of course, we share a great deal more with the species that are most closely related to us for the obvious reason that the amount of time that separates us from our last common ancestor with these species is shorter and thus the amount of change that has accumulated in our respective lineages is less. For example, we share a variety of additional morphological and behavioural traits with our closest living relatives, the great apes (orangutans, gorillas, and chimpanzees). Immunological analyses of gene products (e.g., blood proteins and other macromolecules) as well as direct studies of DNA itself indicate that chimpanzees are the species most closely related to us and that we shared a common ancestor as recently as 4.5 million years ago (R. Martin, 1990). Not surprisingly, then, the list of traits we share with chimpanzees is long and

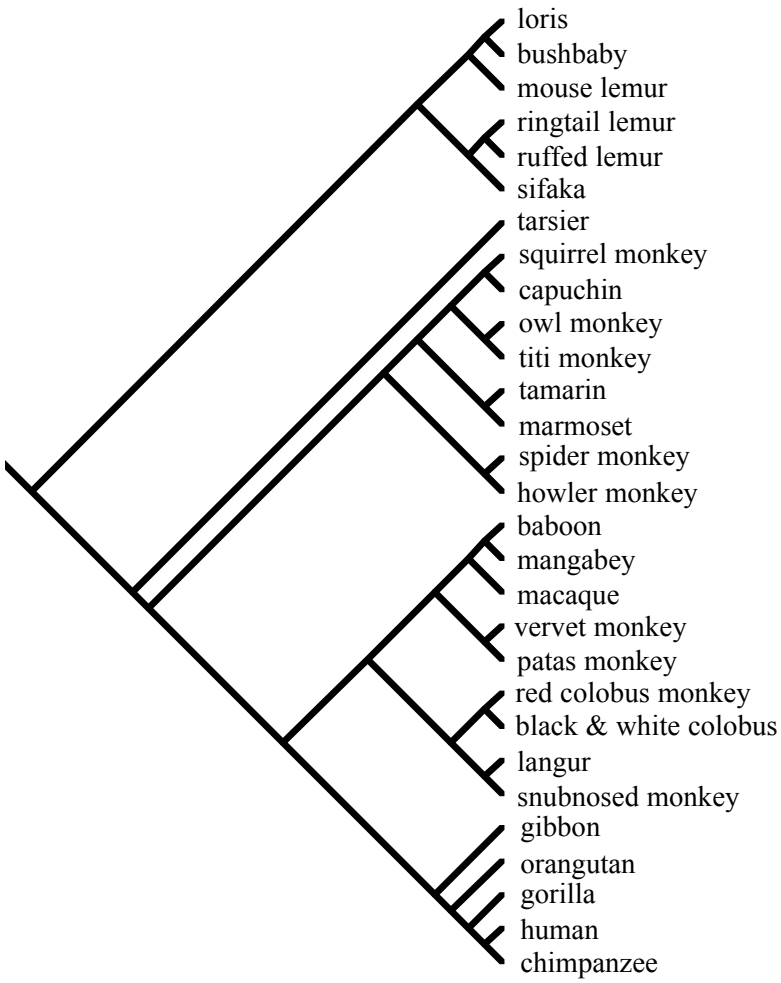


Figure 23.1: Phylogeny of the primates.

includes: a long life span, a long gestation period, a long period of infant dependency, large body size, large brains, a flexible rotary shoulder, an omnivorous diet including significant meat content, large social groups, a patrilineal social structure, similar facial expressions, tool manufacture, hunting, warfare/genocide, and promiscuous mating.

23.2 Explaining Complexity

Thus, the history of humans can be traced clearly to primates and the path is well-marked with evolutionary breadcrumbs—signs of anatomical and behavioural continuity. Despite the overwhelming evidence of our evolutionary history, proponents of human uniqueness resist evolutionary accounts of language. They frequently concede evolutionary continuity in other aspects of our anatomy and behaviour - continuity of the sort just reviewed—but maintain that language is unique, that it has not evolved. In their view, language is simply too complex to have evolved. It is, in fact, not a biological phenomenon at all, but a uniquely human cultural phenomenon.

It is ironic that proponents of human uniqueness choose complexity as the basis for arguing the non-biological basis of language, because complexity has in fact long been considered a hallmark of evolved biological phenomena. In a recent book on language, Steven Pinker (1994) compares language to things like "eyes" that are "organs of extreme complexity and perfection", an expression coined by the grandfather of modern evolutionary theory, Charles Darwin (1859). Pinker's analogy is designed to confront the reader with the overpowering logic that language has in fact evolved using the very quality (complexity) sceptics invoke as proof that it has not. How does the analogy accomplish this? In short, there are only two explanations for extremely complex things: [1] chance; and [2] design. Eyes are very complex things—finely tuned optical devices composed of myriad different but integrated components all operating in synchrony. Eyes are composed of a protective outer layer (the cornea), an iris for varying the pupil aperture and thus light intake, a lens for focusing light on the light-sensitive retina that lines the back of the eye-ball, photoreceptive cells (rods and cones) for receiving the light, an array of nerve cells (ganglion cells) and supporting cell complexes (bipolar cells, horizontal cells, and amacrine cells) for conducting this information to the fibres of the optic nerve, and tiny muscles for varying the iris aperture and the lens shape (not to mention the network of capillaries for delivering blood and providing lubrication for the cornea). Taken together, the various components that make up eyes are highly improbable arrangements of matter. That is, they are extremely unlikely to have come together by chance. The possibility that something as complex as an eye could have arisen by chance is, as Pinker and others before him have remarked, as imponderable as supposing that a strong wind might blow through a junk-yard and from bits and scraps of metal miraculously assemble a fully functioning aircraft. The alternative is that eyes—and other complex and improbable arrangements of matter—arose by design, which in the case of biological phenomena means by evolution, through a long, slow, gradual process of tinkering that accumulates minor modifications of form and function.

But exactly how does evolution generate organs of extreme perfection? Recall that evolution is about descent with modification. It is a logical truism that the organisms

that we see today in the forms that we see them in, are here today in those specific forms because their ancestors successfully reproduced and passed along the traits that aided them in doing so (this is the inescapable logic of Darwin's mechanism of evolution, aptly termed "Natural" Selection). Not all individuals are identical to one another in form or behaviour (another inevitable by-product of sexual reproduction, the shuffling of genetic material every generation). And not all individuals that are born reproduce successfully: some die at birth, some die early from disease, others are preyed upon. Those that survive and reproduce are those that were better able to procure food resources, to avoid predators, and to withstand disease, perhaps because they could run just a little bit faster, or, in the case of eyes, because they could see just a little bit farther or with slightly greater acuity. Whatever slight advantages these individuals possessed would be passed on to their offspring and the process continued with each subsequent generation potentially adding to the slight improvements of the previous ones. Over vast periods of time—literally millions of generations—evolution can produce highly complicated structures, so-called organs of extreme perfection (such as eyes), simply out of the slow, gradual accumulation of minor modifications of structure that provide slight improvements in function.

So, the complexity of a phenomenon, rather than being viewed as an obstacle to evolutionary explanation, is in fact a sure sign that something has evolved—complexity is one mark of an evolved biological phenomena. The only alternative is to argue something like "a strong wind blew through a junkyard and assembled an aeroplane".

23.3 How to Recognise Biological Phenomena

But what about the claim that language is qualitatively different from anything seen in animals and therefore that it must be a cultural rather than biological phenomenon? Certainly the lack of obvious signs of continuity between human language and animal communication systems makes the problem of reconstructing language evolution more difficult. But this should not preclude the possibility that it has evolved. Consider another distinctive feature of human behaviour, namely bipedalism, the characteristic way that we humans stand, walk and run on two legs. Our two-legged posture, like language, is qualitatively different from anything seen among primates, or indeed among all other mammals. Yet this fact does not motivate people to attribute walking to culture. They do not view walking as some wonderful cultural invention. Rather, they have no trouble believing that the capacity to walk bipedally is a part of our biology and like other aspects of our biology that it has evolved. In fact, language appears to fit the bill of an evolved biological phenomenon equally as well.

In his book on language, Pinker highlights a number of diagnostic features of evolved biological phenomena: they are marked by complex design properties; they are supported by specialised structures; they are universally present in the population; and they follow a canalised (i.e., regular) developmental sequence. Clearly, bipedalism fits the bill. Standing and walking upright is a complex way to balance oneself and get around. In fact, bipedal walking is often regarded as 'controlled falling' (doubtless many of you have had occasion to challenge the complexity of bipedalism and averted falling with varying degrees of success). Bipedalism is supported by specialised de-

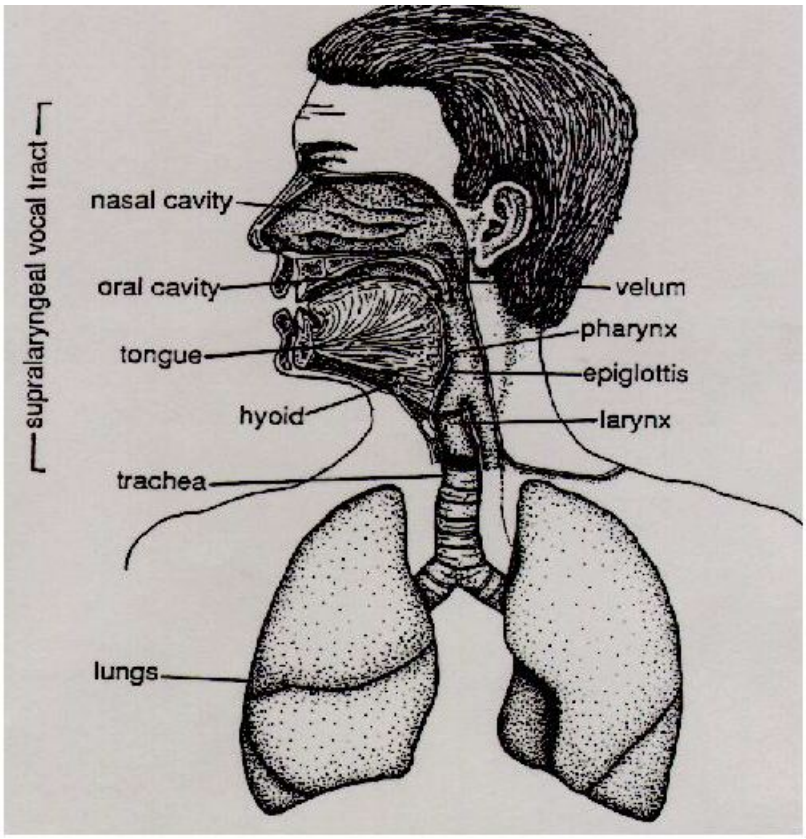


Figure 23.2: Anatomy of the human vocal tract.

sign of the bones of the pelvis, femur and tibia. Our pelvis is short, broad and rounded (relative to those of quadrupedal, or four-legged, animals) in ways that optimise the biomechanical efficiency of attached muscles responsible for maintaining an upright posture and moving our legs. Our femurs (thigh-bones) are angled inward slightly, which shifts our centre of gravity while walking to a more stable central location beneath us (versus to one side or the other). And the articulation of our femurs and tibias (the knee joint) can "lock" allowing us to remain upright with minimal muscular exertion. Walking is universally present—all humans walk upright. And walking follows a canalised developmental sequence: no child can walk at birth, but all acquire the skill with no formal instruction.

Language also fits the bill of an evolved biological phenomenon. Language is indeed complex (on this, we all agree). It uses arbitrary combinations of sounds (words) to stand for (or symbolise) objects and events in the environment. Typically, many such symbolic sound combinations are concatenated into lengthy utterances whose sequential ordering is governed by rules of syntax and grammar that determine their meaning. For example, the statements "Joe ate crow" and "Crow ate Joe" use the same

words to describe very different events whose meaning is interpretable only from the ordering of the elements. Language is supported by specialised structures (see Figure 23.2). We produce speech through the regulated control of airflow across the vocal folds of our larynx (or "voice-box") that sits atop the trachea (the airway leading from the lungs). The resulting air pressure waveform caused by the vibration of the vocal folds is filtered by the cavities of the supra-laryngeal ("above the larynx") vocal tract (neck, mouth and nasal cavities). Movements of the jaw, tongue and lips (the speech articulators) shape the filtered laryngeal waveform into the various vowel and consonant sounds of speech. The wide range of speech sounds we produce is, in part, attributable to the location of our larynx deep in the vocal tract, a position to which it descended only in the last 100,000 years from higher in the vocal tract of primate ancestors. The low laryngeal position in humans greatly increases our risk of choking as water and food particles now slip more easily down the trachea (to the lungs) versus down the oesophagus (to the stomach) where they are supposed to go. However, this cost was more than compensated by linguistic benefits (Lieberman, 1991). The low position of the larynx creates a large resonating cavity (the pharynx) at the back of the mouth and top of the neck that allows us to generate a large variety of speech sounds via movements of the jaw and tongue (which alter the resonating properties of this cavity). Because we can rapidly and dynamically reposition the jaw, tongue, and lips during speech production, we can produce upwards of 40 different speech sounds per second. This rapidly changing speech stream is decoded in listeners first by peripheral sense organs and then by central neuro-cognitive mechanisms. The peripheral auditory sense organs (our ears) capture and transduce the rapidly changing air pressure waveform that constitutes speech into mechanical vibrations in the middle ear and ultimately into electrical impulses in the cochlear nerve of the inner ear (via several intermediate transductions). Cochlear nerve impulses are transmitted to the brain's temporal lobe for additional central processing. Central speech processing operations—such as the phonetic, phonological, syntactic and semantic analysis of speech utterances—appear to have specific localised seats in the brain lateralised to the left hemisphere. For example, specific areas in the left frontal cortex (inferior frontal gyrus, roughly Broca's area) and the left tempero-pareito-occipital junction (posterior superior temporal gyrus, angular gyrus, and supra-marginal gyrus, roughly Wernicke's area) participate in syntactic/phonetic and semantic processing, respectively (see Figure 23.3). Thus, a great number of physical structures of the body participate in language and show evidence of specialisation for this purpose.

Language is also universal. Cultural practises vary markedly around the world. However, all human populations regardless of their culture possess a language that is as rich and fully complex as any other. Finally, language follows a canalised (i.e., regular) developmental sequence. No child anywhere speaks at birth, but the ability emerges reliably in all children around the world in all languages at about the same time . . . and all without formal instruction.

Clearly then, language, like bipedalism, belongs in the realm of the biological, in the realm of things that have evolved. Why then are some people inclined to give two different explanations to two equally distinctive characteristics of human biology—walking and talking? Why do they link walking to biology, but talking to culture?

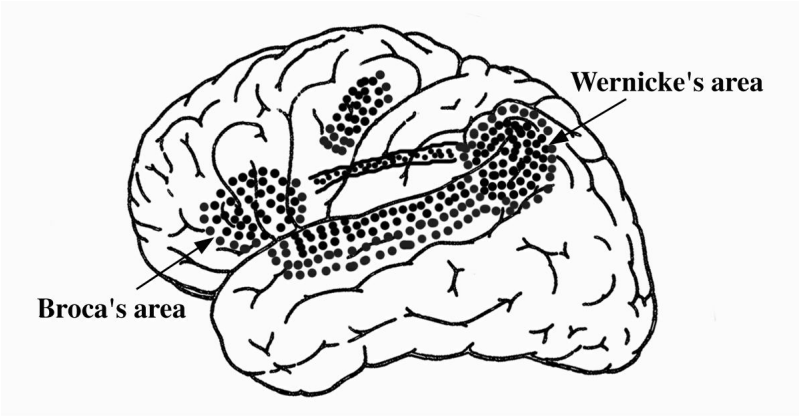


Figure 23.3: Left cerebral hemisphere of human brain with language centres indicated.

23.4 Cartesian Dualism

The view that language is uniquely human can be traced to age old moral pre-occupations having to do with distancing ourselves from the other "beasts" that inhabit this earth, at the same time celebrating our exalted special human status, both practises serving to justify our dominion over and exploitation of everything on this planet. If we aren't special, how can we morally justify our consumption of everything?

Cartesian dualism provides probably the best known rationalisation for this position. Descartes argued that animals were mere automata, organic robots that acted reflexively in response to sensory stimulation, whereas humans were distinctly different, possessed of a rational soul capable of acting on free will (i.e., possessed of a "mind"). Descartes arrived at this view via a circuitous route, however. Ironically, Descartes was initially inclined to the view that human and animal behaviour might be governed similarly by a large set of mechanical reflexes. This possibility occurred to him after viewing an exhibit at a local fair in France that contained a variety of hydraulic human statues connected by pipes to hidden platforms throughout a visitor viewing area. When visitors stepped on one of these platforms, water flowed through the pipes and activated the statues that then moved, made sounds, and behaved in ways that were to all appearances complex, internally motivated, in short just like humans! If a simple system of pipes could produce such complex actions, then perhaps a great deal of animal and human behaviour (however complex it appeared on the surface) could be accounted for by a similarly simple system of mechanical reflexes.

However, the social climate at the time (mid-17th century, the time of the Inquisition) was ill suited for a view that posited identity in behavioural mechanism between animals, machines, and humans (particularly one that appeared to dispose of free will). Descartes pursued a different tack instead. In an exhaustive series of thought experiments designed to doubt everything in experience that was not basic, primal and self-evident so as to arrive at a core of experiential axioms, he explored the fallibility of bodily experience but the veridicality (truthfulness) of thought experience. In brief, he

argued that our bodily experiences can all be doubted because they are brought to us through our senses, which we know to be imperfect. After all, we frequently mistake people on the street. We sometimes even hear unspoken voices (incidentally, Descartes is rumoured to have suffered from schizophrenia). Because all bodily experience derives from sensory experience and because our senses are infinitely fallible, Descartes argued that our very bodily existence can even be doubted. However, at bottom what cannot be doubted is doubt itself, the fact that we have doubts. That is, we cannot doubt that we are a physical being simply by virtue of the fact that we have doubts, or more generally thoughts. This is the cause for Descartes' famous refrain, "Cogito, ergo sum." [I think. Therefore, I am.] The logical conclusion to Descartes was that mind and body must be of different substances. Descartes held that the body consists of a physical substance that obeys basic physical laws (reflexes), but the mind consists of some incorporeal substance, or spirit, that acts independently according to completely different laws. This is the basis for Cartesian mind/body dualism (Descartes, 1637, 1641).

Descartes' dualism was dualistic in its implications. Not only were mind and body distinct, but so were humans and animals. Animals were mere automata governed by basic physical processes like those that moved the hydraulic statues at the french fair. Humans were unique by virtue of possessing a mind that operated independently of the body. In sum, our bodily existence was saved by our doubts, our uniqueness was saved by our minds... and Descartes was saved from certain excommunication and censure unlike his contemporary Galileo (who doubted that the Earth was the centre of the universe)[1].

23.5 Language and Mind

And what has language to do with Cartesian views of uniquely human minds? The connection lies in the relation between language and mind. Language is the medium through which we express our thoughts (of this, there is no doubt). According to some, the relation is reversed— that language *affects* rather than *reflects* our thoughts. The former position, known as linguistic determinism, proposes that language determines thought. That is, our impressions of the world are determined by the language we use (Whorf, 1956). For some, the connection between language and thought is so intimate that the two are taken to be synonymous—that we literally think in language. The latter views are not widely endorsed (that many of our most basic thoughts and feelings are difficult to articulate stands in obvious contradiction). However, they have served as impetus for serious considerations of the dissociability between language and thought, and for research programs designed to characterise the *language of thought* (or mind) as distinct from spoken language (Fodor, 1975).

Irrespective of which way the causal arrow points—whether language affects or reflects thought—language and mind are thoroughly interwoven. As noted above, we place great importance on the mind in defining who we are. In fact, as Cartesian du-

[1] Note that speculation over the connection between mind and body pre-dated Descartes, continued in his wake and is still a lively topic of debate today, though few take seriously the notion of a mystical mind spirit or the view that animals are reflexive automata.

alists maintain, the fact *that* we think lies at the core of what we "think" it means to be human. Given the need to distinguish ourselves from animals—the need to be unique—it is critical to argue that the mind and therefore language must be unique to humans.

And this is why two distinctive human traits—walking and talking— are often given two different explanations: because we are morally obliged to distinguish ourselves from the beasts, because the mind as the seat of our humanity is the best candidate for making this distinction, and because language is intimately connected to mind. Voilà! Human minds and language are unique. Walking simply doesn't have the same cachet, so it can be comfortably relegated to biology.

23.6 Morality versus Science

This is not the place to indulge the moralities of 'human uniqueness' viewpoints. For starters, morality can be a sensitive subject. But far more importantly, morality has little bearing on us here. Morals and ethics are the subject matter of Philosophy and other Humanities—how things should or ought to be. Science is about how things are, and how they came to be this way.[2] Furthermore, science seeks explanations through recourse to general principles and processes that apply to a wide range of phenomena.

It is immediately clear that uniqueness views are unscientific because they offer no explanation in terms of general principles that apply to a wide range of phenomena (Cartmill, 1990). In fact, the 'uniqueness' argument is logically equivalent to the 'chance' argument that Pinker uses as a foil for evolutionary explanations of complex, highly improbable arrangements of matter. Explanations based on 'uniqueness' or 'chance' attribute the cause of a phenomenon to completely unpredictable events. If the causes of a phenomenon are unique and unpredictable, then by definition we can never understand them. And so explanations in terms of uniqueness are really non-explanations, because they add nothing to our understanding. They tell us nothing about how the things we are interested in (in this case, human minds and language) came to be.

Instead, we seek a scientific explanation of how things are and how they came to be the way they are. In the life sciences—where the subject matter is organic life, the study of all things biological— the explanatory framework is evolutionary theory, which can account for the history of organic change and for highly improbable arrangements of matter in terms of general principles and processes that govern descent with modification.

Of course, it is conceivable that not all phenomena will submit to evolutionary explanation. In such cases, we can always resort to uniqueness or chance explanations of their origins. The latter are explanations of last resort, however, because they leave us with nothing to do in the meantime—they are explanatory dead-ends. Neither form of

[2]The fact that morals are not the province of science does not relieve scientists of moral responsibilities. On the contrary, the activities of science are governed by a serious ethical code regulating the conduct and reporting of research. Further, scientists play a pivotal role in shaping societal ethics. Their professional contribution lies in establishing the nature of focal phenomena and how they came to be the way they are as the critical first step in sorting out how they should or ought to be changed, if they can be.

explanation is morally better or worse than the other. But one is scientifically sound and productive in its assumptions, while the other is not. Starting from an assumption of chance or uniqueness suggests no means to proceed, no way to further understand the phenomenon in question because it is assumed to be the result of unique and unpredictable causes that are logically beyond study. The search for understanding ends before it can begin. In contrast, starting from an assumption of biological constitution suggests a very clear path: use general principles and processes of evolutionary theory that have proven applicable to a wide range of biological phenomena to try to learn how the thing came to be the way it is. This approach is productive—it gives us something to do!

23.7 Avoiding the Pitfall

Returning to the issue that framed this essay, "How can we hope to achieve a comprehensive science of human psychology—one that includes an understanding of how we came to be the way we are—if humans, their minds, and language are assumed to be unique?" By now, I hope the answer is clear. We can foster a truly comprehensive science of psychology if we abandon notions of human uniqueness and set about the task of understanding human biology in its entirety including the nature, function and development of human behaviour, human minds, and human language and how they have evolved.

Of course, one might still object, "why is it so important to understand how something has evolved? After all, isn't it enough just to know how it works now? And wouldn't this remove us from the problem that framed this essay?" It might, but at significant cost. Understanding how something has evolved is often an important source of insight into its current design, its current functions and its inherent limitations. Consider again, eyes. Eyes, as we've noted, are extremely complex organs, but they have certain limitations that reflect their evolutionary history (those reading this through eye-glasses will be instantly sympathetic). Our forward-facing eyes evolved for depth perception but at the considerable expense of peripheral vision (ubiquitous crumpled car-fenders testify to this expense). Eyes also come with a blind spot, a quirky result of evolution working with limited raw materials. Eyes work by focusing light on photoreceptive cells (rods and cones) of the retina, the light-sensitive lining that constitutes the back surface of the eye. Ganglion cells collect the unprocessed light information from these photoreceptors and conduct it to fibres of the optic nerve that then transmit the information to the optic lobe of the brain for further processing. Our blind spot derives from the fact that the optic nerve fibres must exit the eye somewhere in order to conduct light information to the optic lobe. They do so at the back of the eye, through a small hole in the retinal lining, where as a result there are no photoreceptive cells at all. Hence, visual information focused on exactly this spot on the retina simply cannot be seen—there are no light detecting cells to see it! Of course, it didn't have to be this way. Eyes could have been constructed without a blind spot, had we been designing them on paper with draftsman's tools, the way we build mechanical lenses. But evolution is blind, limited to working with the materials immediately at hand. The result, a complex seeing organ with specific limitations, a camera with a hole in the film!

Other examples abound. Primate hands are great for grasping (branches, bottles, baseballs), but are not so good for playing piano (why can't our thumb and pinky stretch further!), typing at computer keyboards for extended periods (witness carpal tunnel syndrome), or securing the last peanut from the bottom of the narrow-necked Planter's jar. Brains are good for solving social problems (If Sarah told Bob not to tell Sally about Suzy and Biff sleeping together, but Sally found out about Suzy and Biff, who's in trouble with Sarah?) but are much less proficient at math (If twice a number can be quartered and doubled and then squared and divided by itself but not change, what is that number?). And language is good for many things but it is comparatively clumsy in love. Why? Why are these complex organs good for some things but not others? The answer is that evolution designed them to solve particular adaptive problems using the materials presently at hand. Therefore, to fully appreciate how something works now, we need to know something about how it was designed and why—that is, we need to know something about its evolution.

So, if instead of assuming uniqueness (and thus immediately facing an explanatory dead-end), we assume that human minds and language are biological phenomena subject to evolutionary pressures that govern descent with modification, what should we expect to see in the communicative and cognitive systems of other species, especially other primates? We should see evidence of general processes similar in some respects to those that underlie human minds and language, but different to the extent that the species have diverged over evolutionary time and thus elaborated the processes to meet the different demands of their respective environments. Although the study of nonhuman primate communication is a comparatively new field, the evidence already amassed confirms just that.

23.8 Nonhuman Primate Vocal Communication: Evaluating Continuity with Language

Early views of other primate (and other animal) communication systems, prior to systematic field studies in the wild, assumed that vocalisations were simply hard-wired, reflexive manifestations of internal states of arousal (e.g., fear, hunger, or aggression) that were unaffected by experiential influence and recruited only low level limbic mechanisms that regulate the emotions . . . in short, markedly different than human language (we can now well imagine what might have motivated such a view). More recently, serious research efforts have been undertaken on a variety of different species, and the results have radically changed this view. This research indicates that the functions, mechanisms and development of vocal communication in primates are similar in many respects to those of language. For example, there is now evidence that primates produce vocalisations that function symbolically to refer to objects and events in their environments in a fashion at least crudely similar to the words of language. The best known example comes from vervet monkeys (for a more thorough description of vervet monkeys, their system of vocal communication and comparisons with language, see Cheney & Seyfarth, 1990).

Vervet monkeys are small-bodied (4-5 kg) group-living monkeys that inhabit savannah-

woodland environments in east and southern Africa. Their size and habits expose them to high risks of predation. When the monkeys encounter a predator, they produce loud 'alarm' calls that alert other group members to the danger. The monkeys actually produce three acoustically distinct alarm calls in response to three different classes of predators that prey upon them: leopards, eagles (and other raptors) and large snakes (e.g., pythons). Contrary to early views, these alarm calls are not simply manifestations of a monkey's internal state of arousal, in this case, fear. Although predator encounters can obviously be frightening experiences, the monkeys produce exactly the same alarm call when they spot a leopard in the distance as they do when they encounter a leopard at close range, even though the two circumstances portend very different risks of predation (and thus likely fear). They produce an acoustically distinct alarm call when they spot an eagle, and paralleling the case of leopard alarm calls, they give the same eagle alarm call whether the eagle they have spotted is soaring overhead (i.e., hunting, and therefore a risk) or perched in a tree (i.e., not hunting, and therefore not a risk). Furthermore, when a monkey who is alone encounters a predator, it often does not produce alarm calls even though the risk of predation (and hence arguably fear) is much higher than when in the comparative safety of the rest of the group.

Instead, the alarm calls appear to function symbolically to denote the different classes of predators. Each of the different alarm calls precipitates a behaviourally distinct escape response appropriate to the type of predator. In response to hearing a leopard alarm call, monkeys on the ground dash to the trees and scramble to the safety of the small, terminal branches where leopards cannot go. In contrast, in response to hearing an eagle alarm call, monkeys in the trees rapidly descend from the terminal branches where they are most vulnerable to swooping eagles, while monkeys on the ground run for the cover of a dense bush. (Note that appropriate escape responses to leopards and eagles are diametrically opposed. The small branches of trees are a good place to avoid a leopard but a bad place to be for an eagle. Beneath a bush is a good place to hide from an eagle, but a good place to *find* a leopard!) Finally, in response to hearing a snake alarm call, the monkeys all stand erect on their hind legs and scan the grass around them looking for a snake. Importantly, it is the calls themselves that precipitate these distinct responses and not anything about the caller's subsequent behaviour or the context in which the calls are produced. Experiments in which researchers have hidden speakers and then played back pre-recorded versions of the three different alarm calls to groups of vervet monkeys in the absence of a predator indicate that the calls alone are sufficient to elicit the distinctive responses. Thus, like different human words, different vervet monkey alarm calls symbolise different types of African predator. Subsequent research has revealed similar capacities in other primate species and some non-primate species besides (e.g., Primates: ring-tailed lemurs- Macedonia, 1990; rhesus macaques- Hauser & Marler, 1993; and baboons- Rendall, Seyfarth, Cheney, & Owren, 1999; Non-primates: honeybees- Frisch, 1976; chickens- Gyger, Marler, & Pickert, 1978).

Vervet monkeys use of, and response to, their different alarm calls is not given from birth, however. Although young vervet monkeys appear capable of producing the different alarm calls (at least in broad outline), they do not use them entirely appropriately, nor do they respond to the alarm calls of others in adult fashion (Cheney & Seyfarth, 1990). At first, infants produce eagle alarm calls to a broad class of aerial stimuli, in-

cluding many small raptors that would not prey on vervet monkeys, many non-raptorial birds (e.g., storks, doves), and even such harmless objects as leaves falling from trees overhead! They produce snake alarm calls to long, thin objects on the ground such as sticks. At the same time, infants do not respond to alarm calls the way adults do—that is, with distinct escape responses for each predator type. Instead, they typically look around anxiously for their mother and run to her. Occasionally they behave quite inappropriately, like running under a bush in response to a leopard alarm call. Only gradually do infants begin to limit the range of stimuli that elicit alarm calls to just those that pose a predatory threat, and only gradually do they develop distinctive, adult-like escape responses. The developmental process by which vervet monkeys acquire these abilities is not definitively known. However, these findings clearly indicate that experiential influences do impact vervet monkey vocal communication (similar effects have been identified in other species) and thus that, as in humans, important aspects of vocal usage and comprehension emerge slowly over the course of development.

Additional research suggests that, like humans, monkeys' use of symbolic vocalisations is also based on underlying conceptual representations. In humans, different words, such as "house" and "mouse", are judged to mean very different things even though they sound similar, while the words "smart" and "bright" are judged to have similar meanings even though they sound very different. Our evaluation of what these different words mean is based on the underlying conceptual representations that they instantiate rather than on their acoustic similarity. (If judgements were based on acoustic similarity, then 'house' and 'mouse' would be judged more similar to each other than 'smart' and 'bright' are to each other.) Playback experiments similar to those described above have confirmed a similar phenomenon in vervet monkeys.

In addition to their alarm calls, vervet monkeys produce a host of other vocalisations in a range of other contexts. One specific context involves interactions between neighbouring groups who are invariably intolerant of one another. In this context, the monkeys produce two very different sounding vocalisations: when they first spot another group of vervet monkeys at a distance, they produce a call termed a "wrrr"; when another group approaches to close range, they produce a different call termed a "chutter". Hence, the monkeys have two very different sounding vocalisations that mediate the same context—interactions with other groups. Using a habituation-dishabituation design of playback experiments, researchers have shown that the monkeys respond to these two different calls not simply on the basis of their acoustic structure but on the basis of the underlying concepts that the calls represent (Cheney & Seyfarth, 1990). The habituation-dishabituation experimental design is a technique commonly used in psychology laboratories to test basic perceptual abilities. It is based on the principle that repeated exposure to a single stimulus produces a gradual decrease in responsiveness to that stimulus (i.e., habituation), while subsequent exposure to a different stimulus—one with different physical properties—typically precipitates a dramatic recovery in responsiveness (i.e., dishabituation).

Using this approach, field researchers reasoned that if vervet monkeys compare their calls simply on the basis of their physical (acoustic) properties, then an animal exposed to repeated presentations of one intergroup call, the "wrrr", should gradually cease responding (i.e., habituate), but should show a dramatic recovery of response upon subsequent presentation of the other intergroup call, the "chutter", because the

two calls are physically very different. If, however, vervet monkeys compare their calls on the basis of their underlying conceptual representations, then animals who stop responding after repeated exposure to "wrrrs" should show a transfer of habituation when subsequently presented with a "chutter" (i.e., they should show no dramatic recovery of response) because although acoustically very different these two calls mediate the same event— i.e., have the same underlying conceptual representation (having to do with other groups), like 'smart' and 'bright' above. Vervet monkeys tested in this manner in the wild did indeed transfer habituation between the two different intergroup vocalisations suggesting that their use of these calls is mediated by underlying conceptual structures perhaps similar to those of humans.

Finally, laboratory studies of vocal perception in monkeys have revealed a left-hemisphere bias for processing species-specific vocalisations. This is analogous to the left-hemisphere bias for processing speech in humans described earlier. As in humans, areas of the left temporal cortex of monkeys appear to be especially important in decoding salient features of vocalisations. Neuroanatomical studies have identified additional areas in the left hemisphere of monkey brains that are structurally homologous to speech areas in the left hemisphere of the human brain, although the functional role of these areas remains obscure (in monkeys as it does to some extent also in humans!).

23.9 Conclusions

Evidently, there are important points of continuity between human and nonhuman primate systems of vocal communication. Clearly the evidence reviewed does not establish equality between human language and nonhuman primate vocal communication. Nor should it. Evolution is about descent with *modification*—so we should expect to see some differences... in addition to similarities. And, in fact, many differences between human language and nonhuman primate vocalisations exist. For example, while some nonhuman primate calls possess a symbolic function similar in some ways to human words, the bulk of their calls apparently do not. Furthermore, whereas human language is intentional, in that we speak as a deliberate attempt to inform others, the same does not appear true of nonhuman primates. Calling animals do not seem to appreciate the effect of their calls on others and do not seem to call with the intent "to inform".

These differences do not negate the value of evolutionary explanation, however. The purpose of seeking a scientific evolutionary explanation is not to establish equality between human language and animal communication, or to somehow elevate animal minds and engender a deep appreciation for their abilities (though this might sometimes be a healthy side-effect), or to somehow debase or humble human minds, reducing them to the level of animals. It is not a competition (at least not to anyone but humans). The point of science is to contribute to a comprehensive understanding of who we are and how we came to be this way, however simple or complex, lowly or exalted that may seem to be.

Appeals to uniqueness do not provide that understanding. At different points in the past, the list of human 'uniquenesses' has been long. It has included such things as: humans are the only beings that manufacture tools, humans are the only beings

that hunt cooperatively, humans are the only beings that have language. The list has been seriously shortened. We now know that other primates (and many other animal species besides) manufacture tools, hunt cooperatively, and have complicated systems of communication. Of course, die-hard human uniqueness proponents are quick to move the goal-posts.They are quick to re-define each of these abilities in such a way that only humans can fit the bill. This can always be done. After all, evolution is about descent with modification, and so every species is going to differ at least a little bit even from the species most closely related to it. So, in some sense, it will always be possible to argue that X or Y species is somehow unique. But this is not productive. It doesn't tell us anything about how they came to be this way. The value in finding that some other species manufactures tools, or hunts, or communicates in complex ways, is not simply to match the ante of human abilities. The value in finding similar phenomena in other species lies in the contribution this can make to a general understanding of how such things come to be: how they are designed, what functions they serve, what problems they are designed to solve, and what their inherent design and functional limitations are.

In the case of human minds and language (and other complex things) we are still a long way from understanding how they came to be, what all their functions and limitations are, and what problems they are designed to solve (though we know some of the problems they are *not* designed to solve…recall math and love, above). However, the evidence of continuity between human and nonhuman primates in at least some aspects of communication and cognition is exciting, for it promises to contribute to an understanding of how human language and mind came to be the way they are. It promises to contribute to a comprehensive science of psychology.

Chapter 24

Competition, Cooperation, Social Cognition and the Evolution of Human Psychology: Why Do We Think and Do What We Think and Do?

Drew Rendall

A salesman comes to your door and offers you a set of Wilkinson ever-sharp kitchen knives and a certificate redeemable for a weekend for two at a condominium resort in Fairmont Hot Springs, all for only $100. You pause. Sounds good. Too good, in fact. You think to yourself, what's the catch? You have serious doubts. Why?

In 1977, Craig Packer reported a study of male baboons living in east Africa (Packer, 1977). In 1,100 hours of observation, he noted 20 instances in which a high ranking male, consorting with a sexually receptive female, was confronted aggressively by a coalition of two subordinate males in the group. On six of these occasions, the coalition was successful in displacing the consorting male, one of the two males of the coalition threatening, chasing and otherwise distracting the consort male, while the second male attempted to copulate with the female. This was a landmark finding in the behavioural sciences published in one of science's most prestigious journals, *Nature*. Packer's results were remarkable not because we knew so little about the behaviour of wild baboons. On the contrary, baboons are one of the best studied species of non-

human primate, and at that time a great deal was already known about their natural behaviour. Packer's results were remarkable because they appeared to represent a bona fide instance of "true" cooperation. Why should this be so remarkable?

24.1 Competition: The Default Behavioural Program

Darwin stands out as one of the most important scientific thinkers in recent history, his single most important contribution, of course, being his theory of evolution by means of natural selection that was published in 1859 under the short title, *The Origin of Species*. Prior to Darwin, many scholars had debated the diversity of life on earth and had speculated about its origins. Why were there so many different types of organisms? Why were they distributed around the globe the way they were? Why were there so many in some places and so few in others? And why were some found only in one place and others only in another? A variety of scenarios had been proposed, but none offered a systematic account that could make sense of the similarities and differences among organisms, nor the reasons for them. After decades of careful research, and only then at the considerable coaxing of his colleagues, Darwin published his hypothesis (subsequently labelled a theory), which was thorough, profound and best of all, wonderfully simple.

The gist of it is as follows: organisms need a variety of resources to survive and reproduce (e.g., food, water, and refuge); such resources are in short supply and organisms are capable of reproducing at rates that quickly outstrip them; hence, life is a struggle in which (obviously, or "naturally") the more skilled competitors secure more resources for survival and reproduction; more skilled competitors will thus be favoured, or "selected" evolutionarily, in that their progeny will be more numerous in succeeding generations; whatever traits that make certain individuals more successful in competition for limited resources will be preserved among their descendants (i.e., offspring resemble their parents). Simple.

These are the core elements of Darwin's proposal and they help to account for the diversity of life on earth and its causes because, as you might expect, the particular traits that are favoured in a given environment will depend on the conditions that prevail in that environment and they will be different from the traits that are favoured in a different environment where different conditions prevail. Any given environment offers a variety of ways to make a living (a variety of 'niches'), and selection will favour organisms that compete most successfully in some one particular niche, leaving other organisms to compete in a different niche. Over time, organisms differentiate along lines compatible with their different niches as selection favours changes that improve each organisms' ability to compete in its particular niche. Some (rich) environments offer more possible niches and so sustain greater numbers of different organisms, other (poorer) environments offer fewer niches and thus support fewer different organisms. Still other environments offer completely different sorts of niches and so support completely different types of organisms. In this way, "from so simple a beginning endless forms most beautiful and most wonderful have been and are being evolved" (Darwin,

1859: closing sentence of the *The Origin of Species*).

In the century and a half since Darwin's proposal, these basic points have been confirmed over and over again in all manner of organisms, animals and plants alike. Darwin's theory of evolution by means of natural selection has had an increasingly broad impact such that it is now embraced by virtually all fields of natural science, including psychology.

The central upshot of Darwin's work that concerns us here is that competition is a ubiquitous feature of daily life for all organisms because natural selection favours individuals who procure even slight advantages that translate into greater reproductive success. Among nonhuman primates, including baboons, competition is seen in many arenas. For example, individuals compete for limited food resources. Fruit, a staple food for many species, occurs in small clumps that are patchily distributed throughout a forest. Individuals in a group compete to gain and maintain access to fruit patches, displacing those already in the patch and repelling others who try to join. A consistent result of such competition in many species is the development of fairly rigid dominance hierarchies within the group. Animals of high rank typically have priority of access to the best resources. Of course, adequate food supplies are critical to reproduction, particularly in females who must gestate, lactate and then care for offspring. Without sufficient energy reserves, females will not even ovulate. Compared to low-ranking females, high-ranking females often have higher reproductive success: they reach sexual maturity at a younger age, they have shorter intervals between successive births, they produce more babies during their lifetime as a result, and more of their babies survive to adulthood in order to reproduce themselves. Much of this reproductive advantage stems from superior resource access for high-ranking females. It can also result from direct reproductive competition among females. In some species, high-ranking females target aggression at low-ranking females and sometimes also their young infants. The resulting social stress (as well as direct physical injury) compromises the reproductive function of subordinate females to varying degrees (as noted above). In a few species, the reproduction of subordinate females is completely suppressed: there may be several adult females in a group, but only one—the highest ranking (or 'alpha') female—actually breeds.

Males also compete, most conspicuously for access to sexually receptive females, because male reproductive success is more limited by breeding opportunities than it is by access to food resources. Success in such competition favours large males who can physically overpower smaller males and thereby gain access to more females. Females also often prefer to mate with large, physically dominant males with a demonstrated ability to secure resources. Not surprisingly, priority of access to adult females translates into increased reproductive success for males, and over evolutionary time selection has so favoured large males in reproductive competition that they are considerably larger than females in many species and possess other morphological traits that equip them for reproductive competition, such as large canine teeth.

Other examples of competition abound and can be observed in virtually every facet of life. For example, low-ranking female monkeys even compete among themselves for opportunities to groom high-ranking females due to the possible benefits that such friendly behaviour might some day return to them in the eyes of these high-ranking females. In fact, so pervasive is competition that you might say it is the default be-

havioural program. Darwinian natural selection favours organisms that compete in their own self-interests. And that is why Packer's baboons seemed so remarkable. They were cooperating! . . . (if ultimately only to compete for an opportunity to mate)

24.2 Cooperation: The Evolutionary Challenge

Cooperation, in everyday parlance, typically implies individuals joining forces to achieve some mutual goal that is then shared. In this sense, cooperation is not remarkable and poses no problem for classic Darwinian natural selection favouring individuals who act in their own self-interests. It is common to see multiple individuals acting in concert toward some mutual goal whose benefits they all reap. Predator defence strategies of herd-living ungulates, such as muskox, are a familiar example. Faced with a predator, muskox aggregate in a tight circle. While an individual on its own stands little chance of defending itself, a tight circle of horned animals provides a more formidable defence. The normally vulnerable hindquarters are turned toward the centre of the circle, while the horned heads and front hooves face out and can be tossed to and fro and kicked outwards. In this way, the protective self-interests of each individual are served by the collective behaviour of many. In the field of animal behaviour, such cooperative behaviour is termed "mutualism" because all participants benefit simultaneously.

What Packer's baboons did was different. One subordinate male heeded the solicitations of a second subordinate male, and the two confronted a higher ranking male whom neither could have defeated on its own. At this point, the recruited male antagonised the higher ranking male, threatening and chasing him, while the recruiting male attempted to mate with the female. In this case, then, one male helped another at some cost to itself (in terms of time and energy but especially risk of injury attacking a high-ranking male). The other got an opportunity to mate. In short, one male paid a cost, the other received a benefit.

This sort of cooperation is often termed 'altruism' because it involves self-sacrifice, foregoing some opportunity oneself so that another may benefit, or (as in this case) actively helping another at some cost to oneself. You might think of it as pure or "true" cooperation in the sense that the helping behaviour is not motivated by any immediate self-interest. It is the mirror image of competition or selfishness. It is selflessness.

Perhaps obviously, such true cooperation—altruism—is not favoured by Darwinian natural selection. Individuals that help others at a cost to themselves will be "selected out" because they increase the ability of others to survive and reproduce at the same time that they reduce their own ability to do both. The simplest way to see the logic of this would be to imagine an individual who chose to forego eating when food was in short supply so that others might eat. Obviously, any individual who behaved so altruistically would suffer seriously and die, while those who did not forego eating would continue to survive and reproduce. However minimal was the reproduction of the latter individuals (given the acknowledged food shortage), it would still exceed that of the altruist who forewent eating entirely. The result: there would be fewer and fewer altruists in each succeeding generation until eventually they would be selected out entirely. This is an extreme example, of course, but the logic applies no less surely in all other cases. A small amount of feeding restraint would produce a small decrement in

reproductive success, but over sufficient time, this too would lose out to self-interested individuals who did not restrain themselves at all. Quite simply, altruistic behaviour (true cooperation) is not evolutionarily stable.

However, just as surely there do appear to be examples of organisms cooperating. How can this be? For a long time after Darwin, examples of cooperation were a puzzle for evolutionary theory. It was not until 1964 that a neo-Darwinian named William D. Hamilton resolved the problem (Hamilton, 1964). Like Darwin's original thesis, Hamilton's insight had broad impact but was elegantly simple. Hamilton pointed out that the entity on which natural selection operates is not, in fact, the individual but rather the gene. That is, when considering the evolution of any particular form of behaviour, what is important ultimately is not simply the reproductive success of the individual performing the behaviour but rather whether or not, as a result of the behaviour, the underlying gene guiding the behaviour increases its representation in future generations. And there are two routes by which genes can secure representation in future generations: directly, by the straightforward route of their bearer's reproduction; and indirectly, through the reproduction of relatives who share identical copies of the gene by virtue of descent from a common ancestor. Thus, genetic representation is more inclusive than that that results from an individual's own reproduction. It also includes a portion of the reproduction of an individual's relatives. This more inclusive measure of evolutionary success is termed, 'inclusive fitness'.

Here was a gene's eye view of evolution that made perfect sense (once realised) given the discovery early in the 20th century that genes are the building blocks of organisms as well as the units of inheritance. Offspring are genetic copies of their parents. In sexually reproducing organisms, 50% of an offspring's genetic makeup is a copy of its mother's genes, the other 50% a copy of its father's genes (owing to the fact that in the meiotic cell division preceding gamete formation the full complement of genetic material in each parent is halved in anticipation of being fused with a gamete from the other parent, in order to preserve in the fertilised offspring the original full complement of genetic material). In this way, members of a family are genetically related to one another. The degree to which family members are related varies and can be expressed in terms of the proportion of genetic material they share. For example, parents are related to each of their offspring by a factor of 0.5—that is, they have in common 50% of their genetic material. Full siblings (who share both the same mother and the same father) on average also have in common 50% of their genetic material. Half-siblings (who have only one parent in common) share only 25% of their genetic material. Cousins share only 12.5% of their genetic material. And so on . . . more distantly related individuals having in common a smaller and smaller proportion of their genome.

Another way to understand genetic relatedness is in terms of the probability that any two individuals will share a gene for a particular trait. In the case of full siblings (or parents and offspring), this probability is 0.5, matching the proportion of the total genome they share. In more distantly related individuals this probability declines. Among unrelated individuals, the probability is zero.

The solution to the cooperation puzzle in all of this lies in the observation that a trait that imposes some cost on its bearer can still evolve if it provides a more than compensatory benefit for a close relative (the more than compensatory benefit being

measured in terms of gene copies governing the behaviour in question). Imagine a case in which you could help your brother in such an important way that it allows him to successfully have and rear three children (say, for example, you could save him from drowning). Doing so cost you the ability to have one child of your own (because of some physical injury sustained in the process of saving your brother). The cost to you (in the sterile accounting of evolutionary genetics) is one copy of the gene governing this behaviour.[1] However, the benefit to your brother is three children, each of which has a 50% probability of carrying a copy of the gene for this behaviour. On average, then, the behaviour results in 1.5 copies of the gene governing it transmitted to the next generation through the reproduction of your brother. So, although the behaviour serves to decrease its representation in the next generation via its bearer's reproduction, it more than compensates for this loss by a greater increase in its representation in the next generation through the reproduction of relatives. Hence, the behaviour will be "selected for" . . . it will persist.

In other words, helping your kin is a lot like helping yourself. There will be certain circumstances in which helping them to reproduce—even at some cost to yourself—will actually serve to increase the number of copies of your own genes that are contributed to future generations. This insight translates into a generalisable rule—Hamilton's Rule—for predicting the circumstances under which cooperative behaviour of any particular sort should evolve. Basically, cooperative behaviour will be favoured by evolution when the benefit to the recipient of an altruistic act, discounted by the degree of genetic relatedness between altruist and recipient, exceeds the cost of the act to the altruist (formalised in mathematical notation: $rB > C$, where r = degree of genetic relatedness, B = benefit to the recipient, and C = cost to the altruist).[2] The above scenario is obviously an extreme example to illustrate Hamilton's Rule, but the logic applies perforce in cases where the genetic stakes are not so high, the only difference lies in the size of the numbers that will be involved in the accounting!

In sum, then, selection will favour cooperative, or altruistic, behaviour among genetic relatives, when it actually increases the number of copies of the altruist's genes

[1] Note that explanations that imply "genes *for* behaviour" are short-hand expressions used for expository convenience only. In no way does Hamilton's insight imply that there are literally genes *for* behaviour, or even more restrictively that there is a single gene for any particular behaviour. Hamilton's insight implies only that there is a genetic basis for behaviour generally, which we know to be true for a wide array of species including humans. The link between genes and behaviour is obviously complicated, involving many intermediate factors related both to the organism's biology and physiology and to the environment (and interactions between them). However, the link is no less sure and no less important because of these complexities. Therefore, the core of Hamilton's insight applies perforce, and explanations of it often make use of short-hand expressions that may imply "genes for behaviours" so as to avoid the convoluted wording otherwise needed to properly explain the complicated relation between them each time the issue is raised.

[2] Although discussion may seem to imply sophisticated decision-making processes even including mathematical analyses of cost-benefit ratios, no conscious evaluation or mathematical calculations are ever implied. Things like 'costs' and 'benefits' and formulas that relate them—like Hamilton's Rule—are concepts that we use in order to formalise our scientific analysis of behaviour and make clear the various assumptions we are using to do so. They do not imply that the organisms concerned possess anything like the same concepts, or go through anything like the same evaluative steps we do in analysing their behaviour. Thus, the implied "decisions" about when and with whom to cooperate are not meant to imply conscious, rational decision-making of the sort we find familiar. Rather, we assume only that organisms that behave in ways that maximise the benefits they receive in respect of the costs they pay will be favoured by natural selection, irrespective of the actual mechanisms that guide them in their tendency to behave in such ways.

contributed to the next generation via the reproduction of kin. Appropriately, Hamilton termed this evolutionary mechanism "kin selection". Somewhat ironically, this kind of cooperative behaviour is not really altruistic at all. That is, at bottom, it is not self-sacrificing, but rather genetically selfish, and in this sense, it is completely in keeping with classic Darwinian natural selection.

Hamilton's theory of kin selection was a critical addendum to Darwinian natural selection that has helped to refine modern evolutionary theory. In the 35 years since Hamilton's thesis was published (kin selection was actually Hamilton's doctoral thesis), innumerable examples have been documented of "cooperative" behaviour among kin (i.e., nepotism). In fact, among nonhuman primates, including baboons, kinship is one of the most important organising principles of social life. In many species, females remain in their natal group throughout life, while males disperse at sexual maturity to join other groups. Thus, females form the core of related individuals in a group. Accordingly, it is the females that show the highest levels of friendly affiliation and the highest levels of (kin selected) "cooperative" (genetically selfish) behaviour. In fact, in many groups, there are typically two or three distinct female lineages, or matrilines; hence, some females within a group are closely related to one another, while others are not. Females in such groups show a complex mix of cooperative and competitive behaviour, being affiliative toward related females in their own matriline and antagonistic toward unrelated females from rival matrilines. Related females feed together, rest together, and socialise and groom one another at high rates. They also support one another in aggressive encounters with females from rival matrilines. Unrelated females spend far less time together and are typically antagonistic when they do.

Males, on the other hand—being unrelated—are almost always competitive, often fiercely so. Fighting is common among males, often over access to females, and can cause serious injuries and even death. This is not invariably the case, however. Males are sometimes affiliative and cooperative. For example, in chimpanzees, it is the males who show high levels of affiliation and cooperative behaviour, while females are relatively intolerant of one another. However, in the case of chimpanzees, contra the many other primate species just noted, it is the males who remain in their natal groups and form the core of related individuals, while females disperse at sexual maturity. And this really is the clincher that relatedness guides affiliative and cooperative behaviour, for when the patterns of dispersal (and thus relatedness) among the sexes are reversed, so too are the patterns of affiliation and cooperation.

Kin selection makes an even more specific prediction, in fact, namely that the level of cooperative behaviour between individuals should vary as a function of the degree of relatedness between them: cooperation should be highest among close kin and decrease with diminishing degrees of genetic relatedness. After all, the proportion of the genome shared among distant kin is small; therefore, the benefit accruing to a relative far removed would have to be very high indeed to justify the cost of cooperation. Either that, or the cost of cooperation would have to be trivial. In contrast, between very close kin, comparatively small benefits (when only minimally discounted by the high degree of relatedness between participants) will exceed the cost of most cooperative acts. Exactly this pattern is observed. Cooperation is highest among close relatives, such as between parents and offspring, siblings, grandparents and grandchildren, and it decreases steadily with increasing genetic distance, as among cousins, and among

nieces/nephews and aunts/uncles.

Thus, cooperative behaviour in nonhuman primates (and many other organisms) conforms well to the predictions of kin selection: cooperative behaviour is seen primarily among close kin.

24.3 Reciprocal Altruism

But Packer's male baboons were not close kin, or related at all (at least they were unlikely to be). Baboons are one of the many primate species in which females form the core of the group and males disperse. So the adult males that ultimately end up residing in a group are normally unrelated. Hence, cooperation among male baboons does not jibe with the principles of either classic Darwinian natural selection, or Hamiltonian kin selection. If males are unrelated to one another, then the act of helping another at your own expense (in this case, the risk of serious injury) cannot provide any genetic returns. As we've noted, cooperation of this sort—true cooperation, or altruism—should not persist.

It was not long after Hamilton resolved one facet of the apparent cooperation problem, by observing that it obtained primarily only among kin, that another neo-Darwinian, Harvard biologist, Robert Trivers resolved the remaining paradox of apparent cooperation among non-kin (Trivers, 1971). He demonstrated that such cooperation could nevertheless accord with Darwinian principles so long as certain conditions were met. Briefly, cooperation among unrelated individuals can be sustained so long as the cost of helping another individual on one occasion is more than repaid by the benefit received from reciprocated help from that same individual on some future occasion. This was a sort of "I scratch your back, you scratch mine" type of explanation, and it was referred to as Reciprocal Altruism because it involved an iterated sequence of reciprocated altruistic acts. So long as the balance sheet remains in the black—that is, the benefits received by an individual outweigh the costs paid—it will pay individuals to cooperate. Here again, cooperation of this sort can be viewed as ultimately selfish because in the end each cooperating individual receives more than it gives. Nevertheless, it could be evolutionarily stable... in principle.

However, as Trivers acknowledged (and you might imagine, perhaps from personal experience), this sort of cooperative system is highly unstable because there is an ever-present temptation to defect—to avoid paying the cost of helping after having already reaped the benefit of being helped. In fact, extensive mathematical modelling and computer simulations were undertaken to test the conditions under which such a system of cooperation actually could evolve. Only a single scenario proved stable, a straightforward Tit-for-Tat strategy in which altruistic acts were reciprocated in the strictest fashion, individuals alternating back-and-forth between the roles of helper and helped. Under these circumstances, individuals will continue to support one another so long as each cooperates. However, at the first sign of defection in one, the other will also defect... and the whole system gets derailed.

Furthermore, notwithstanding its inherent instability, a system of cooperation based on reciprocal altruism challenges the odds because it also depends on several fairly stringent conditions, including: [1] the need for repeated encounters, so that altruistic

acts performed now even have a chance of being reciprocated in the future. Individuals will be reluctant to cooperate with strangers, for example, because the chance of meeting again and having one's help reciprocated is low; [2] the ability to remember past encounters and their details (whether or not others cooperated); [3] the ability to recognise specific individuals so as to distinguish those with whom one has interacted (and cooperated) from those with whom one has not; [4] the ability to detect and punish cheaters (if only by withholding future support). Collectively, this set of conditions describes a fairly complicated social problem.

And *these* are the reasons why Packer's study was so remarkable. Partners in the baboon coalitions alternated roles in apparently tit-for-tat fashion. Males solicited help preferentially from other males who had previously solicited their help. Here, then, was a bona fide instance of true cooperation, demonstrating that a stable system of reciprocal altruism was possible not only in principle but also in practise. By extension, it also implied that baboons must have the requisite social problem-solving abilities to support such a complicated system of reciprocity.

24.4 Primate Social Cognition

In fact, we now know that the basic conditions for reciprocity obtain in baboons and many other primates (as well as non-primates). Thus, primates are long-lived and often form cohesive social groups whose membership is stable, both conditions providing effectively infinite opportunity for repeated encounters with the same individuals. In fact, many species are very gregarious with high rates of social interaction. Individuals interact repeatedly every day and as a result form strong and enduring social relationships. The members of social groups also clearly recognise one another on an individual basis. Individual recognition is implied most basically in their linear dominance hierarchies: each individual occupies a specific social rank, higher than some and lower than others. Not only do individuals know their own rank relative to others, but they know the ranks of others relative to each other. For example, if in a group of monkeys animals ranked 4 and 7 are sitting together and animal 3 approaches the pair to be groomed, animal 7 invariably moves off while animal 4 never budges. Animal 7 evidently knows that it is the lowest ranked animal of the three and thus ought to be wary, while animal 4 knows that although it is subordinate to the approaching animal 3, animal 7 is more subordinate still, so it can stay put. Clearly, this behaviour requires the animals to be able to individuate one another in addition to knowing each other's rank.

In fact, monkeys are extremely sensitive to rank and its effects. Subordinate female monkeys sometimes compete subtly for opportunities to groom higher ranking females seemingly in an effort to curry favour with them. In return for such grooming, higher ranking females appear more tolerant of subordinate groomers in other contexts (e.g., feeding), and are even motivated to support them in subsequent agonistic encounters. The latter effect was demonstrated in an experimental study of wild vervet monkeys in which a concealed loudspeaker was used to play back the recruitment screams of certain females to other females in the group in the presence or absence of recent grooming between them. Recruitment screams are given during a fight to solicit aid from others.

Female subjects in these experiments were more responsive to the recruitment screams of other females if those females had just recently groomed them, suggesting that the strategy of subordinate females ingratiating themselves to dominants through grooming can be effective. (Incidentally, this too is an instance of reciprocal altruism—the exchange of grooming for help in a fight.)

More direct evidence for individual recognition derives from the fact that social relationships among group members are highly individualised. That is, relationships show special qualities depending on the identity of the participants. As already noted, affiliative and cooperative behaviour are highest among kin and vary according to degree of genetic relatedness, while aggression is most severe among nonkin. Individuals often also have preferred grooming partners and partners with whom they travel and socialise. In addition, individual males and females in some species sometimes form especially close bonds, termed friendships. Male and female "friends" remain close to one another throughout the day, sit and rest together, groom each other at high rates, support one another in fights with others, and may huddle and sleep together at night. Such friendships can endure for many years.

Recognition of group members extends to the vocal domain. Analyses have shown that many primate vocalisations show consistent individual differences in their acoustic structure, and several studies have shown that the animals can use these differences to recognise one another at a distance. Rhesus monkeys, for example, produce a harmonic vocalisation that researchers term a "coo" (because of how it sounds). The monkeys produce coos at high rates when foraging in dispersed fashion in dense vegetation where visibility is limited. The calls function to maintain vocal contact when individuals are out of sight, which is important for highly nepotistic species such as rhesus monkeys who often need to recruit aid from kin in competition with other group members. Using field experimental techniques, researchers have played back to isolated females the pre-recorded coos of either a close female relative or a familiar but unrelated female in the group (Rendall, Rodman, & Emond, 1996). They find that females respond dramatically to the calls of their female kin but tend to ignore the calls of nonkin. Moreover, they appear to be able to distinguish between the calls of individual female relatives.

Not only can monkeys recognise one another, but they also appear to understand something about each others' social relationships. In a set of field experiments similar to those just described, Robert Seyfarth and Dorothy Cheney played back to female vervet monkeys the scream vocalisations of either their own infant or an unrelated infant in the group (Cheney & Seyfarth, 1990). Mothers responded preferentially to their own infant's calls. More interestingly, other females in the vicinity responded by looking toward the mother even before the mother herself had responded. These females were able thus to recognise the calls of various infants in the group and match them to their appropriate mothers. In similar fashion, Verena Dasser tested the ability of monkeys housed in a large captive colony to match slides of infants in the colony to their appropriate mothers (Dasser, 1988). They could, and could do so even when the slides of mothers and infants were several years old—i.e., the monkeys retained a memory for the appearance of others at younger ages.

This latter result implies that monkeys also have the requisite memory skills required for reciprocity. Recall that reciprocal altruism requires an ability to remember

past encounters, the identity of the participants, and the outcomes so as to adjust future behaviour accordingly. Additional evidence for these skills comes from natural patterns of reconciliation and redirected aggression observed in several species. In the course of daily activities, fights routinely break-out over access to limited resources of one sort or another (e.g., food, water, preferred resting sites, access to preferred social partners). Opponents in such fights often remain hostile to one another for several hours, such that if they encounter one another later they are apt to avoid each other or resume aggression. Many times after such a fight, the antagonists, or their kin, will subsequently redirect aggression at the relatives of their former opponents, even hours after the original fight. However, at times, rivals immediately reconcile by directing affiliative gestures toward one another, which function to restore baseline levels of tolerance such that when these individuals encounter one another again later they are likely to behave amicably.

In sum, primate social life appears to be a complex balance of competition and cooperation, some individuals competing (nonkin), others cooperating (kin), and still others competing to cooperate (e.g., low-ranking females competing to groom high-ranking females), or cooperating to compete (e.g., subordinate males cooperating to displace a dominant male for an opportunity to mate). Taken together, the various findings indicate clearly that primates possess the requisite memory and recognition abilities to support a system of reciprocal social exchange as well as many other cognitive abilities sustaining other complex patterns of social behaviour.

In fact, findings like these have led many to speculate that primate cognition is especially well developed in the social domain. Primate evolutionary history has been characterised by a disproportionate increase in brain size, meaning that, as an order, primates have larger brains for their body size than do most other animals. Researchers have hypothesised that the selective pressures that account for this increased brain size are social in nature—in essence, that primate brains have evolved in large part to solve problems in the social domain. Alternative hypotheses tend to stress environmental factors such as the ecological challenges posed by foraging for patchily distributed and ephemeral fruit resources, or the technological challenges involved in manufacturing and using tools, or constructing elaborate nests of vegetation in which to sleep.

Formal tests of these hypotheses are few, though evidence of the sort just reviewed certainly suggests that primates are highly gregarious animals with a well developed social acumen. Overall, primates seem no different than many other animals in their ecological and technological skills: many other animals are fruit specialists like primates and meet the challenges that fruit harvesting poses without dramatic increases in brain size; and with few exceptions, primates do not make extensive use of tools, or build nests of more complicated design than any other animals. However, primates do seem to distinguish themselves from other animals in their range and degree of socially complex behaviour. Furthermore, social problems also seem to be inherently more complicated than most ecological and technological problems because conspecifics (unlike features of the physical environment) are reactive to one's behaviour. Thus, social problem-solving involves predicting and interpreting the behaviour of others so as to adjust your own at the same time that your adjustments are eliciting behavioural changes in those whose behaviour you're trying to predict. Hence, the social world is constantly changing in an integrated but endless spiral of causality.

One recent effort to test the social hypothesis more formally was undertaken by Dorothy Cheney, Robert Seyfarth and Joan Silk (Cheney, Seyfarth, & Silk, 1995). They focused on one very important cognitive ability—the ability to understand the relation between cause and effect. This relation is fundamental to many aspects of life: it is a core property of the physical world (e.g., moving objects that contact others cause the latter to move; gravity causes suspended objects to fall; wind causes trees to move). It is also fundamental to social behaviour (e.g., socially dominant individuals cause subordinates to behave submissively).

Previous research had shown that monkeys do not perform well on tests of the cause-effect relation in the physical domain (reviewed in Tomasello & Call, 1997). One standard test involves presenting a monkey with a clear plastic tube with a food item in the middle, and providing several suitable poking sticks. The tube is too narrow to accommodate the monkey's hand, but large enough for a stick to be inserted. Subjects try in vain to jam their hand in the tube, to bite the tube open, or to break it by pounding or jumping on it. After much trial-and-error, subjects begin to explore the tube with one of the sticks (after first using it to try to break open the tube) and eventually learn to poke the food out of the tube. The tube is then modified to include a small "trap" well at some point along its length such that if the food is pushed in one direction it will fall into the well and be lost, but if pushed in the opposite direction it will emerge out the end of the tube as before. Subjects proficient on the original task, now exposed to this simple modification, have great difficulty, pushing the food into the trap as often as not (i.e., behaving at random). Gradually, they begin to pay very close attention to the direction the food is travelling down the tube as they push, and if it appears headed for the trap, they stop and push from the opposite direction. Next, the tube is modified one more time, this time simply rotating the tube so that the trap is on the top and thus at no risk of capturing the food no matter which direction it is pushed from. In this condition, subjects continue to proceed very cautiously as in the last condition, paying very close attention to the direction the food is travelling. If the food is approaching the trap—even though the trap is now on the *top* of the tube—subjects immediately stop and push from the opposite direction. Analogous sorts of tests of cause-effect reasoning in the physical domain have returned the same general result. The monkeys simply don't get it!

For comparison, Cheney, Seyfarth and Silk designed an analogous test in the social domain (Cheney et al., 1995). In baboons (as in many other species) dominant animals attacking subordinates produce harsh threat vocalisations that elicit shrill fear calls from the subordinate victims. The relation between these calls is obligatory: dominant threat calls elicit or cause fear calls in subordinates. The reverse never occurs. That is, subordinate threat calls never elicit fear calls from a dominant. The investigators capitalised on this relation to test the baboons' understanding of it. Using previously recorded vocalisations from animals in one group, they constructed sequences of causally consistent (a dominant threat call paired with a subordinate fear call) and causally inconsistent calls (a subordinate threat call paired with a dominant fear call). When played back to other individuals in the group, the causally consistent sequences elicited very little response. However, the causally inconsistent sequences elicited much stronger responses suggesting that the baboons were struck with the social anomaly that this situation implied, a contradiction of the normal causal relations

that underlie their social behaviour.

Clearly this is only one test that favours the social hypothesis. More research is needed. However, in combination with their many other complex patterns of social behaviour, it lends credence to the hypothesis that the trend toward increased brain size in primates might reflect cognitive specialisations for solving problems in the social domain.

24.5 Human Evolutionary Psychology

The primate trend toward increased brain size reaches an extreme in the human primate. Humans have extraordinarily large brains, about seven times what would be expected for a mammal of our body size. (Our brains are more than three times the size of that of a chimpanzee, our closest living relative.) What accounts for this dramatic increase in brain size? What are large human brains *for*? What problems did they evolve to solve? Questions like these are the focus of a relatively newly recognised subfield of psychology called Evolutionary Psychology. Actually, evolutionary psychology has deep roots dating to Darwin and before, but only fairly recently has it begun to crystallise into a distinct subfield that pursues issues of human cognition within an explicitly evolutionary framework.

The most fundamental tenet of evolutionary psychology is that the brain is an evolved organ (just like every other organ of the body), and that its design (like the design of every other organ of the body) is functionally specialised. Like the functional specialisation of other bodily organs (e.g., the heart for pumping blood, and the liver for filtering it; the lungs for breathing oxygen and the blood for delivering it to the body), the brain is functionally specialised to solve a variety of specific problems. The brain is not a single, general-purpose problem-solving device. Rather it is a set of multiple distinct but inter-related problem-solving devices each specialised to operate on informational content specific to particular domains of experience. Furthermore, only problems that represent recurrent evolutionary challenges (as opposed to problems of a transient or fleeting nature) are likely to be reflected in the brain's design. After all, evolutionary change of any kind occurs very gradually over long periods of time in response to enduring selective pressures. Circus bears can walk bipedally, like humans; however, such dancing bears show none of the functional modifications of the pelvis and lower-limb required for habitual bipedal locomotion that humans show, because bears have been dancing in circuses for only a short time, while humans have been walking upright for more than 4.5 million years.

Thus, evolutionary psychologists operationalise the human brain as a collection of more-or-less specialised cognitive modules dedicated to particular problems that have represented recurrent challenges over (at least) the last several million years of human evolutionary history. There is some evidence indicating that many of these recurrent challenges have, as in other primate species, been social in nature.

Consider the problems described in 24.1. They illustrate the Wason card selection task, a standard experimental technique for studying human reasoning, which has often been assumed to exemplify a classic general-purpose cognitive process. In problem 1, the task is to detect violations of an office filing rule that involves cross-coding client

files in two different ways. In Problem 2, the task is to detect violations of an alcohol consumption rule. In Problem 3, the task is to detect violations of a general cost-benefit rule.

In fact, all three problems test exactly the same reasoning skill - the ability to detect violations of a logical relation of the form: "If P, then Q". In other words, if some particular condition "P" holds, then some other condition "Q" must also hold. In all three problems, the correct answer involves turning over exactly two cards: the card that corresponds to "P" and the card that corresponds to "not Q". That is, in order to detect whether or not the obligatory relation linking "Q" to "P" has been broken, one must turn over the card corresponding to "P" to ensure that there is indeed the required "Q" on the reverse. The card corresponding to "not Q" must also be flipped to ensure that there is NOT a "P" on the other side. Because all three problems test the same logical relation, performance should be equally good on all three. . . but this proves not to be true!

In the first problem, the solution translates into turning over cards 2 (P) and 3 (not Q) to ensure that the file coded 'D' is, in fact, also coded with a '3' on the reverse and (evidently, counter-intuitively) to ensure also that the card coded '7' is NOT coded with a 'D' on its reverse. Turning over either of the other two cards is irrelevant with respect to the prescribed rule. This is probably obvious for the card coded 'F' because the rule said nothing about files coded with this letter. It may seem less obvious (in fact, surprising to many people) for the card coded '3'. However, turning over this card is not informative with respect to the rule either, for although the rule prescribed that all 'D' files had to bear a '3' code, it did not prescribe the converse, namely that all '3' files had to bear a 'D' code. So, it does not actually matter what is on the reverse of the card coded '3'.

In the second problem, cards 1 and 3 must be turned over (which probably seems patently obvious): it is necessary to check that the beer-drinker is of legal age and also that the 16-year old is not consuming alcohol. How old the coke-drinker is and what the 25-year old is drinking are irrelevant.

In the third problem, the solution requires flipping cards 3 and 4 to ensure that those who have not paid the cost have not accepted the benefit, and to ensure that those who have accepted the benefit have also paid the cost. Whether or not those who have paid the cost have also accepted the benefit is irrelevant to the rule. It is likewise irrelevant what those who forego the benefit ultimately do vis a vis paying the costs.

Tests like this (using many different forms of question) have been administered to scores of subjects and the results are always the same. People have great difficulty with problems like that exemplified by Problem 1 but no trouble with problems like those exemplified by Problems 2 and 3. As already noted, all three problems embody exactly the same relation—a rule prescribing an obligatory connection between two things or events. The only difference among them is that Problems 2 and 3 frame this relation in a social context. That is, in the form of a social 'contract' or exchange. That, and the fact that, by comparison to Problem 1, Problems 2 and 3 are embarrassingly simple! They seem completely intuitive. This result suggests that human reasoning skills are not, in fact, 'domain-general', but rather are 'domain-specific', meaning that they operate best on informational content in specific domains rather than being general-purpose processes for reasoning about problems across domains with different informational

Problem 1

Part of your new job at an accounting firm is to make sure that client documents are organised correctly. The firm has numerous clients that it organises according to different rating and coding rules. Your job is to make sure that the documents conform to the following rule:

"If a client has a 'D' rating, then its documents must be marked code '3'."

You suspect that the person you replaced did not categorise the documents correctly. The cards below have information about the documents of four clients. Each card represents one client. One side of a card tells a client's letter rating and the other side of the card tells that client's number code.

Indicate only those card(s) you definitely need to turn over to see if the documents have been categorised correctly.

F	D	7	3

Problem 2

In its crackdown against drunk drivers, provincial law enforcement officials are revoking liquor licences left and right. You are the owner of a restaurant, and you'll lose your operating license unless you enforce the following law:

"If a person in your restaurant is drinking alcohol, then she/he must be over 18 years old."

The cards below have information about four people seated in your restaurant. Each card represents one person. One side of a card tells what a person is drinking and the other side of the card tells that person's age.

Indicate only those card(s) you definitely need to turn over to see if any of these people are breaking the law.

Beer	Coke	16	25

Problem 3

It is your job to enforce the following law:

"If you take the benefit, then you must pay the cost."

The cards below have information about four people. Each card represents one person. One side of a card tells whether a person accepted a benefit and the other side tells whether that person paid the cost.

Indicate only those card(s) you definitely need to turn over to see if any of these people are breaking this law.

Paid Cost	Didn't Accept Benefit	Didn't Pay Cost	Accepted Benefit

Table 24.1: Reasoning problems adapted from the work of John Tooby and Leda Cosmides.

contents.

It is important to note that the ability to detect violations in these tests does not vary with the detail provided, nor with content familiarity. Problem 1, for example, arguably provides the most detail but yields poor performance, while Problem 3 is worded in the most abstract fashion (i.e., devoid of details), yet performance is good. Furthermore, experimenters who conduct these tests are careful to draft many different forms of the questions to control for differential familiarity with their content. Nevertheless, the results are the same. In fact, even very unfamiliar social problems, describing social contracts that are completely foreign (even exotic!) to subjects (e.g., "If a man eats cassava root, then he must have a tattoo on his face.") are solved easily, whereas highly familiar but non-social problems (e.g., "If one goes to Boston, then one takes the train.") prove difficult. Thus, familiarity with social problems cannot by itself account for our skill at detecting violations of social contracts, a fact further evidenced by the behaviour of small children. Even very young children with comparatively little social experience are expert at detecting social violations, as anyone who has tried to "short-change" a small child on the last slice of pizza, or the larger of two ice-cream cones, can attest! Instead, we just seem to be inherently good at reasoning about social contracts—we are inherently good at detecting "cheaters".

Formal tests like these using the Wason card selection task are not conducted in isolation. Rather they have been inspired and buttressed by many other sorts of results that together support the view that human psychology has been shaped in large part by evolutionary challenges faced in the social domain. Humans, like other primates, are intensely social and engage in many complex social practises; hence, the social problems so generated have been a recurrent feature of our evolutionary history. Other aspects of our environment—such as its ecological and technological bases—have changed markedly over the last several million years and continue to change rapidly today! However, our capacity for group living and intense sociality have not changed. The functional demands of social problem solving, therefore, have been a continuing selective pressure that has favoured cognitive adaptations for negotiating social exchange. Problems 2 and 3 above are comparatively easy, then, because they tap special cognitive mechanisms for reasoning about social contracts and ferreting out cheaters to them.

Interestingly, although we are adept at detecting cheaters, we are far less adept at detecting altruists. In identical versions of the Wason card selection task, subjects do poorly if asked to detect altruists rather than cheaters—that is, to identify individuals who pay the cost but do not take the benefit to which the cost entitles them. At first, this result may seem surprising given the fact that we have specialised social reasoning skills. However, when you think about, detecting altruists does not represent a serious selective pressure. Failing to detect an altruist is not costly to you for it is the altruist who pays the cost (not you) without taking the benefit. Failing to detect a cheater, on the other hand, is costly to you because the cheater takes the benefit (from you) without paying the cost. This imbalance in the ability to detect cheaters versus altruists suggests that the cognitive mechanisms underlying social exchange are highly specific. Not only is the cognitive architecture of our minds not designed as a general-purpose device for solving problems across different informational domains, but even within more specific domains, such as that of social exchange, evolved cognitive mechanisms

are quite finely tuned to specific dimensions of those domains, in this case, to detecting cheaters. Cheating is the aspect of social exchange that represents the evolutionary challenge, and it is to this aspect of the problem that we have specific cognitive mechanisms.

In fact, cheating is not the only part of the social exchange problem that is critical and a specialised cheater detection module is but one possible evolved social cognitive module. Examples of other hypothesised social cognitive modules include: a module for recognising faces, important for discriminating among individuals for various social purposes but especially social exchange and reciprocity (as outlined earlier); a module for language, critical for mediating social interaction including social contracts; and a module for mate selection, decisions about mating obviously being a recurrent part of our evolutionary history. There is a variety of evidence in support of these (and other) social cognitive modules. Faces, for example, are highly salient stimuli. Faces are obviously the focus of our social attention, but this facile observation does not simply reflect social experience. Faces are extremely salient even to new-born infants. Small babies—only hours old—preferentially orient to schematics that display landmark facial features (i.e., eyes, a nose and a mouth) over schematics displaying other things. They are also sensitive to the proper orientation of these landmark facial features, such that disrupting the normal configuration of eyes, nose and mouth eliminates the preference for looking at them, which suggests that human infants are innately equipped with cognitive mechanisms for recognising faces. (Perhaps a cognitive specialisation for recognising faces accounts also for the peculiar human tendency to "see" faces in very obscure places, such as clouds, impressionistic art, and even mountain-scapes!)

Language too is central to human social behaviour (see Chapter on Language, Evolution and Human Uniqueness). Language is present in every population around the world, and despite their superficial diversity all languages embody the same core elements (e.g., symbolism, grammar, and syntax) and share a common underlying organisational structure. Furthermore, although children are not born talking, every child the world over acquires language at about the same age, following a similar developmental trajectory, and all without formal instruction, even if their abilities are compromised in other domains. Thus, mentally challenged children acquire language just as surely as other children, suggesting that language has important biological underpinnings and recruits specialised cognitive processes.

Finally, large-scale studies of mate selection in many populations around the world have revealed remarkably consistent sex differences in the cues that males and females find attractive in potential mates, the differences correlating with the different resources that limit reproductive success in each sex. Thus, across cultures, females prefer males of high status with demonstrated access to material resources (which often translates into a preference for older, established males), while males prefer females with high fertility (which almost invariably translates into a preference for younger females irrespective of socio-economic position). Despite marked differences in the details of cultural and mating practises around the world, then, males and females consistently show the same psychological mechanisms governing mate attraction.[3]

[3]This does not mean that there are not a host of other variables involved ultimately in decisions about mating (or marriage). Obviously, these issues are incredibly complicated and are influenced by numerous

Although the emphasis here has been on social cognitive modules, potential cognitive specialisations need not be limited to the social domain. Indeed, modules specialised for other problem domains have been proposed (e.g., food choice, and habitat selection). Evolutionary psychology is a young subfield and its full depth and breadth have yet to be realised. However, it is already beginning to yield important insights into the evolution of human psychology—into why we think and do what we think and do. Like any new field of enquiry, one test of its utility is its compatibility with accumulated knowledge in related fields. In this respect also, evolutionary psychology looks promising, as many of its findings are compatible not only with basic evolutionary theory that guides research in most of the behavioural sciences but also with empirical findings in closely allied subfields of psychology, namely neuroscience.

Increasingly, research in neuroscience is revealing functionally specialised areas of the brain, some of which correspond to hypothesised social cognitive modules. Recent work has revealed individuals with very specific cognitive deficits resulting from highly localised brain damage. For example, some individuals suffer from a deficit that compromises their ability to recognise faces, a deficit termed prosopagnosia. Although these individuals can recognise other objects without difficulty, they cannot recognise the faces of very familiar people such as their own husbands. (An analogous deficit in recognising voices has also been hypothesised.) Research has also identified numerous brain areas involved in language processing, damage to any one of which can cause specific language deficits (generally termed 'aphasias'). For example, diffuse damage in the left frontal lobe often impairs speech production (traditionally called Broca's aphasia), while damage in the left temporal or parietal lobe can impair speech comprehension (Wernicke's aphasia). More localised damage within these areas can cause even more specific language impairments such as the inability to name familiar objects (anomia), or the inability to comprehend the grammatical structure of speech (agrammatica).

Research like this supports the tenets of evolutionary psychology that the brain is structured as a collection of distinct cognitive modules specialised for processing certain kinds of information. In many ways, evolutionary psychology and neuroscience are complementary, evolutionary psychology suggesting specific problem domains that represent plausible evolved cognitive modules and neuroscience evaluating the structural realisation of such modules. As yet, localised brain areas for detecting cheaters have not been identified, and it is possible that they never will be. Some cognitive processes may be highly distributed in such diffuse neural networks that they are robust to damage to localised circuits or simply resist efforts to localise them. It is interesting to note, however, that schizophrenics who typically manifest compromised reasoning skills have no trouble with social contract problems like those illustrated in Box 1, confirming that cognitive mechanisms for the latter (however diffusely distributed they may prove to be) are nevertheless insulated from other reasoning domains. Thus, in combination, evolutionary psychology and neuroscience are mutually informative and should ultimately contribute to an ever-clearer picture of the functional organisation of the human brain and the problems it has evolved to solve.

---

cultural practises in addition to basic aspects of biology. Nevertheless, it is illuminating to find such consistent differences despite tremendous cultural diversity, differences that align well with sex differences in evolutionarily recurrent selective pressures affecting reproductive success.

24.6 Conclusion

Competition and cooperation are ubiquitous features of primate social life. The complex interplay between these conflicting influences on behaviour has, over the course of primate evolutionary history, shaped the structure and function of human psychology. The challenges of negotiating social exchange have been a recurrent selective pressure favouring cognitive adaptations for reasoning about social problems. These social cognitive adaptations represent a collection of distinct but integrated processing modules, each dedicated to specific informational content within the social domain. Much of the human psyche, then, constitutes an evolved cognitive architecture designed to solve various problems associated with being a highly social organism whose survival and reproduction hinges on negotiating encounters with social companions.

Chapter 25

Memory Systems in the Brain
Rob Sutherland

One of the more important insights gained by modern psychology is that our ability to remember reflects the functioning of several, separate memory processes and systems in the brain. Our ability to remember what we ate during our last meal, the face of our paternal grandmother, how to maintain our balance during turns while ice-skating, how we felt when we scored our first "A" on a university test, or where we have just placed our pen are all instances of memory, unified by the fact that it is the same mind who "has" each of these memories. That they are unified in any important way may be illusory. Each represents a complex set of processes occurring in connections between neurons contained within anatomically distinct, functionally separate brain regions.

Consider the case of patient HM who contributed more to the modern understanding of human memory systems than any other person. For years HM suffered epileptic seizures many times each day. Available treatments were ineffective. While still a young man, HM was referred to a Boston neurosurgeon, W. Scoville, who was pioneering a surgical procedure in which brain tissue was removed to reduce seizures. Scoville performed a radical version of the procedure, termed "medial temporal lobectomy", during which nearly all of HM's medial temporal lobe was removed, importantly, in both left and right hemispheres of his brain. In part prompted by complaints from HM and his family about memory problems after the surgery, Scoville carried out many standard tests of psychological functions, including general intellectual competencies, attention, digit span, and the like. HM passed all of these with flying colours, yet complaints of memory problems persisted. It was not until a young neuropsychologist from McGill University, Brenda Milner, conducted a more thorough and appropriate battery of psychological tests that the nature of HM's plight was discovered.

25.1 HM's Memory Deficits

Milner showed that HM was able to remember normally any new information only for as long as he continued to think about it—very soon after he turned his attention to something else, he was unable to recall the new information. For example, if Milner introduced herself and carried on a normal conversation about any topic HM would have no difficulty in holding up his end. However, if, in mid-conversation she drew his attention to something occurring out of HM's window, when she would subsequently ask if he remembered her name or what they had been talking about he would have completely forgotten, often denying that they had ever spoken before. Thus, when HM engaged in effortful attention to new information his recall was normal. This unaffected memory process is termed *working memory*. HM appeared to have a specific impairment in forming new long-term memories. This deficit is termed *anterograde amnesia*. Milner further demonstrated that HM could recall accurately events from most of his life before his surgery and he remembered his own telephone number and home address and the like. However, memories from the 5 year period before surgery were almost entirely gone. This loss of memories of events from before brain injury is termed *retrograde amnesia*. Milner concluded that Scoville had removed a brain region essential for acquiring and—at least initially—storing long-term memories. For several reasons Milner believed that a structure called the hippocampus was the critical area within the medial temporal lobe. Subsequent work has revealed that Milner was essentially correct that the hippocampus is critical for certain forms of memories, but so are other medial temporal lobe structures, albeit in somewhat different ways. Thus, HM may be viewed now as a patient who has lost functioning of several important memory systems (described below).

25.2 Certain Forms of Memory are Preserved

A student of Milner's, Suzanne Corkin, made another important discovery about HM's memory. She showed that there were two kinds of new memories that HM had no trouble forming and retaining. The first involves learning new skills. The classic example is learning to trace a complex figure when only a mirror-image of the subject's hand movements are available—the subject cannot see his or her hand directly, but must learn to make movements based upon the inverted visual information from the mirror. With extended practice people learn this skill and retain it for a long time—so does HM. In fact, HM and other patients with his form of amnesia have learned and remembered a surprisingly wide range of new skills—without being able to report a single memory from any of the previous practice episodes! When questioned about how they are able to perform these skills almost flawlessly they answer something along the lines of, "I guess I am naturally good at this kind of thing."

25.3 Priming

The second kind of preserved learning is termed *priming*. When subjects are shown specific pictures, words, or other items, very quickly or within a long list of similar material they are often unable to explicitly recall whether they had seen a specific item on the list, say a picture of an aeroplane or the word "clarinet". Despite a complete absence of explicit memory for these items, it is possible to show that people retain long-term information about the items. This retention can be tested in a few indirect ways. One way is to show a very degraded, fragmentary version of the original item, say a fragmented picture of the identical aeroplane. It has been shown that the prior, "forgotten", exposure significantly enhances the ability to identify the degraded picture (this enhancement is very specific to the original picture, it does not generalize to all pictures or even pictures of other planes or the same plane in a different orientation). A second indirect way to test for retained information can be illustrated with the example of the word item "clarinet". This test is a simple spelling test that takes advantage of the fact that there are many homophones in the English language. A person is asked to spell a word like "reed". Because the words "read" and "reed" are spoken identically, the request is ambiguous and one might expect the person to respond with the higher frequency spelling, "read". However, even if a person does not remember that the word "clarinet" had previously been presented, the person's spelling is often biased toward "reed", the lower frequency spelling that was *primed* by the earlier presentation of a reed instrument.

Indirect tests of memory are often thought to tap long-term memory that is unconscious or not reportable. The contemporary term for this form of memory is *implicit memory* and it is largely, if not completely, preserved in amnesia. In contrast, direct tests of memory, involving items such as, "tell me what you ate for breakfast", or, "what was the topic of our most recent conversation?", or, "where do you live?", measure *explicit memory*, information that can be consciously recalled and reported to others. The preservation of memory for skills and of priming in the face of devastated explicit memory after medial temporal lobe damage represents a good illustration of two facts. Our brains contain multiple, independent memory systems. And, much of what we do as a consequence of prior experience occurs without conscious awareness.

25.4 Functional Brain Imaging

The branch of science that deals with complex psychological processes, such as memory, attention, perception, and so on, with a focus on brain mechanisms is called *cognitive neuroscience*. Most of our information in cognitive neuroscience has depended upon careful behavioural studies of patients and nonhuman animals with brain injury—accidental or experimental—going back to clinical studies in the 1800s by Alzheimer, Korsakoff, and others. Contemporary work on memory draws on a very wide array of methods, including those from cognitive psychology, neurocomputational modelling, neurophysiology, neuropharmacology, and physics. Functional brain imaging provides cognitive neuroscientists with a set of new, rapidly evolving, extraordinarily powerful methods. Given that the mind is the brain in action, functional brain imaging allows

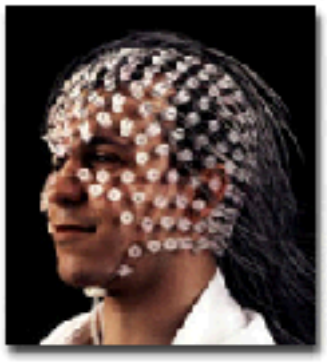


Figure 25.1: Dense-array ERP sensors.

us to actually observe the mind/brain as it is thinking, feeling, planning, fantasizing, and remembering. We will discuss two of these methods in some detail: event-related potentials (ERPs) and functional magnetic resonance imaging (fMRI).

25.5 ERP

Neurons are specialized for processing information, in large part because of their ability to undergo carefully regulated, rapid shifts in the electrical properties of their membranes. These changes (the action potentials and synaptic potentials described in other chapters) generate electrical currents that produce measurable changes in the electrical field on the surface of the head. The size of these changes depends upon many factors, including how many neurons are generating a similar signal, how far these neurons are from the recording contacts on the head, how synchronized are the signals from different neurons, what other signals are being generated in other neurons, and the amount and type of tissue through which the electric field changes are distorted between the signal and the recording contacts. Making accurate inferences about the location of a neuronal signal at any point in time is a dauntingly complex task and there is much current work directed toward improving this accuracy. Some of the improvements are based upon better mathematical approaches to "source localization" and some, such as the *dense-array system* pictured in Figure 25.1, are based upon improvements in contacts/detectors (such as dramatically increasing the number of surface contacts). A major advance was the development of signal averaging techniques.

25.6 ERP: Signal Averaging

Consider a memory task in which familiar and novel pictures of faces are shown to a person and the person's task is to push one button if the face is familiar and a different

button if the face is novel. Rather than measuring brain activity during presentation of each face, ERP workers average the responses at each contact over all of the face presentations of a certain type. The assumption is made that there is lots and lots of processing going on during the face task, most of which is irrelevant to recognizing faces. Subject may be thinking about what they did last night, what they are going to eat for lunch, why this experiment is taking so long, and so on. A further assumption is made that this irrelevant processing will not have a consistent temporal relationship to the presentation of each face. In contrast, relevant activity in visual areas and mnemonic systems involved in face recognition is quite likely to have a relatively constant temporal relationship to when each face is presented or perhaps to when a response is made by the subject indicating that this is a familiar or novel face. What one is able to see after such averaging at each scalp contact is a flat line before face presentation followed by reliable fluctuations in the local electric field beginning soon after face presentation that continue to change for several hundred milliseconds until just before the response is made. The signals can be averaged by time-locking to any event, be it stimulation (face presentation), movement (button push), or meaningful action (correct familiarity response).

Because the scalp contacts are in slightly different locations relative to the neuronal signals within the brain, it is often possible to accurately determine the location of the individual waves or potentials. This is especially true for signals that arise from neurons of the cerebral cortex that are close to the skull. Using this strategy and superimposing inferred locations onto a subject's 3-D MRI scan we can see a variety of processing stages for faces. First, we can identify regions of posterior visual cortex that respond to faces very early on, but these regions show the same response to non-face stimuli that have very similar visual features. Next, areas of temporal cortex (fusiform and lingual gyri to be exact) show the first face-specific signals, but these areas respond equally to all faces, be they familiar or novel. Finally, beginning around 400 ms after the face is presented there are medial temporal regions that respond differently to familiar versus novel faces—these would correspond to the circuits where memories first interact with the specifics of face perception.

The ERP technique has three important advantages compared with most other methods. One is temporal resolution. Neuronal activity occurs very quickly and ERP techniques allow for accuracy below a millisecond. A second is validity. ERPs measure the electrical activity from neural information processing directly. A third is safety and ease of use. There is no need for potentially harmful and expensive radioactive isotopes, injections, or non-physiological applied electromagnetic fields. An important short-coming of the technique is the blurring of electrical field flux by various tissues (fluids, bone, skin) between the generator of the signal and the scalp contacts and mathematical uncertainties associated with source localization.

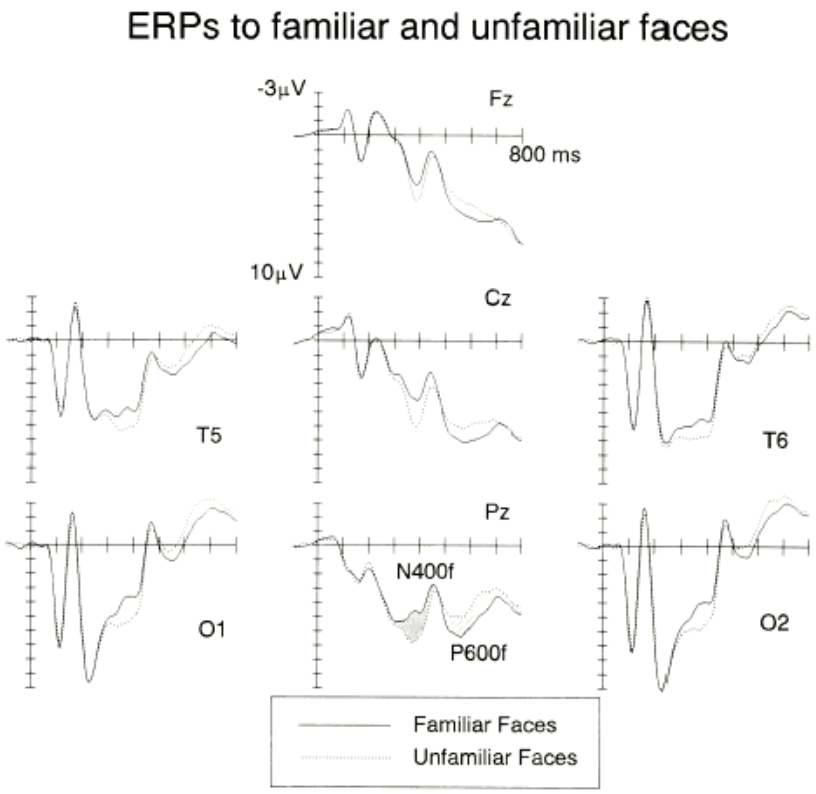


Figure 25.2: Average signals shown at 7 of the ERP scalp contacts for familiar and unfamiliar faces (from Eimer, 2000).

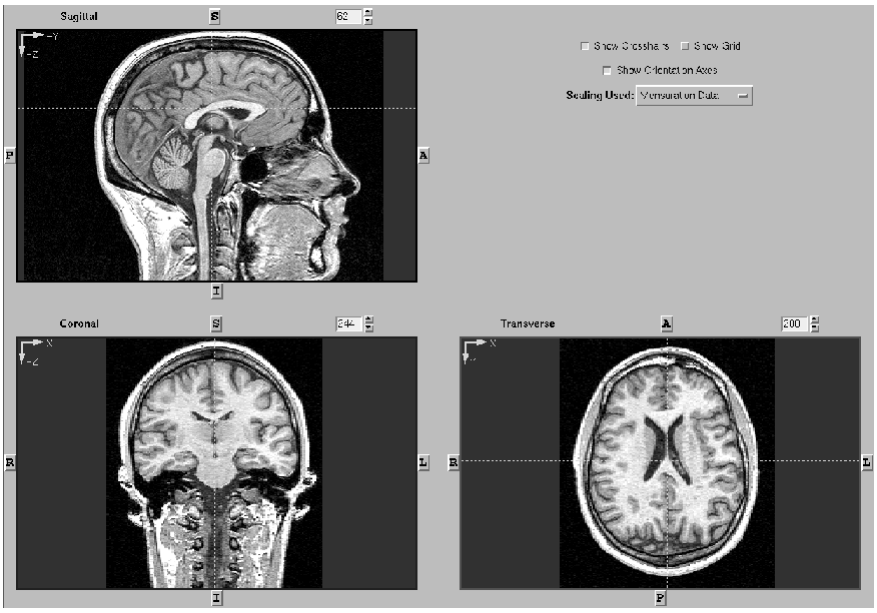


Figure 25.3: A MRI of the brain viewed from the three standard orientations

25.7 fMRI (Functional Magnetic Resonance Imaging)

fMRI is a newer technique for measuring brain activity [1]. An example image is shown in Figure 25.3 and the machine that generates the images is shown in Figure 25.4. fMRI is not beset by the source localization problems of ERP, but it has inferior temporal resolution (at best a few seconds is its quickest response). fMRI takes advantage of the fact that the brain has a regulatory system that delivers extra oxygenated blood to areas where neurons have been especially active (because neural activity places demands on pathways related to oxidative metabolism). Most of the oxygen in blood is carried by hæmoglobin molecules and if we were to look carefully at a capillary that irrigated a region of fusiform gyrus just as we looked up from a book to stare at our grandmother's face we would notice a decrease for a few seconds in the number of oxygenated hæmoglobin molecules followed by a longer lasting increase in local oxygenated hæmoglobin.

It is this later upswing in oxyhemoglobin that creates the BOLD response (Blood Oxygen Level Dependent effect) that is the measured signal in most fMRI. Essentially, MRI depends upon recording radio-frequency signals generated by aligned hydrogen atoms "spinning" coherently in a strong magnetic field. Oxy- and deoxyhæmoglobin get into the picture because these two molecules differ in how much they disturb the local magnetic field, which changes the coherence of the signalling hydrogen atoms.

[1] see http://www.fmrib.ox.ac.uk/fmr_intro/ for a description of the basics

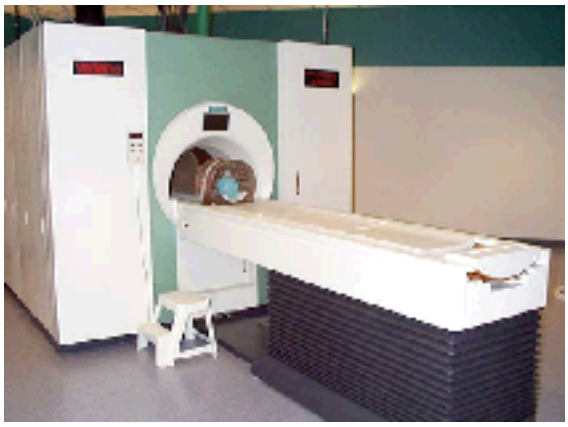


Figure 25.4: MRI scanner

Deoxyhæmoglobin perturbs the local field more than oxyhemoglobin, thus there is an increase in the MRI signal as the relative amount of oxyhemoglobin goes up. More neuronal activity = more oxyhemoglobin = greater MRI signal intensity. We can be much more sure of where neuronal activation is occurring with fMRI (to within about 0.5 mm) than we are about when the activation occurred, since the BOLD response unfolds over many seconds. Just like with ERPs, we can locate the sources of inferred activity within a 3-D MRI scan as shown in Figure 25.5.

25.8 What About Memory Systems?

Functional brain imaging studies, especially fMRI, have confirmed many of Milner's conclusions about memory. In particular, neuroimaging studies have shown a sharp dissociation between those areas that are involved in working memory and those that subserve long-term memory processes. Milner would predict that medial temporal lobe regions, especially hippocampus, would be strongly activated in the latter but not the former tasks. Based upon a great deal of other work with frontal lobe lesion patients it would be predicted that a specific region of frontal cortex would be activated by working memory tasks and not by long-term memory tasks. What do we see with fMRI studies? When we show short lists of words or other items and soon after ask subjects to identify whether or not particular items were on the list, there is a prominent activation of the dorsolateral prefrontal cortex (and little activation of hippocampus). In contrast, when people are engaged in committing new information to memory, especially new spatial information (see Figure 25.6 for an example), or when they are retrieving this information from long-term memory there is a prominent activation of hippocampus and a less robust activation of the dorsolateral prefrontal cortex.

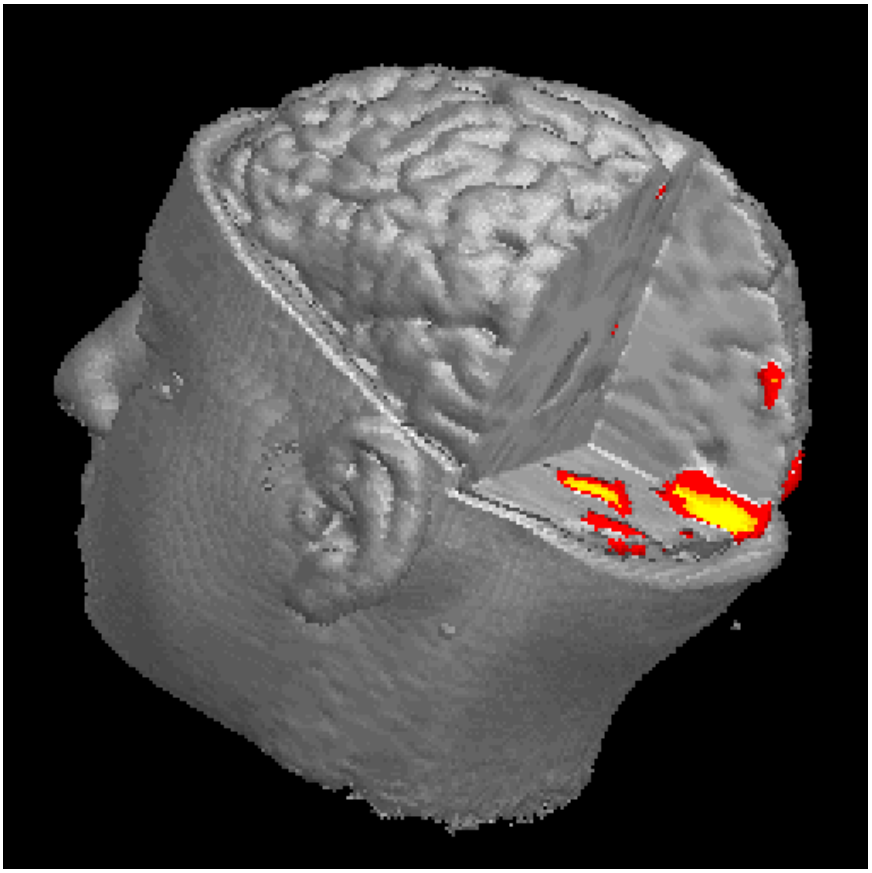


Figure 25.5: BOLD signal generated by presentation of a complex visual cue superimposed on a 3-D structural MRI.

Figure 25.7 provides one comprehensive overview of the different memory systems we have mentioned. This taxonomy of Squire and Knowlton (1994) refers to explicit and implicit memory as declarative and non-declarative respectively. Within the medial temporal lobe structures under the Declarative label you will see the amygdala. This structure, located just in front of the hippocampus, is especially activated by emotion-laden events. We know from experimental lesion work with nonhumans that damage to the amygdala impairs learning various associations between cues and emotions, like fear or pleasure, without affecting other forms of associative learning. Consistent data have been collected in fMRI studies with humans, in which various kinds of learned emotional responses activate amygdala neurons (and not hippocampus or dorsolateral prefrontal cortex). If we turn our attention to the other side of the memory taxonomy figure, under skills and habits the cerebellum is listed. Experimental or accidental damage to the cerebellum has little or no effect on emotional learning, or working memory, or formation of new explicit memories. Instead there is a very selective deficit in classical conditioning. Conditioning involving discrete conditioned responses like eye blinks or leg flexion is abolished by cerebellar damage. Likewise, classical eye-blink conditioning has been shown in human fMRI studies to selectively activate portions of the cerebellum.

25.9 Conclusion

Not too long ago, a significant number of well-educated, intelligent scholars wrote definitive treatises concluding that we would never, using the tools of science, be able to learn the composition of stars. Their argument was based upon the impossible distances from earth to stars and upon the scarcity of useful information that could be gathered. Despite their arguments, we now have amazingly accurate information about stellar composition, not to mention stellar life-cycles and origins. In a similar vein, we often hear that the human mind and self are impossibly complex and that there are insurmountable difficulties in obtaining good useful information. The burgeoning field of cognitive neuroscience, with its newly acquired and maturing methods of lesion analysis, functional brain imaging, and neurocomputational modelling, are leading us to solve problems in psychology where the imaginations of a prior generation saw only impenetrable dark.

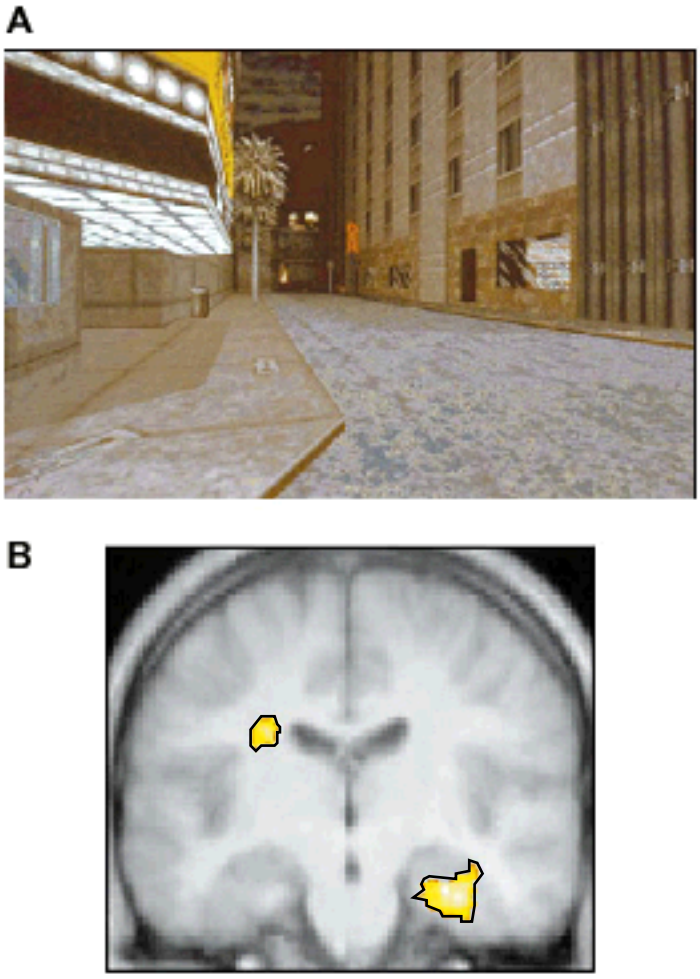


Figure 25.6: Panel A shows a computerized street scene that must be committed to memory during a spatial mapping task. Panel B shows activity in right hippocampus during performance.

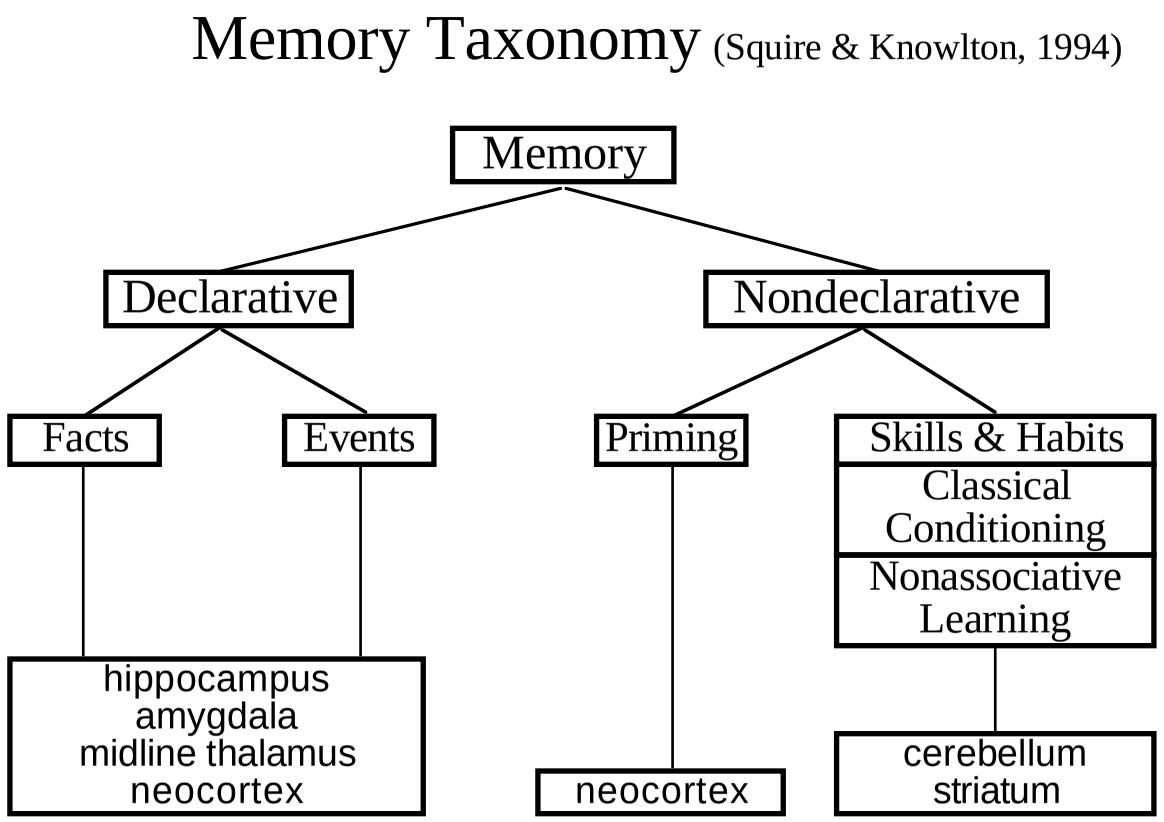


Figure 25.7: A current taxonomy of memory systems.

Chapter 26

Ethical Considerations in Psychology and Neuroscience Research
Rob Sutherland

The most important event in how we attempt to answer psychological questions was the change from a reliance upon the methods of insight, introspection, and learning the doctrines of long-dead authorities to the application of direct observation and analysis of the phenomena of behaviour and brain. An important implication of this aspect of the scientific method is that we use humans and other animals as the subjects for our observations and experiments. This use has brought psychologists and neuroscientists into regular contact with the realm of societal norms about what is ethically justified and what is lawful.

Cultures in different places and our own cultures in different times have evolved numerous rules about how people and other animals should be treated. In some it is unethical for strangers to look upon a woman's face; in others, it is ethically justified to abandon a baby before the age of one, or to own another person as property; in still others, it was a moral imperative to torture and even kill people who did not submit to a particular version of Christianity.

Over the second half of the 20th century scientists witnessed major changes in the understanding of the ethical status of human and nonhuman animal participants in research. This chapter tries to identify some of the major events and ideas in this recent evolution that set the ethical context for our research.

In 1984 I visited Dachau, the first of the German concentration camps active during World War II in a town very near Munich. The camp is maintained as a musem exhibiting artifacts and photographs of the horrifying treatment of prisoners and as a memorial to the unknown thousands and thousands of victims, nearly all of them German left-

wing labour activists, communists, gypsies, homosexuals, or Jews, whose cremated remains are buried in mass graves around the camp. I was not prepared for the graphic and extensive documentation of the scientific research conducted on the prisoners. One series of experiments I vividly remember. The participants were placed into air-tight chambers and exposed to suddenly lowered air pressure as a laboratory simulation of very high altitude. The experimenter wanted to learn at what altitude people would die and how the low pressure affected various organs including the brain. At specific pressures the tops of the participants' skulls were quickly cut off to immediately examine the condition of the brain, regardless of whether the person was already dead. I knew that experiments like this one had been conducted, but seeing the artifacts and photographs generated a strong sense of what these experiments must have been like to witness.

The victors at the end of WWII conducted show trials in Nuremberg, Germany involving various members of the Nazi government, military officers, and officers who participated in concentration camp activities. The testimony concerning mass executions, genocide, treatment of concentration camp prisoners, and medical experiments provoked a sense of outrage and urgency to create an international legislative solution to problems of this sort. One outcome was a codification in 1944 of ethical principles that has become the foundational ethical guide to researchers. This set of principles came to be called the Nuremberg Code.

The principles are:

1. from the participant. Informed consent includes freedom from coercion, fraud, deceit, and enough understanding of the nature, methods, purpose, and duration of the research to make an enlightened decision. It includes information about inconveniences, duress, and health effects from the research. It is the responsibility of the experimenter to determine that the conditions of informed consent are satisfied.

2. The study should yield results *for the good of society*, not obtainable by other means.

3. Based upon design and prior knowledge of the disease and animal experimentation, the *anticipated results should justify* the experiment.

4. Avoid *unnecessary* physical and mental suffering.

5. No experiment should be conducted if there is a *good chance of death or disabling injury*.

6. The *degree of risk* should not outweigh the humanitarian benefit.

7. Preparations should be taken to prevent even remote possibilities of subject injury or death.

8. Only *qualified scientific personnel* should conduct the experiment.

9. Subjects have the *right to quit* the experiment if they perceive physical or mental effects making it impossible to continue.

10. If at any stage in the course of an experiment the scientist believes that injury, disability, or death may result, he or she must *terminate the study*.

The World Medical Association at its General Assembly in Helsinki, Finland in 1964 expanded upon and added to the 10 principles of the Nuremberg Code. The Helsinki Declaration was subsequently amended several times, most recently in 2000. The most important additions include:

1. The design and performance of each experimental procedure involving human subjects should be clearly formulated in an experimental protocol. This protocol should be submitted for consideration, comment, guidance, and where appropriate, approval to a *specially appointed ethical review committee*, which must be independent of the investigator, the sponsor, or any other kind of undue influence. This independent committee should be in conformity with the laws and regulations of the country in which the research experiment is performed. The committee has the right to monitor ongoing trials. The researcher has the obligation to provide monitoring information to the committee, especially any serious adverse events. The researcher should also submit to the committee, for review, information regarding funding, sponsors, institutional affiliations, other potential conflicts of interest, and incentives for subjects.

2. Medical research is only justified if there is a reasonable likelihood that the populations in which the research is carried out stand to benefit from the results of the research.

3. The right of research subjects to safeguard their integrity must always be respected. Every precaution should be taken to respect the *privacy* of the subject, the *confidentiality* of the patient's information and to minimize the impact of the study on the subject's physical and mental integrity and on the personality of the subject.

4. For a research subject who is legally incompetent, physically or mentally incapable of giving consent, or is a legally incompetent minor, the investigator must obtain informed consent from the *legally authorized representative* in accordance with applicable law. These groups should not be included in research unless the research is necessary to promote the health of the population represented and this research cannot instead be performed on legally competent persons.

5. Research on individuals from whom it is not possible to obtain consent, including proxy or advance consent, should be done only if the physical/mental condition that prevents obtaining informed consent is a *necessary characteristic of the research population*. The specific reasons for involving research subjects with a condition that renders them unable to give informed consent should be stated in the experimental protocol for consideration and approval of the review committee. The protocol should state that consent to remain in the research should be obtained as soon as possible from the individual or a legally authorized surrogate.

The major changes introduced by the Helsinki Declaration include a much more complete elaboration of the meaning of informed consent and consent by legal guardians,

establishing the principle of the right to privacy or confidentiality, the establishment of a central role of the ethical review committee in overseeing research on humans, and clarifying the idea that the research should have promise in benefiting specifically the population studied. Further evolution of the guidelines awaited a few landmark human studies.

From 1932 to 1972, 399 poor black sharecroppers in Macon County, Alabama were denied treatment for syphilis and deceived by physicians of the United States Public Health Service. As part of the Tuskegee Syphilis Study, designed to document the natural history of the disease, these men were told that they were being treated for "bad blood." In fact, government officials went to extreme lengths to insure that they received no therapy from any source. Needless to say these men suffered horribly and unnecessarily in order to provide information to researchers funded by the National Institutes of Health in the U.S. This study and several others led to a major commission to re-establish basic ethical principles. In 1978, the U.S. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research submitted its report entitled *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research.* The Report, named after the Belmont Conference Center at the Smithsonian Institution where the discussions which resulted in its formulation were begun, sets forth the basic ethical principles underlying the acceptable conduct of research involving human subjects. Those principles, respect for persons, beneficence, and justice, are now accepted as the three quintessential requirements for the ethical conduct of research involving human subjects. The principles can be thought of as underlying the specific guidelines developed in the Nuremberg Code and the Helsinki Declaration and are in no way in conflict.

Respect for persons involves a recognition of the personal dignity and autonomy of individuals, and special protection of those persons with diminished autonomy.

Beneficence entails an obligation to protect persons from harm by maximizing anticipated benefits and minimizing possible risks of harm.

Justice requires that the benefits and burdens of research be distributed fairly.

The Report also describes how these principles apply to the conduct of research. Specifically, the principle of *respect for persons* underlies the need to obtain informed consent and privacy; the principle of *beneficence* underlies the need to engage in a risk/benefit analysis and to minimize risks; and the principle of *justice* requires that subjects be fairly selected.

The principles first enunciated in the Nuremberg code and elaborated upon in the Helsinki Declaration and the Belmont Report continue to provide the benchmarks against which contemporary experiments' ethical considerations are measured. However, it is important to note that unethical researchers will continue to find ways to perform unacceptable experiments off the ethical radar screen. For example, the U.S. government in the 1950s and 60s conducted a series of experiments to study the brainwashing effects of high doses of LSD on psychiatric hospital inmates and ordinary citizens in the U.S. and Canada. Furthermore during the same era, small quantities of dangerous radioactive isotopes were inserted chronically into unsuspecting children

who were undergoing routine surgical procedures such as tonsillectomy. The purpose was to study effects of long-term radiation exposure. Both of these kinds of experiments were funded by U.S. granting agencies, but were performed under the protection of government secrecy laws. It was not until decades later that the information finally became public.

26.1 Research with Non-human Animals

The ethical principles guiding research with nonhuman animals seem to be less clear. For example, there is a sizeable contemporary contingent who see no harm in and great benefit to banning all research with nonhuman animals. At the same time, some believe there should be no constraint at all on nonhuman animal research. In the context of little consensus it is worthwhile to clarify the historical and contemporary arguments for and against such research in a point/counter-point fashion.

26.1.1 Arguments for nonhuman animal research

Naturalism

This position holds that animals kill one another, eat one another, and inflict pain and suffering in their natural life histories. Humans, just like all other animals have always exploited other organisms in similar ways. Experiments with nonhumans represent extensions of the natural use of animals to promote human lives. [Even if this position were correct, one would need at least one additional valid argument to have a good ethical justification, namely, establishing that something that is natural is automatically good. No one has come up with a satisfactory argument equating the two and there are at least several examples of things that are natural that we do not believe are good. For example, rape and murdering children are considered by many psychologists to be natural, but they are certainly not ethically good.]

Souls

A position most closely associated with the French luminary René Descartes is that only humans are given souls by God. Nonhuman animals have no souls. The immaterial human soul enables us to experience modalities such as suffering, love, moral certitude, etc; in the absence of a soul an animal has no such inner experience. Animals exhibiting pain responses are showing us only “sham” pain, not “felt” pain. Nonhuman animals are like zombies (i.e., without the inner experiences—only the overt responses). Thus, in the absence of the ability to really suffer, any kind of experimentation is justified. Contemporaries of Descartes described his ability to inflict extreme pain upon a nonhuman animal while laughing at its tormented responses. [Today the concept of an immaterial soul has no scientific currency at all. As the knowledge of the relationship between neural circuitry and pain experience has expanded, we can see that at least all vertebrates have the kind of neural machinery that is, in humans, the basis for pain and suffering.]

Utilitarianism

This argument is based upon the idea that we should behave in such a way as to maximize the general happiness and minimize the general suffering. One must do a cost/benefit analysis for each research project. This position is very much in keeping with the spirit of the Nuremberg Code and Helsinki Declaration. The public consensus is that human happiness and reduction in suffering is more important than the unhappiness of nonhuman animals created by experiments. [What principles does one use to weigh the suffering of a nonhuman animals against the benefit to human happiness or reduction in human suffering?]

26.1.2 Arguments against nonhuman animal research

No benefit

One of the most often repeated arguments against nonhuman animal research is that it produces no benefits—that the results do not improve human (or nonhuman animal) health or happiness. For example: "Animal experiments mislead doctors and the general public. Diseases such as cancer, that are artificially induced in laboratory animals, have no relationship to the diseases which afflict humans. These are largely caused by lifestyle and pollution[1]." Therefore the suffering of the nonhuman animal subjects is unjustified. The real motive is profit for experimenters and/or sadism. [Virtually all of the current therapeutic drugs, surgical techniques, transplant procedures, vaccinations, diagnostic procedures, and new safety features in homes and vehicles were developed using research with nonhuman animals. If you encounter people espousing this view ask them if they would sign a card to carry in their wallet or purse, stating that if they are found unconscious or unable to speak on their own behalf that they want no procedure, treatment, or product used for their recovery that was developed using research with nonhuman animals.]

Lack of necessity

Many argue that research with nonhuman animals is unnecessary because the same information could be obtained by studying people, human tissue, human DNA, or computer models. [Think about where our information about DNA and its function came from. Where would we get the basic physiological information to put into our computer models? If you really understand very little about, say, pain pathways in the brain, how could you model them? Our ability to create experimental systems that replace nonhuman animals as subjects is based upon a copious amount of valid experimental data from experiments with nonhuman animals.]

Animal rights

This argument is based upon an analogy to certain aspects of the history of human rights. There was a time in this century in North America when certain groups of citizens were denied fundamental rights, such as the right to vote, etc. Two large groups

[1] Animal Liberation NSW website

were women and African Americans. The rationale was that these two groups essentially lacked some qualities that are prerequisite to exercising these rights in a responsible fashion. As it became clear that there were no such essential differences that would exclude them from having these rights, the fundamental rights were extended to them. So too, the rights that humans have in respect to being subjects in experiments are denied to nonhuman animals supposedly because nonhumans lack some essential quality. However, there is no relevant essential difference between humans and nonhuman animals. Therefore, nonhuman animals must be granted the same rights as humans. [This is an erroneous reconstruction of how women and African Americans won their rights. Rights are not all or none—perhaps there is some category of rights different from that of an intact adult human that nonhuman animals might properly fit.]

Animal liberation

A more extreme position is that by freeing nonhuman animals from all of our exploitation of them as pets (read slaves or prisoners), food, clothing, and subjects in experiments, we will also improve our treatment of other humans, and that this ethical elevation will continue to have positive vibrational effects such that there will be a continuing animal ethical evolution to the point that the lions will lay down peacefully with lambs, each of them enjoying a strict vegan diet. To give a flavour of this position:

> One of the spiritual experiences I had was when my dog Mitra was dying. I took him to his beloved Sangre de Cristo Mountains and there, as his true home, I laid him down in the soft grass among the aspen trees. I saw the magic of death unfolded as his body became lighter and lighter as if he allowed his corporeal existence to enter the magic of transformation. He was ready to abandon the density of his animal form for the final reunion with the great mystery. As Homo sapiens, we are beginning to examine our own moral and spiritual consistency. Animal liberation is human liberation. Every tree, every rock, every animal should be an object of reverence, this was how our ancestors lived. To guard the rights of nature and animals can serve as our own passage to spiritual maturity[2].

Also along these lines. we find people suggesting that nonhuman animals can have the capacity to show human, moral reasoning, and to enter into social contracts with humans:

> St. Francis once converted a wolf to reason. The wolf of Gubbio promised to stop terrorizing an Italian town; he made pledges and assurances and pacts, and he kept his part of the bargains. But St. Francis only performed this miracle once, and as miracles go, it didn't seem to capture the public's fancy. Humans don't want to enter into a pact with animals. They don't want animals to reason. It would be an unnerving experience. It would bring about all manner of awkwardness and guilt. It would make our treatment of them seem, well, unreasonable. The fact that animals are

[2]The Animal Liberation Movement and Global Spirituality, Dr. Mira Foung

> voiceless is a relief to us, it frees us from feeling much empathy or sorrow. If animals did have voices, if they could speak with the tongues of angels—at the very least with the tongues of angels—it is unlikely that they could save themselves from mankind.[3]

Animal welfare

A final position that accepts that some research with nonhuman animals is justified but only if consistent with strict guidelines is termed the Animal Welfare position. This position has many similarities to our current state of guidelines with humans who cannot give informed consent. It is based upon the Russell-Burch three Rs: reduction, refinement, and replacement. Reduction refers to taking whatever steps possible to ensure that the fewest animals are used consistent with obtaining valid scientific results. Refinement refers to adopting increasingly precise, less harmful or stressful, and less invasive procedures. Replacement refers to using invertebrates, tissue cultures, and computer models, whenever these are feasible, instead of a vertebrate animal. Prior to conducting an experiment, the procedures, rationale, and design must be evaluated, just as human work is, by an ethics review committee that is independent of the researchers proposing the experiment. Something akin to informed consent is provided by this committee whose job it is to ensure that the accepted guidelines and ethical principles are not violated. There are several extensive documents outlining contemporary principles and practices of nonhuman animal research. Two of the best representatives are the National Institutes of Health guidelines and the guidelines of the Canadian Council on Animal Care. Both of these guidelines are followed by researchers at the University of Lethbridge.

The following principles are excerpted from the NIH guidelines for nonhuman animal research (`http://oacu.od.nih.gov/regs/guide/guide2.htm`):

1. Animals selected for a procedure should be of an appropriate species and quality, and of the minimum number required to obtain valid results.

2. Proper use of animals, including the avoidance or minimization of discomfort, distress and pain, is imperative.

3. Procedures with animals that may cause more than momentary or slight pain or distress should be performed with appropriate sedation, analgesia or anesthesia. Surgical or other painful procedures should not be performed on unanesthetized animals paralyzed by chemical agents.

4. Postoperative care of animals should minimize discomfort and pain and, in any case, should be equivalent to accepted practices in schools of veterinary medicine.

5. Animals that would otherwise suffer severe or chronic pain or distress that cannot be relieved should be painlessly killed at the end of the procedure or, if appropriate, during the procedure. If the study requires the death of the animal, the animal must be killed in a humane manner.

[3] 'Do creatures have the same rights that we do?' By Joy Williams (C) 1997 *Harpers Magazine.*

6. Living conditions should be appropriate for the species and contribute to the animals' well-being. Normally, the housing, feeding and care of all animals used for biomedical purposes must be directed by a veterinarian or other scientist trained and experienced in the proper care, handling and use of the species being maintained or studied. In any case, appropriate veterinary care should be provided.

7. Exceptions to these principles require careful consideration and should only be made by an appropriate review group such as an institutional animal care and use committee.

Canada too has adopted national guidelines developed and promoted by the Canadian Council on Animal Care (`http://www.ccac.ca/`). Both public and private sector research are bound by these rules. The rules are summarized below from a CCAC statement approved in 1989:

> The use of animals in research, teaching, and testing is acceptable only if it promises to contribute to understanding of fundamental biological principles, or to the development of knowledge that can reasonably be expected to benefit humans or animals.
>
> Animals should be used only if the researcher's best efforts to find an alternative have failed. A continuing sharing of knowledge, review of the literature, and adherence to the Russell-Burch 3R tenet of Replacement, Reduction and Refinement are also requisites. Those using animals should employ the most humane methods on the smallest number of appropriate animals required to obtain valid information.
>
> The following principles incorporate suggestions from members of both the scientific and animal welfare communities, as well as the organizations represented on Council. They should be applied in conjunction with the Canadian Council on Animal Care's *Guide to the Care and Use of Experimental Animals.*
>
> 1. If animals must be used, they should be maintained in a manner that provides for their physical comfort and psychological well-being, according to CCAC's *Policy Statement on Social and Behavioural Requirements of Experimental Animals.*
> 2. Animals must not be subjected to unnecessary pain or distress. The experimental design must offer them every practicable safeguard, whether in research, in teaching or in testing procedures; cost and convenience must not take precedence over the animal's physical and mental well-being.
> 3. Expert opinion must attest to the potential value of studies with animals. The following procedures, which are restricted, require independent, external evaluation to justify their use:
>   (a) burns, freezing injuries, fractures, and other types of trauma investigation in anesthetized animals, concomitant to which must

be acceptable veterinary practices for the relief of pain, including adequate analgesia during the recovery period;

(b) staged encounters between predator and prey or between conspecifics where prolonged fighting and injury are probable.

4. If pain or distress is a necessary concomitant to the study, it must be minimized both in intensity and duration. Investigators, Animal Care Committees (ACC), grant review committees and referees must be especially cautious in evaluating the proposed use of the following procedures:

   (a) experiments involving withholding pre- and post-operative pain-relieving medication;

   (b) paralyzing and immobilizing experiments where there is no reduction in the sensation of pain;

   (c) electric shock as negative reinforcement;

   (d) extreme environmental conditions such as low or high temperatures, high humidity, modified atmospheres, etc., or sudden changes therein;

   (e) experiments studying stress and pain;

   (f) experiments requiring withholding of food and water for periods incompatible with the species specific physiological needs; such experiments should have no detrimental effect on the health of the animal;

   (g) injection of Freund's Complete Adjuvant (FCA). This must be carried out in accordance with *CCAC Guidelines on Acceptable Immunological Procedures.*

5. An animal observed to be experiencing severe, unrelievable pain or discomfort, should immediately be humanely killed, using a method providing initial rapid unconsciousness.

6. While non-recovery procedures involving anesthetized animals, and studies involving no pain or distress are considered acceptable, the following experimental procedures inflict excessive pain and are thus unacceptable:

   (a) utilization of muscle relaxants or paralytics (curare and curare-like) alone, without anesthetics, during surgical procedures;

   (b) traumatizing procedures involving crushing, burning, striking or beating in unanesthetized animals.

7. Studies such as toxicological and biological testing, cancer research and infectious disease investigation may, in the past, have required continuation until the death of the animal. However, in the face of distinct signs that such processes are causing irreversible pain or distress, alternative endpoints should be sought to satisfy both the requirements of the study and the needs of the animal.

8. Physical restraint should only be used after alternative procedures have been fully considered and found inadequate. Animals so restrained, must receive exceptional care and attention, in compliance with species specific and general requirements as set forth in the *Guide*.
9. Painful experiments or multiple invasive procedures on an individual animal, conducted solely for the instruction of students in the classroom, or for the demonstration of established scientific knowledge, cannot be justified. Audiovisual or other alternative techniques should be employed to convey such information.

Chapter 27

Reproduction and Evolution
Paul L. Vasey

Charles Darwin, the founding father of modern evolutionary thought, once remarked that "The sight of a feather in a peacock's tail whenever I gaze at it, makes me sick!" What was it about those feathers and that tail that troubled Darwin so?

Although the peacock's elaborate train of iridescent feathers is spectacular, to be sure, Darwin realized that there was a price to be paid for such flamboyant beauty. Significant amounts of energy must be diverted during development to produce such an exaggerated characteristic and once developed, the enormous tail (which can weigh up to four pounds) has to be carried around, necessitating even more energy expenditure. Walking on the ground and having to *schlep* such a tail around would seem cumbersome enough, but imagine how difficult it must be to take off in flight, let alone stay airborne! In addition, it seems likely that the showy tail feathers would attract a lot of attention; just think about the reaction they elicit from zoo-goers. For all intents and purposes, the peacock's tail would act like a lighthouse beacon calling forth unwanted predators who would be all too willing to partake in a tasty meal. Once located, the long, flowing tail would appear to be a perfect handle for some predator to grab hold of while dispatching its peacock prey. In other words, the peacock's tail seems to hinder its ability to survive rather than enhance it. Given all these costs, why on earth would the peacock tail have evolved at all?

The second problem Darwin faced with the peacock's tail was that only half of the members of the species—the males—possess the showy train of feathers. Darwin also noticed that when it came to their tails, peacocks behaved differently from peahens. Peacocks displayed their tails by fanning them out and then vibrating the feathers so that they shimmered, while turning around in a circle. The existence of sexually dimorphic morphological and behavioural traits (i.e., traits that differ between males and females) puzzled Darwin greatly. If both males and females lived in the same environment, then they presumably faced the same types of problems in terms of survival.

Yet peacocks look and behave like nature's Las Vegas showgirls, while females of the species, peahens, look more like nature's chimney sweeps. Indeed, it would not be surprising at all if the causal observer mistook the peacock and the peahen to be entirely different species. Why would the sexes have evolved sexually dimorphic characteristics? Why did peacocks have that bedazzling train of tail feathers, while the drab peahens had none?

To make matters worse, Darwin realized that the peacock's tail was not an idiosyncratic and isolated case of evolution gone wrong. Rather, it was only the tip of the iceberg. Examples of exaggerated, sexually dimorphic morphological and behavioural traits abound in the animal kingdom: the horns of male mountain sheep, the chest beating of the adult male gorilla, the croaking of male bullfrogs, and the facial hair of human males. The challenge for Darwin was to make sense of these intriguing characteristics.

27.1 Darwin's Theory of Sexual Selection

In 1871, Darwin published his book, *The Descent of Man, and Selection in Relation to Sex*, in which he outlined his now famous "Theory of Sexual Selection". Darwin formulated sexual selection theory to explain the enigma of the peacock's tail, and other exaggerated, sexually dimorphic traits like it. Such characteristics could evolve, Darwin suggested, even if they contribute absolutely nothing to survival. In fact, they might even be detrimental to the individual's survival. But how could individuals with such obvious "handicaps" exist? Simply put, Darwin reasoned that exaggerated, sexually dimorphic traits like the peacock's tail, existed because they enhanced males' ability to acquire female mates (i.e., reproductive partners), not because they helped males survive.

Sexual selection, Darwin argued, is a process of *differential reproduction among males of the same species that results because some males are better able to acquire reproductive partners than others*. What this means is that males of a particular species vary greatly in terms their lifetime reproductive success (i.e., the number of surviving offspring produced during a male's lifetime). In humans, for example, many men fail to reproduce at all. At the other end of the spectrum, the most reproductively successful man ever known to exist, Moulay Ishmail "The Bloodthirsty", former emperor of Morocco, sired 888 children. In between these two extremes, any number of possible reproductive outcomes exists for human males. Clearly then, the potential exists for males in most species to vary enormously in terms of their reproductive output. The reason for such variation is simple: not all males are equally adept at acquiring female reproductive partners; some are better than others.

How exactly do males acquire mates? Darwin argued that males compete among themselves (i.e., intra-sexually) in order to obtain access to female mates. He defined mate competition in rather broad terms, *as any behaviour that a male performs which increases the likelihood he will fertilize a female, at the expense of a rival male doing the same*. As such, Darwin recognized that mate competition could be quite large in terms of its scope. Males could, for example, compete intra-sexually for *direct* access to fertile females. However, fertile females need not be the immediate objects

of rivalry for mate competition to take place. Males can compete for female mates in more *indirect* ways as well. For example, they can vie for *control of resources* (e.g., food, nesting sites, shelter from predators) that females require. This control, in turn, would allow males to obtain exclusive access to any female that utilizes the resources under the males' control. Males could also compete for high *dominance* rank within their social group that would, in turn, allow them to obtain access to female mates prior to their subordinate rivals.

Darwin recognized that mate competition could take on a number of different forms. Most obviously, males can engage in contest competition by threatening or fighting with each other in order to: (1) gain exclusive access to a female mate, (2) obtain the resources needed to attract females mates, or (3) increase their position in the male dominance hierarchy. Contest competition involves males interacting directly with each other, but it does not necessarily require that they come into physical contact. For example, male fiddler crabs posses a sexually dimorphic pincer than constitutes one third of their body weight. By simply waving their enlarged pincer in the air, males can threaten same-sex rivals and cause them to flee the area. Contest competition can, however, escalate to the point of all-out physical aggression between males. Head-butting contests, during which male mountain sheep slam into each other's horns head-on, are a classic example. Physical aggression can be severe, resulting in deep flesh wounds and, in some cases, even death. Male macaques (a type of Old World monkey) have sexually dimorphic canine teeth that they use to slash the skin of other males during competitive interactions. In the macaque species I study, Japanese macaques, I have seen males lose fingers as a result of intense competition. As such, an entire gradient of contest competition exists starting from low-level, non-contact threats right up to high-level contact aggression.

Intra-sexual mate competition among males also can be manifested when males attempt to attract discriminating females so that the females will choose them as mates. This type of competition, termed *courtship competition*, often involves males performing highly stylized courtship displays that involve particular vocalizations, gestures, movements and facial expressions. For example, mountain sheep perform gentle "foreleg kicks" during which a male raises one front leg up and touches the underbelly of a female. In many bird species, such as the peacock, males can pique fertile females' sexual interests by performing ritualized courtship "dances" during which they display their elaborate plumage.

It is important to note that although successful competition among males can ensure the victor *access* to fertile females, taken alone, it does not guarantee that a male will actually copulate with those females it is able to access. For a successful copulation to occur, Darwin believed females had to discriminate carefully among males and choose the one with which they would copulate based on a pool of potential suitors. Thus, males are affected by sexual selection not only because they engage in intersexual competition for female mates, but also because females exercised *mate choice*, favouring some males as mates, while ignoring others.

Armed with sexual selection theory, Darwin argued that morphological and behavioural sex differences were readily explainable, and indeed, expected. Certain morphological and behavioural characteristics such as the horns of the male big-horn sheep and their head-butting contests, could evolve simply because they enhanced the males'

ability to compete with other males for: (1) direct access to female mates, (2) the resources needed to attract female mates, and/or (3) the dominance rank needed to access females prior to male rivals. Other morphological and behavioural characteristics, such as the peacock's tail and its courtship "dance", could have evolved, because they enhanced the males' ability to attract discriminating females into choosing them as their mates. Darwin reasoned that males who possessed such characteristics would access more female mates and, in turn, reproduce more than those that did not. The advantageous traits, if heritable, would be passed onto offspring and spread throughout the population.

27.2 Females and Sexual Selection

Darwin argued that sexual selection affected males, but it did not affect females. He based this argument on his belief that females did not compete for mates. Mate competition, Darwin believed, lead to inter-individual differences in life-time reproductive success because some individuals were better competitors than others. In the absence of inter-individual variation in reproductive success, sexual selection could not favour certain traits in certain females at the expense of others because all females reproduced at similar rates. Based on this of this logic, sexual selection could not be an important mechanism driving the evolution of females of the species. But maybe Darwin missed something.

Subsequent research demonstrated that sexual selection could operate on females because not all males were equally valuable as mates. In a series of experiments, Maynard Smith (1955) demonstrated that female fruit flies could discriminate among various male mates, preferring males who performed a particular courtship dance. Males that were not inbred were the best dancers, and it was these males that females preferred as mates. Interestingly, these same males had higher sperm counts than did inbred males. Thus, female fruit flies apparently choose males based on their courtship behaviour, which in turn was an honest indicator of their fertilizing ability. As such, females that preferred fancy dancers ended up having more of their eggs fertilized and, by extension, a higher lifetime reproductive success.

Maynard Smith's fruit-fly research demonstrates that for females, mate quality was more important than mate quantity. Choosing to mate a disadvantaged male imposed negative consequences on a female's reproductive success, whereas, choosing to mate with a more valuable male conferred reproductive advantages. So what should females choose? In theory, females should discriminate among potential male mates in favour of those that contribute the most to offspring quality and survival (Trivers, 1972). Thus, females should choose a male based on (1) ability to fertilize eggs (i.e., a mature, sexually competent male of her species), and (2) quality of genes (i.e., ability of genes to survive, reproductive ability of genes, compatibility of genes). In addition, in those species in which males help care for their young, females should prefer males based on the quantity and quality of their parental care (i.e., willingness and ability to invest, compatibility of parental attributes).

27.3 Parental Investment Theory

Although Darwin stated that males typically engage in mate competition, while females exercise mate choice, he never explained why these sex differences in patterns of mate acquisition existed in the first place. One hundred years after Darwin published his theory of sexual selection, Robert Trivers (1992) formulated his ground-breaking "Theory of Parental Investment" to explain why males and females differ in terms of how they acquire mates. Trivers (1972) argued that sex differences in parental investment could explain sex differences in patterns of choice and competition. He defined parental investment as *any investment by a parent in an offspring that increases the offspring's changes of survival (and hence, future reproduction) at the cost of the parent producing another offspring*.

Trivers (1972) noted that in the "typical" species, females provide considerably more parental investment than males. Gestation, parturition (i.e., giving birth), lactation, protection, feeding and carrying are all reproductive "costs" that females incur in relation to having offspring. Females that become pregnant can potentially incur these costs for years. In contrast, males can simply walk away from a casual copulation having lost only a few minutes of their time. One consequence of this sex difference in parental investment is that females are occupied for much of their lives producing and caring for offspring, while males are not. As such, females have only the potential to produce a small number of offspring relative to males who can produce many (remember Moulay Ishmail "The Bloodthristy" and his 888 children!). Another consequence of this sex difference in parental investment is that at any one time, there will be fewer fertile females in a population relative to reproductively active males. In other words, a lot of males will be looking to copulate with fertile females, but only a few fertile females will be available to do so. It is not surprising then that females are sometimes characterized as valuable "resources" that limit male reproductive success.

Given these sex differences in parental investment and the consequences they entail, males and females typically employ different mate acquisition strategies. Generally, males have a relatively low threshold for sexual arousal, they copulate relatively indiscriminately, and they compete intra-sexually for reproductive access to females. Females, on the other hand, exhibit sexual restraint relative to males, and they rarely, if ever, compete for reproductive access to males. Instead, they discriminate among potential reproductive partners in favour of those males which contribute the most to the quality and survival of their offspring (Trivers, 1972).

An examination of "atypical" species in which males provide the bulk of parental care, lends support to Trivers' (1972) theory of parental investment. Such "sex role reversed" species are few and far between. However, examples include pipefish, mormon crickets, moorhens and certain species of frogs. In such species, it is the males that are choosy about the females with which they copulate, while the females compete among themselves for access to a male—just as parental investment theory would predict (Gwynne, 1991).

27.4 Sexual Coercion

How the sexual selection process occurs has been overwhelmingly conceptualized in terms of two interactive mechanism: intra-sexual competition among males for female mates, and choice of male mates by females (Darwin, 1871). Although other mechanisms such as male choice of female mates have long been recognized (Darwin, 1871, p. 263), they have attracted relatively limited research attention. More recently, however, there has been a growing recognition that these alternative mechanism for sexual selection need to be formally acknowledged and investigated.

In particular, researchers have begun to ask what happens when a male is unable to access a female mate: is it because he is unsuccessful at contest competition, courtship competition, or both? Studies show that in a wide variety of species the solution to this problem is a violent one: sexual coercion (Smuts & Smuts, 1993). Sexual coercion occurs when a male uses force, or the threat of force, against a female to increase the likelihood that he will mate with her and some other male will not.

Perhaps the most dramatic example of sexual coercion in animals comes from orangutans (Mitani, 1985). In this species, most copulations by sub-adult males and nearly half of all copulations by adult males occur after the female's fierce resistance (e.g., slapping, screaming, hitting, and bitting) has been overcome through violent restraint. At the other end of the spectrum, males can coerce females into mating with them simply be following the female continuously for days on end. I watched a Japanese macaque male do just this. He shadowed an uninterested female for days, always remaining within a few meters of her. Because this male was dominant, no other males dared to approach the female, and she was simply unable to mate, or have normal social interactions with other group members. The harassing male would even disrupt the female from foraging for food. Finally, after five days of constant stalking, the beleaguered female gave up and allowed the male to copulate with her. Having obtained what he wanted, the male abandoned the female's side immediately after he ejaculated.

In all instances, males use sexual coercion to override female choice. In theory, males that behave in this manner are more likely to reproduce, relative to males that do not and, as such, sexual selection will favour the evolution of sexual coercion in males. Indeed, a survey of the animal kingdom indicates that sexual coercion is a widespread behavioural pattern. Researchers have suggested that the success of sexual coercion is why males tend to have lower thresholds for sexual arousal than females. If we look at humans, for example, men are much more willing than women to have sex with a member of the opposite sex one hour after meeting him or her for the first time (Buss & Schmitt, 1993).

So, males are thought to benefit reproductively from sexual coercion. Females, on the other hand, incur costs when they are the victims of sexual coercion; they are unable to select carefully the male with whom they will produce offspring. In theory then, sexual selection should favour the evolution of female counter-strategies against coercive males. In many species, this result is exactly what we see. For example, in Japanese macaques, females sometimes respond to the screams of a victim of sexual aggression by "mobbing" her aggressor. Faced with tens of females pouncing on him, the male typically beats a hasty retreat.

27.5 *Inter*-Sexual Mate Competition?

Evolutionary studies of sexual behaviour tend to focus on reproductive sex. This emphasis is understandable to some extent, because differential reproduction is the engine that drives evolution. Without reproduction there can be no evolution. Consequently, the study of reproductive sex makes good sense from an evolutionary perspective, but it is important to remember that reproduction does not occur in isolation. Animals engage in a myriad of non-reproductive sexual behaviours as well (Bagemihl, 1999). These behaviours can influence patterns of potentially reproductive sex, and by extension, evolution. As such, any comprehensive understanding of reproduction and evolution necessitates understanding the relationship of reproductive sex to various forms of non-reproductive sexuality.

The species I study, Japanese macaques, represents an excellent model for examining the relationship between reproductive and non-reproductive behaviour, because females of this species engage in heterosexual and homosexual behaviour through their lives. More specifically, they engage in mounting with both males and females within the context of consortships—temporary, but exclusive sexual relationships which can last for several days. Females can mount, and be mounted, hundreds of times during these relationships (see Figure 27.1). This behaviour presents somewhat of a dilemma for male Japanese macaques. Unlike most species in which males only have to compete with other males for reproductive opportunities, in Japanese macaques, males have to compete with both males *and*females (Vasey, 1998). In other words, male Japanese macaques have to engage in *inter*-sexual (i.e., male-female) competition for mates, in addition to engaging in intra-sexual competition.

What does inter-sexual competition for mates look like in Japanese macaques? Typically, a male competitor will intrude on a female homosexual consortship and attempt to disrupt the females' sexual activity. During these intrusions, the male will target one of the female consort partners as his competitor, and the other as his desired mate. The intruder then attacks the female competitor and solicits the desired mate for sex. Following about one third of these intrusions, the female competitor retaliates against the male by attempting to drive him away from her female consort partner.

As with intra-sexual competition for mates, it appears that inter-sexual mate competition involves behaviours that influence the acquisition and retention of reproductive partners through interactions with conspecifics (i.e., members of the same species). It is not immediately apparent, however, why male Japanese macaques would engage in inter-sexual mate competition. After all, female consort partners can't reproduce together, so they pose no direct reproductive threat to the males. Moreover, by engaging in these competitive interactions, males risk injury. In light of such considerations, shouldn't males avoid inter-sexual mate competition altogether and instead simply approach potential mates immediately after their homosexual consortships have terminated? This possibility might be the case were it not for the fact female homosexual consortships in Japanese macaques occur in the context of *intra*-sexual mate competition. In such a context, the potential always exists for males to lose copulations to their same-sex rivals. Consequently, males able to access mates engaged in homosexual consortships, prior to same-sex rivals, would gain a reproductive advantage.

My research shows that inter-sexual competition for mates occurs in about sixteen

Figure 27.1: Female homosexual mounting in Japanese macaques. The mountee facilitates the mount by reaching back to grasp the mounter. The mountee's immature offspring huddles below his mother (Photo by author).

other species of birds and mammals, including humans. Among the Indians of the north-west Amazon, for example, both males and females try to seduce young unmarried men that visit from other villages (Sorenson, 1984). Women attempt to ambush the visiting men if they happen to see them going to bathe alone. Men, in turn, try to seduce the visitors when they go to bathe to keep them away from the local women. This strategy effectively lessens the chances that the visiting men will be seduced by village women.

27.6 Conclusion

Sexual selection has favoured sex differences in patterns of mate acquisition. In "typical" species, males compete intra-sexually for female mates. Females, on the other hand, carefully discriminate among potential male suitors. When males cannot access female mates through contest or courtship competition, they may attempt to coerce them sexually. Females may, in turn, demonstrate counter-strategies to circumvent sexually coercive males. These sex differences in mate acquisition tactics ultimately stem from sex differences in patterns of parental investment. In some species, such as Japanese macaques and humans, the ability to compete for mates is a more generalized capacity that extends into the inter-sexual arena.

Chapter 28

Sex and Gender Diversity
Paul L. Vasey

Imagine you walk into a government office to renew your driver's license and are required to fill out a personal information data sheet. One of the questions reads as follows:

> Please check one of the following:
> Gender: ☐ Male
> ☐ Female

You hesitate and then write "Neither" beside the question.

What's the problem? The problem is that the question asks for information about your *gender*, but then provides you with only two possible answers, both of which pertain to *sex*.

Many of us confuse sex with gender, and vice versa, in our everyday lives. We often make no distinction between the two terms, and use them interchangeably, as if one were synonymous with the other. In reality, however, sex and gender do not refer to the same phenomenon and one should not be equated with the other. This distinction isn't simply an arcane linguistic point of interest for grammar mavens. As I hope to demonstrate, the definitional sloppiness associated with sex and gender is particularly problematic for psychologists and other scientists who are interested in understanding the combined influences of biology and culture on human behaviour.

28.1 Defining Sex and Gender

The terms sex and gender have quite different etymologies. *Sex* derives from a group of Indo-European words that collectively refer to "dichotomous relationships in which two categories are divided, one from the other." It comes from the Latin "sexus,"

meaning "either of two divisions of organic nature distinguished as male or female respectively." The word *gender*, on the other hand, is derived from a group of Indo-European words that collectively refer to "systems of classification in which there may be many categories." It comes from the Latin "genus," meaning "kind or race" (Walker & Cook, 1998).

In contemporary terms, the correct use of the term *sex* is in reference to an individual's biological status as defined by one, or some combination of, the following parameters: (1) sex chromosomes, (2) sex hormones, (3) gonads, (4) internal reproductive organs, (5) gametes, (6) external genitalia, and (7) secondary sexual characteristics. All cultures recognize the two sex categories *male* and *female*. There is no culture that recognizes only one sex.

Inclusion into the category of male or female depends on the constellation of biological parameters that an individual possesses. Males are typically characterized by *all* of the following: XY sex chromosomes, a preponderance of circulating androgen, two testes, two vas deferens, two seminal vesicles, two epididymis, two Cowper's glands, one prostate gland, sperm production, a penis, a scrotum, and various secondary sexual characteristics including the development of facial hair and increased shoulder width at puberty. Females, in contrast, are typically characterized by *all* of the following: XX sex chromosomes, a preponderance of circulating estrogen, two ovaries, two fallopian tubes, one uterus, one cervix, ova production, a clitoris, a vagina, two labia majora, two labia minora and various secondary sexual characteristics including the development of breasts and a widening of the hips at puberty.

Although we currently have the technology to assess each of these biological parameters, we typically do not examine all of them when making sex determinations. Rather, the precise combination of biological criteria used to define whether one is a male, or a female, varies cross-culturally and historically (Dreger, 1998a; Fausto-Sterling, 2000). In our culture, for example, sex is assigned at birth following inspection of the external genitals. After puberty, secondary sexual characteristics (e.g., breasts, facial hair, shoulder width, hip width, bone and muscle structure) provide important clues about the sex that we perceive an individual to be (Devore, 1989).

In contemporary terms, the correct use of the term *gender* is in reference to culturally constructed categories of people that are determined, in part, by perceived sex distinctions. In other words, gender is a system of separating people into groups based, in part, on the sex that others perceive them to be. All cultures recognize the two gender categories, *boys/men* and *girls/women*, which are subdivided on the basis of age. There is no culture that recognizes only one gender.

As with sex, inclusion into a particular gender category is initially assigned at birth on the basis of external genitals. For example, following inspection of the external genitals, babies with penises and scrotums *become* baby boys and are traditionally dressed in blue clothing that is deemed masculine and thus, appropriate attire for boys. Babies with clitorises, vaginas and labial folds *become* baby girls and are traditionally dressed in pink clothing which is deemed feminine and thus, appropriate attire for girls.

Once this initial assignment occurs different "rules" of conduct (i.e., *gender role expectations*) exist for the members of these two gender categories. Gender role expectations vary cross-culturally and historically. In our culture, for example, the expectation exists that boys and men will not wear dresses, nor will they carry their books

on their hips when walking; girls and women will. Similarly, it is expected that girls and women will not aspire to become Catholic Popes or swim topless at the public pool; boys and men will. Behaviours that are considered to be appropriate for boys and men are labeled *masculine*, whereas, those that are considered appropriate for girls and women are labeled *feminine*. Following infancy, gender is primarily attributed to an individual on the basis of the constellation of masculine and feminine behaviours they actually manifest in public (i.e., their *gender role enactment*,Devore (1989)). More often than not, individuals' gender role enactment conforms to their culture's gender role expectations. As such, individuals that are perceived to be boys/men typically engage in masculine behaviours, whereas individuals that are perceived to be girls/women typically engage in feminine behaviours.

28.2 Changing Sex

Whether individuals can change their sex or not depends on how sex is defined. If, for example, we define sex solely on the basis of sex chromosomes then we would have to conclude that an individual cannot switch from one sex to another. Presently, we simply do not have the technology to alter an individual's complement of chromosomes. Similar arguments against the possibility of ever changing one's sex can be made if sex is defined in terms of gonads, internal reproductive organs and/or gametes.

These internal markers of sex may be ascertainable in the laboratory, but in our day to day lives, they simply don't count (Devore, 1989). In Canada, for example, sex is *legally* defined on the basis of external genitals: males have penises and scrotums, while females have vaginas, clitorises and labial folds. Secondary sexual characteristics such as facial hair, breasts, bone and muscle structure are also important criteria that we use in social situations to determine what sex we think an individual is likely to be. If we define sex solely in terms of these external criteria (which is exactly what we do on a day-to-day basis), we would have to conclude that individuals *can* change from one sex to another. Indeed, legally speaking, this is very much the case in Canada.

For example, *transsexuals* are individuals that feel uncomfortable with their natal sex (i.e., the sex they were born as) and prefer to live as members of the other sex (Bolin, 1996; Devore, 1998). They undertake a variety of hormonal and surgical procedures aimed at making their external genitals and secondary sexual characteristics (e.g., body hair, breasts, facial structure, trachea or "Adam's apple", etc.) conform in appearance to those of the opposite-sex into which they were born. Many post-operative transsexuals go on to live as full-fledged members of their chosen sex and "pass" entirely unnoticed by other members of their society (Bolin, 1996; Devore, 1998).

The existence of transsexuals underscores the important point that an individual's *sex identity* (i.e., the sex that they feel themselves to be) may not be concordant with the biological reality of their natal sex.

28.3 Changing Gender

Gender is often incorrectly assumed to be a natural and inevitable social manifestation of an individual's biological sex. It is important to note, however, that masculine and feminine behaviours can be "performed" by either sex. Consequently, it is entirely possible for a male to engage in predominantly or exclusively feminine behaviours and to pass unnoticed in public as a girl/woman. Likewise, it is entirely possible for a female to engage in predominantly or exclusively masculine behaviours and to pass unnoticed in public as a boy/man (Devore, 1989). Indeed, some *transgendered individuals* (or transgenderists) are uncomfortable with the gender they were assigned at birth and, consequently, they switch to living, part-time or full-time, as members of the opposite-gender (i.e., *gender crossing*, (Bolin, 1996)). Transgendered individuals that cross from one gender to the other may pass entirely unnoticed by other members of their society and, in doing so, live in accordance with the gender role expectations associated with their chosen gender (Bolin, 1996). Others may endeavor to live unnoticed in their preferred gender, but be much less successful at passing, often with violent results. A recent, real life example of the latter is Brandon Teena, a gender-crossing female who was murdered in Falls City, Nebraska, because she was transgendered. A fictionalized account of Brandon Teena's life and death is depicted in the Hollywood film "*Boys Don't Cry.*"

Individuals need not engage in gender-crossing to be considered transgendered. For example, some individuals may be uncomfortable with their assigned gender and may, as a result, engage in various combinations of masculine and feminine behaviours that results in others perceiving them as ambiguously gendered (i.e., *gender mixing*, (Devore, 1989; Bolin, 1996)). For example, *androgynous individuals* are simultaneously, highly feminine and highly masculine (Bem, 1974). The model/singer/actress Grace Jones is a good example of an androgynous individual with her crew-cut hair and full, red lips. *Gender undifferentiated individuals* are simultaneously not particularly masculine and not particularly feminine (Bem, 1974). The character, "Pat," from the television program *"Saturday Night Live"* is a good example of a gender undifferentiated individual. The show's running joke was that everyone tried increasingly inventive ways to determine whether Pat was a man or a woman, but Pat's ambiguous behaviour squelched these efforts and left everyone scratching their heads and wondering. By enacting gender in this manner, androgynous and gender undifferentiated individuals draw attention to themselves and live publicly in ways that flout gender role expectations for how they should behave.

It is important to note that transgendered individuals are not transsexuals and thus, they do not desire to alter their external genitals. As such, they demonstrate that *gender identity* (i.e., the gender an individual perceives himself or herself to be) is not necessarily concordant with sex identity. In other words, individuals may be uncomfortable with their assigned gender, but perfectly comfortable with their assigned sex. A great example of this occurs in the film, "*The Crying Game.*" One of the characters, Dil, is a gender-crossing male who passes very well as a woman. When confronted about the reality of "her" biological sex, Dil seems quite unconcerned about "her" penis and simply comments "Details, details, baby!"

28.4 Are Two Sexes Enough?

In Western society sex is conceptualized in a dichotomous manner. There are two sexes and they are diametrically opposed to one another. You are either a male, or you are the "opposite-sex," a female. There are simply no other options. Indeed, this seems so self-evident that most people would think it is not even worth mentioning. You can't be both sexes; you can't be neither sex; and you certainly can't be some third option. Or can you?

Consider, for example, the following case of an intersexed individual, Chris, who has *Androgen Insensitivity Syndrome* (AIS). Chris has XY sex chromosomes and testes, both of which are male-typical. Chris's body produces male-typical sex hormones, called androgens, but is "blind" to their presence because the receptors that "read" androgens are not functioning. The internal reproductive organs that Chris possesses are male-typical to some extent, but they never developed in the "absence" of androgens. Chris has female-typical external genitals and, as such, she possesses a vagina, a clitoris and labial folds. At puberty, Chris's body responded to small amounts of naturally occurring estrogen by developing female-typical secondary sexual characteristics such as breasts and broad hips. Chris is infertile and has never menstruated. So what sex is Chris? Is she a female? Or, is he a male? Or is Chris something else altogether? How you answer this question depends entirely on the primacy you give to the various criteria used to measure biological sex.

Intersexed individuals, like Chris, have a mixture of male and female sex characteristics. Their very existence suggests that some individuals are neither male, nor female. Interestingly, when questioned about their sex identity, many intersexed individuals express the belief that they are neither male, nor female, but something else that they call "intersexed" (Kessler, 1998). If we accept the argument that intersexed individuals are neither male, nor female, then we would have to conclude that sex is not dichotomous as is commonly assumed, but rather, that it exists on a continuum. According to this perspective, males and females are the most common expressions of sex and they exist on opposite ends of a sex continuum. However, a myriad possible, albeit less common, intersexed bodies exist in between these two extremes (Kessler, 1998; Fausto-Sterling, 2000).

Our dichotomous view of sex can have negative repercussions for individuals that cannot be easily categorized as males or females. In Western cultures, including here in Canada, a "state of emergency" is declared when an intersexed baby with ambiguous genitalia is born (Fausto-Sterling, 2000, : p. 275, note 1). A team of medical experts is rapidly assembled and the child's genitals are surgically "normalized" as soon as possible so that they resemble male- or female-typical genitals. Follow-up studies of intersexed individuals that have undergone "corrective" surgery indicate that they suffer significant physical and psychological effects as a result of these operations (Dreger, 1998b). An intersexed child's ambiguous genitals do not threaten that child's life, so we might ask why are such invasive surgeries tolerated given that they seem medically unnecessary and cause so much physical and psychological damage? Some researchers argue that we permit the surgical "normalization" of intersexed infants because their ambiguous genitals threaten our culture's expectations concerning the dichotomous nature of sex. Faced with this threat to the "logic" of our two-sex system, we attempt

to make intersexed infants either male or female.

The dichotomous view of sex that characterizes Western culture is not universal. In some areas of the Dominican Republic, for example, three sexes are recognized: males, females and guevedoche (gway-va-doe'-chay – meaning "eggs at twelve"). Guevedoche are individuals with XY sex chromosomes that have a rare condition known as *5-alpha reductase deficiency*. Individuals with this condition do not produce 5-alpha reductase, an enzyme that converts testosterone into 5-alpha-dihydotestosterone. In the absence of 5-alpha-dihydotestosterone, the external genitals of XY individuals do not develop in a male-typical fashion and appear ambiguous at birth. At puberty, high blood level concentrations of androgens result in some penile growth, the descent of the testes (aka the "eggs"), and the development of male-typical secondary sexual characteristics. In the Dominican Republic these individuals are commonly identified at birth as guevedoche, not as males or as females (Herdt, 1996). Typically, they are raised as girls and then at puberty, they adopt masculine gender identities and behave as men. Thus, the Dominican Republic provides us with an example of a culture that has three sexes (male, female, guevedoche) and two genders (boy/man, girl/woman). Males and guevedoche both become men. This underscores an important point: the criteria we use to determine sex, gonads and genitals and the like, are biological realities to be sure, but the manner in which sex is actually defined is culturally determined (Herdt, 1996; Fausto-Sterling, 2000).

28.5 Are Two Genders Enough?

As with sex, Western society has traditionally conceptualized gender in a dichotomous manner. This perspective holds that there are two genders and they are diametrically opposed to one another. You are either a boy/man or, you are the "opposite-gender," a girl/woman. There isn't any middle-ground. We might identify the singer k.d. lang as a "butch" (i.e., masculine-acting) woman, but she is a woman nonetheless. As such, most people in your Psychology 1000 class would say that you can't be both of these genders; most would say you can't be neither of these gender; and most would certainly deny the existence of some third option.

In Western cultures, individuals that behave in a gender atypical manner are pathologized by the psychiatric profession. In Canada, for example, children that behave in a manner that is more consistent with the expectations of the opposite gender to which they were assigned at birth are diagnosed as having a mental disorder called "Gender Identity Disorder" (GID) (for a full list of the diagnostic criteria for GID see the 4th Edition of the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders). Children diagnosed with GID are "treated" within a clinical setting with various therapies aimed at making them behave in ways that conform to Western culture's gender role expectations. One such therapy, for example, encouraged boys with GID to walk like the 1970's television character, Steve Austin–*"The Six Million Dollar Man"* (Hay, Barlow, & Hay, 1981). It seems reasonable to expect that the enforced repudiation of one's gender identity and accompanying behaviours would cause such children a great deal of distress. Reflecting on similar treatments, Daphne Scholinski (1998, p. *x*), a woman who was institutionalized throughout much of the 1980's

after being diagnosed with GID as an adolescent, writes of feeling "…so false your own skin is your enemy." Interestingly, research done by my colleagues and myself indicates that there is no scientific support for the conclusion that gender-atypical children diagnosed with GID have a mental disorder (Bartlett, Vasey, & Bukowski, 2000). Why then do we tolerate imposing such behavioural therapies on children given the potentially negative message it sends to them about their self-worth? My colleagues and I have argued that, quite simply, we pathologize such children because they threaten to destabilize the particular gender role boundaries Western culture legitimizes.

The dichotomous view of gender that characterizes Western culture is by no means universal. Some cultures recognize three or four genders, as opposed to two (Nanda, 2000). These "alternative" gender categories are defined according to various parameters in association with sex distinctions. These parameters can include: (1) the type of work performed, (2) the religious activities in which one engages, (3) the position employed during sexual intercourse, (4) whether one is sexually abstinent or not, (5) one's sexual orientation, (6) the clothing and appearance one adopts, and (7) the mannerisms (e.g., timbre of voice, gait, posture, gestures) that characterize an individual. In a number of non-Western cultures, alternative genders are referred to with specific linguistic terms (e.g., alyha, hwame, hijra, sadhin, bakla, kathoey, mahu, fa'fafafine, etc.) which denote that they are neither men nor women, but rather a separate gender category (Nanda, 2000).

In Oman, for example, three genders are recognized: boy/man, girl/woman, and xanith (za'-nith; Wikan, 1977). Xanith are biological males that members of their culture recognize as being neither boys/men, nor girls/women. In Omani culture, the ultimate criterion for defining a male's status as a man is his ability to engage in penile penetration of sexual partners, be they female or male. Consequently, males that do not penetrate their sexual partners, but instead take the receptive role during anal intercourse, are not men, they are xanith. Thus, it is sexual position (in association with biological sex), not sexual orientation, that defines one's gender in Omani culture.

In terms of their gender role enactment, xanith engage in a mix of masculine and feminine behaviours (Wikan, 1977). For example, Omani culture is characterized by a great deal of sex segregation; men and women do not interact except if they are close kin. Instead, men interact primarily with other men, and women interact primarily with other women. If a man who is not a close kin member comes to a woman's home and her husband is not there, the woman refuses to let the man in. Likewise, an Omani women would never speak to a man on the street unless the man was a brother, father, husband or son. But xanith are not men, nor are they women, and as such, they are free to interact with *both* men and women in private and in public spaces.

In other ways, the gender role enactment of xanith cannot be described as gender mixing because the behaviours they manifest are regarded as being neither masculine, nor feminine, in Omani culture. For example, Omani men wear "masculine" white robes and women wear "feminine" patterned robes with bright colours. Xanith do not "mix" these masculine and feminine styles of dress. Instead, they dress in a manner that is neither masculine nor feminine; they wear unpatterned robes with pastel colours. Similarly, Omani men brush their hair back from their face, whereas, women part their hair in the middle. Xanith, on the other hand, part their hair on the side.

The Omani case is instructive because it provides us with an example of a culture

that has two sexes (male, female) and three genders (boy/man, girl/woman and xanith). Males can become either men or xanith. This underscores the important point that gender is culturally constructed and as such, males do not necessarily have to be men (and females do not necessarily have to be women). Moreover, the xanith demonstrate that alternative ways of thinking about gender do not necessarily require that we make reference to masculine or feminine qualities, or even some combination of the two. Instead, gender can be enacted in culturally specific ways that are considered to be neither masculine nor feminine. As Westerners it can be extremely difficult to conceptualize that such possibilities even exist.

28.6 Concluding remarks

Sex and gender do not refer to the same phenomenon and they should not be used as synonyms. Sex refers to an individual's biological status as defined by various parameters including: sex chromosomes, sex hormones, gonads, internal reproductive organs, gametes, external genitals and secondary sexual characteristics. Gender refers to culturally constructed categories of people that are determined, in part, by perceived sex distinctions. By differentiating between these two phenomena, we can better disentangle the combined influences of biology and culture on human behaviour.

In Western society, we tend to think about sex and gender as dichotomous. Everyone is either a male or female and, likewise, everyone is either a boy/man or girl/woman. All males are boys/men, and all females are girls/women. By examining the exceptions to these sex and gender "rules"—transsexuals, transgenderists, intersexuals and alternative genders like xaniths—we can better understand the rules themselves.

28.7 Website Resources

For information on male-to-female transsexuals you can visit the transsexual women's resource website at: <www.annelawrence.com>.

For information on female-to-male transsexuals you can visit the FTM international website at: <www.ftm-intl.org>.

For information about transgender individuals you can visit the website: <www.gendertalk.com>.

For information about intersexuality you can visit the Intersexed Society of North America's website at: <www.isna.org>.

Chapter 29

Subliminal Messages
John R. Vokey

Is it possible that the meaning or content of an event can affect people's behaviour without their being aware of the event? Can briefly-flashed messages in film and video commercials, embedded sexual imagery in print and video advertising, sub-audible messages in self-help audiotapes, or backward or barely audible messages in rock music lead people to make purchases they hadn't intended to, stop smoking, or commit suicide? Is it likely that such techniques are actually more effective than promotional messages of which people are aware? Subliminal persuasion refers to the use of hidden or otherwise imperceptible or masked stimuli to manipulate viewers or listeners to behave in ways they otherwise would not. Of the many folk psychology beliefs, the notion that such methods are an effective means of controlling people's behaviour is one of the most pervasive, and an inevitable topic of discussion in any course in introductory psychology.

Part of the wide-spread fascination with subliminal persuasion is undoubtedly its insidious nature and its ready confirmation of what many already believe to be the unscrupulous methods of advertisers, public relations experts and large corporations. A large part probably is also attributable to the phenomenal popularity of the books and college-circuit lectures of Wilson Bryan Key (1973, 1976, 1980, 1990), a former professor at the University of Western Ontario, and the leading proponent of the belief of a major conspiracy among advertisers and product manufacturers to manipulate the unsuspecting public through subliminal methods.

Another part is no doubt due to the rapidly expanding, $50-million per year market in subliminal self-help tapes, which have produced numerous testimonials to their effectiveness in promoting, among many other things, weight loss, breast enlargement, improvement in sexual function, self-esteem, and improved bowling scores (Pratkanis, 1992; T. E. Moore, 1992). And part, too, is most likely attributable to the influence in books and well-publicized public sermons and lectures of fundamentalists and other conservative religious and political zealots who have promulgated the belief of oc-

cult, satanic messages in rock music,[1] and who, in addition to promoting mass record-burning rallies to destroy the offending material, have lobbied, sometimes successfully, for the requirement that all such rock music be affixed with prominent warning stickers attesting to its subliminal content, as a "consumer protection act" (McIver, 1988; Vokey & Read, 1985).

We will touch briefly on each of these subliminal techniques, but first we will discuss the one study most often cited by proponents of these beliefs as "proof" of the effectiveness of subliminal persuasion.

29.1 The Vicary "Eat Popcorn/Drink Coke" Study

Popular discussion of subliminal persuasion inevitably appeals to James Vicary's notorious "Eat Popcorn/Drink Coke" study of the mid-1950s. Indeed, only months after results of the study were made public in newspapers and magazines, a survey of the American public revealed that already 41% of respondents had heard of subliminal advertising; by the 1980s over 80% reported being aware of the term, with roughly 70% of those believing it to be effective in increasing sales (Pratkanis, 1992).

Unfortunately, except for a credulous summary prepared for an article in a magazine for high-school students, there has never been a primary publication of the Vicary study, despite repeated demands at the time from sceptical, professional advertisers and research psychologists that Vicary do so. Consequently, the study has never been subjected to any proper scientific review. However, what can be gleaned about the study from the published reports in newspapers and popular magazines of the day is as follows. In 1956, in a movie theatre in Fort Lee, New Jersey, James Vicary—a social psychologist and advertising expert—and his Subliminal Projection Company, conducted six weeks of studies, involving thousands of unsuspecting movie-goers, to test a device that secretly flashed the messages "Eat Popcorn" and "Drink Coke" for a third of a millisecond[2] every five seconds during the film. Vicary claimed an almost 58% increase in sales of popcorn, and an 18% increase in Coke sales (Pratkanis, 1992), although why the technique was so much more effective for sales of popcorn than for sales of coke was never explained.

Reports of the study provoked immediate outrage. In an influential article that reflected the deep concerns of many Americans, Norman Cousins (1957) warned of the serious consequences of such a device, and argued that it should not be allowed. The U. S. Federal Communications Commission (FCC) undertook an immediate investigation, and ruled that the use of Vicary's techniques could result in the revocation of a license to broadcast[3]. Members of the U. S. National Association of Broadcasters were

[1] The primary meaning of the term "occult" is simply "hidden"; thus, to say that there are "occult" messages in rock music recordings is to say nothing more than that the recordings contain "hidden" messages, as is common to other such creative productions as poetry, visual, and performance art. Unfortunately, many automatically read the secondary meaning into the term, that of "preter-" or "supernatural", and from there it is but a simple, uncritical slide to "satanic messages".

[2] That's 0.00033 of a second!

[3] Despite an announced "deep concern" about the use of subliminals, the FCC subsequently did not pass

prohibited from using subliminal advertising, and it was banned outright in Britain and Australia.

Probably the most important fact about the 1956 Vicary study, however—and one never mentioned by proponents of subliminal persuasion—is that it apparently never happened: in a 1962 interview, Vicary admitted that he had made the whole thing up! His company did have the claimed device, but as Vicary stated in the interview "... we hadn't done any research, except what was needed for filing a patent. I had only a minor interest in the company and a small amount of data—too small to be meaningful." So much for the thousands of subjects, and the large increases in sales (Pratkanis, 1992).

More than the fact that the study had never happened, is the fact that in subsequent attempted demonstrations of his machine as a way of responding to critics, Vicary frequently failed to get the machine to work at all, and when he did, no one in the audience felt the least bit compelled to comply with the flashed messages. Furthermore, when the machine was finally subjected to a controlled test by an independent research company, no increase in the sales of either popcorn or coke was observed (Pratkanis, 1992). In 1958, the Canadian Broadcasting Corporation (CBC) tested Vicary's claims by subliminally flashing the message "Call Now" 352 times during the popular Sunday-night program *Close-Up*. No increase in telephone calls was observed, and no one called the station, although when asked to guess the message, almost one-half of the roughly 500 viewers who sent letters claimed to have been made hungry or thirsty during the show. Not one letter writer, however, guessed the correct message (Pratkanis, 1992). Another company, Precon Process and Equipment, began in 1957 to use subliminal messages on billboards and in movies, and received a patent for the technique in 1962 (McIver, 1988). The patent was awarded because the device could do what was claimed for it—in this case, insert subliminal messages—not because the subliminal messages themselves were shown to be effective.

29.2 *Mad Max, ALF* and Garfield

Because of the credulous reports of the now-apocryphal Vicary study, many advertising, television, radio, and film companies began using subliminal messages, often in the belief that it would enhance the effect of some scene or message, and many still do. In a recent investigation of alleged subliminals in video and film, for example, Poundstone (1993) freeze-framed his way through such films and television shows as *Mad Max* and the cartoon version of *ALF*, both of which clearly do evince the subliminals claimed for them. In *Mad Max*, the subliminal (a close-up of a face showing an impossible widening of the eyes of a biker so that even the conjunctiva were showing just before he crashes into an oncoming truck) was clearly intended to enhance the impact[4] of the scene, but most of the subliminals Poundstone found, such as the many in the ALF cartoons, were clearly inserted as gags by the producers. In one fleeting image in

any such regulation, preferring instead to "pass the buck", arguing that the use of subliminals properly was under the jurisdiction of the U. S. Federal Trade Commission. Indeed, years later, the FCC allowed the use of a subliminal in an announcement about a murder. Apparently, the briefly-flashed phrase "Give yourself up" was unsuccessful in convincing the murderer to turn himself in (McIver, 1988).

[4] I know; it's a bad pun, but I couldn't resist it.

ALF, for example, there is a picture of Garfield the cat with one of his normally bulging eyes missing to reveal a pink cavity, passed out in front of an open refrigerator full of wine bottles, one of which is spilling its contents. Taped to the refrigerator is a note reading "I ♡ ALF".

Of course, the existence of such subliminals proves only that people use them and believe them to have the intended effects (for the non-gag subliminals, anyway; the others are probably intended more as inside jokes), not that they actually are effective. George Miller, for example, who produced the Mad Max subliminal, undoubtedly believed that it would enhance the effectiveness of the scene, but that doesn't in any way prove that it does; the scene may have had exactly the same effect without the subliminal. As we will see, this fallacious confusion between the simple demonstration of the use of subliminals in some medium and the conclusion that they are therefore effective is common to the claims of many in this area. But first, we will attempt to define the term "subliminal".

29.3 The Meaning of "Subliminal"

The term "subliminal" is derived from the construct of a "limen of consciousness", a threshold or line separating conscious from unconscious. The concept dates back to the literal beginning of psychology as an empirical science separate from philosophy in the seminal writings of Johan Friedrich Herbart (1776-1841).[5] Herbart argued that ideas (i.e., both perceptions and thoughts) differed from one another in strength, and inhibit or suppress one another in a dynamical fashion, competing with one another to achieve enough strength to rise above the "limen of consciousness" and, hence, be consciously experienced. Ideas below the line still exist, in this view, and through collateral inhibition can influence what other ideas, including themselves, are subsequently consciously experienced.[6] In this way, perceptions obtained only subliminally (i.e., below the "limen of consciousness") can still affect what we experience consciously (i.e., think about) and how we behave.

It is this idea or something very similar to it that informs most popular discussion of subliminal persuasion or perception. Yet, except for Freudian psychology (which is not a credible scientific theory of perception or cognition), few models of perception and cognition take such notions seriously. First, contrary to the common caricature of psychology in the popular media, no modern theory posits "an unconscious", that is, a mind-entity separate from consciousness such that perceptions, ideas, beliefs and desires can slip or exert influence from one to the other. Rather, perceptual and cognitive processes can and often do occur without our awareness, and without our having to or, in many cases, even being able to consciously control them.

[5] Herbart argued strongly that psychology should be an empirical science—that is, based on experience, divorced from philosophy. Hence, he is fairly considered to be the father of the science of psychology. However, Herbart also argued that psychology could not be an experimental science; that aspect of psychology did not occur until Wilhelm Wundt in Europe and William James in the U. S. set up the first psychological laboratories in 1875.

[6] Herbart's notions of a dynamic unconscious clearly foreshadow the Freudian concepts of the dynamic subconscious, and related ideas such as repression, although Freud explicitly denied the connection.

Think of riding a bicycle, catching a ball, or even reading this text. Not only do they happen "automatically" (after you have become competent at them) and without our awareness of how we are doing them, but any attempt to gain conscious control of them usually results in the process coming to a grinding halt (i.e., and you falling off the bicycle). The important point is that these processes may occur *unconsciously*—that is, without all the internal chatter that normally accompanies what we refer to as "conscious" processes, but there is no reason to suggest that they therefore occur in some special mind-place called "the unconscious", any more than you would want to say that the internal workings of your toaster, car, or computer occurred in an "unconscious"; although, they certainly occur unconsciously, that is, without awareness—unless you are willing to believe that your computer, say, is sentient.

Second, few current perceptual or cognitive theories hold to the idea of an absolute sensory or information threshold dividing those events we are aware of from those we are not. Rather, it is viewed as a continuum along which the amount of information or sensation we require for a given event on a given trial before we are willing to say we are aware of it varies as a function of the context, the event itself, task demands, payoffs, and the like. This style of decision making about sensory and other kinds of events occurs because the event itself always occurs in the context of both internal and external "noise"—a background of half-formed, fleeting thoughts, impressions, sensations, desires, and beliefs of varying and unpredictable intensity against which we must detect the event in question. Sometimes, for example, this background "noise" is low enough that even relatively weak signals can be confidently detected; other times, it completely overwhelms even the most intense of signals so that a confident decision is impossible; unfortunately, we are rarely completely sure which state we are in, sometimes mistaking noise for signal and other times mistaking signal for noise, so we are constantly having to balance the costs of deciding that the event happened when in fact it didn't against the costs of failing to detect the event when it did. Thus, to say that you are "aware" of a given event is to say that *in your opinion* for the demands of the particular task in question you have gathered sufficient sensory or other evidence to conclude that the event did happen.

From this perspective, experimental psychologists can define two different "thresholds" of awareness. The *subjective threshold* refers to that level of information or stimulus intensity at which the individual claims to be "just guessing" or responding at a chance level—at which it just "feels" as if you have no information to decide one way or another about some event. Note that this threshold corresponds to a testimonial; events subliminal in this sense of being below the subjective threshold are *events whose presence observers do not report* (T. E. Moore, 1992).

In contrast, the *objective threshold*, refers to that lower level of information or stimulus intensity at which the observer is no longer able even by guessing to discriminate between events, say, the presence or absence of a signal, at a level above chance. Events subliminal in this sense are *events whose presence the observers can't detect* (T. E. Moore, 1992). The results of research using events below the objective threshold are clear: there is no compelling evidence for unconscious or subliminal perception of them, no evidence that the meaning or content of such events can affect people's behaviour (Holender, 1986; Greenwald, 1992). Put simply, if an event is below an individual's objective threshold, it has no effect on him or her at all.

The subjective threshold is different; here we can demonstrate an effect on the observer. Consider the results of what was one of the first psychological experiments performed in America. In this experiment, Pierce and Jastrow (1884) investigated their own abilities to discriminate tiny differences in pressure with their fingertips, and found that the accuracy of their decisions was still well above chance—that is, above what we would now call the "objective threshold"—even when they were convinced they were just guessing. In a similar experiment reported a few years later, Sidis (1898) found that subjects shown cards containing a single character at a distance between the cards and the subject at which the subjects saw nothing more than a blur or a spot, and therefore were convinced they were just guessing, were still able to name the characters at a level above chance. Numerous experiments of this type continue up to the present day, virtually all with the same result: by adjusting the intensity or duration of the target stimulus, people brought to the point of being convinced that they can no longer detect differences between events, say a video clip or an auditory tape with and without an embedded message, still obtain enough information—in their view "unconsciously"—to perform better than chance when forced to guess.

This, then, is subliminal perception in the sense of an apparent dissociation between awareness and the acquisition of information, but two points need to be noted about it. First, it is a long-established, unsurprising phenomenon in psychology, and no special devices or sophisticated processes are needed to produce it. Indeed, many of the techniques at issue here, such as Vicary's messages flashed for one-third of a millisecond or the messages on self-help auditory tapes, are well-below people's objective thresholds, and hence are not capable of producing this or any effect. Second, there is no convincing evidence that the effect when it occurs at all extends much beyond improving the accuracy of people's guesses in forced-choice tasks. That is, there is no good evidence that the meaning or content of such events can affect people's behaviour. And, given the extreme difficulty of ensuring that the stimuli used are simultaneously above the objective threshold but below the subjective threshold, many cognitive scientists seriously doubt whether there has been a convincing demonstration of the effect even on forced-choice guessing: just because individuals claim to be "just guessing" doesn't mean that they had no awareness of the event; maybe they were just being overly careful in attributing awareness, saying that they were guessing unless they were absolutely sure. In fact, doing so on only a few trials would be enough to establish the effect (Holender, 1986; Greenwald, 1992; T. E. Moore, 1992).

Thus, *if* it can be shown for some event that it is above observers' objective thresholds, *and if* it can be shown that it is simultaneously below their subjective thresholds (and we are willing to accept that being below the subjective threshold completely exhausts all possibilities for awareness), *and if* we can show some effect on the observers' behaviour consistent with the meaning of the event that doesn't also occur in the absence of the event, *then* we would have a demonstration of subliminal perception and subliminal influence.[7] None of the popular claims for subliminal influence come even

[7]Note that failing to be able to meet this criterion does not mean that unconscious perception doesn't happen, only that it doesn't happen without concomitant awareness. That is to say, it may be that there are unconscious effects, but that they are always associated with awareness; adjust conditions sufficient to eliminate the awareness, and you automatically eliminate unaware effects as well. If the unaware effects are never different from what would be expected from exercising the aware knowledge, it would be impossible

close to meeting this criterion. As noted, many fail because the events or messages are below the objective threshold, many to point of not existing at all! Others contain messages that are not obviously subliminal: observers are generally quite aware of them. And none of them present any scientific proof that the messages have the effects claimed for them. It is to those specific claims that we now turn.

29.4 Wilson Bryan Key and Subliminal Advertising

In four books and innumerable college and university lectures[8] over the last two decades, Wilson Bryan Key[9] has made a career out of the claim that advertisers have resorted to subliminal advertising to influence the buying public. He provides an extensive litany of such messages he claims to have found in print, film, and television as proof not only that advertisers use such techniques, but also that they must be effective, else why would so many different companies be using them? Key's primary claim is that advertisers use a variety of subliminal techniques to embed sexually-explicit words and symbols into the pictorial content of their advertisements and, indeed, the products themselves (such as Ritz crackers)[10] to manipulate the public by subliminally capitalizing on the public's obsession with sex. In Key's view, subliminal sexual imagery is much more persuasive than the more explicit use of sexual symbols, such as scantily-clad women that until recently seemed to populate beer and automobile commercials.

A paradox of Key's books and lectures is that they consist almost entirely of exposés of the subliminal content. The reader is regaled with photograph after photograph of advertisements highlighting the "subliminal" imagery that Key sees in them: male and female genitalia, phallic symbols, and numerous occurrences of the word "sex". But of course if Key and his audience can see these images, then they are hardly subliminal, subjectively or otherwise. In fact, what Key appears to mean by "subliminal perception" is more akin to the processes of selective attention, not perception without awareness (Creed, 1987).

But that's not the worst of Key's odd beliefs. He claims that the brain comprehends the totality of a complex stimulus such as print advertisements "at the speed of light", which is certainly more than a little faster than the neurons of my brain operate (and, I'd wager, Key's as well). According to Creed (1987), Key claimed in one of his lectures—in direct contradiction to basic neurophysiology and brain function—that the "unconscious brain" could perceive, analyse, and fully comprehend an advertisement in "less than a millisecond". No evidence is offered for any of these claims.

---

ever to demonstrate them. But if this were the case, it is not clear what it would mean to say that there were such effects.

[8] Creed (1987) reports that his U. S. university paid Key $2000 plus expenses for his public lecture. Ignoring expenses and assuming that this rate of pay is typical of the 40 or so lectures per year Key claims to give, then Key's income from lectures alone is over $80,000 per year.

[9] A not uncommon mistake is to reverse Key's first and second names; in fact, despite the careful eye of a professional editor, this error crept into one of our own publications discussing Key's work (Vokey & Read, 1985, p. 1232). We suspect a subliminal influence of the Beach Boys…

[10] When asked recently in court if he really believed that the word "sex" was baked, as he had claimed, many times into the surface of Ritz crackers, Key replied, "Yes, and on both sides too"!

Furthermore, no independent scientific evidence is offered either about the actual existence of the "subliminal" images that Key sees in the advertisements, nor their effectiveness. In fact, his examples appear to be little more than constructions and projections of his own fantasies. For example, one of Key's (1973) principal examples of subliminal manipulation is the phrase "u buy" printed backward in an advertisement depicting four types of rum. None of the dozens of my colleagues and students have been able to find this message when shown the ad, yet Key claims that 80% of the students in his studies must have unconsciously perceived the message, because that's the percentage that preferred the rum so labelled when asked to choose one of the four. In his view, the facts that the preferred rum comes in a larger bottle, is only one presented in a fancy brandy-snifter rather than a hi-ball glass, is of a darker colour than the other three, and is the only one that has the words "extra special" clearly written on the label—among other things—could not possibly account for the marked preference of his students. No, to Key, it is only the backward subliminal phrase "u-buy" that only he can see that can explain the preference.

In another example, Key (1973) asserts that the explanation for why 95% of college-age males in another of his studies were able one month later to remember a *Playboy* advertisement depicting a naked woman was that the ad also contained subliminal sexual imagery. That most young men would probably remember a picture of a beautiful, naked, young woman after a month even without subliminal sexual imagery seems to have escaped Key's notice, as has the use of control conditions, generally. In our own research on the issue (Vokey & Read, 1985), we could find no effect on memory of embedding the word "sex" in photographs compared, as controls, to the embedding of nonsense trigrams or no embedding in copies of the same photographs, despite the embedded words being perceptible (i.e., above the objective threshold) when pointed out to the subjects, none of whom was aware of the embedding (i.e., below the subjective threshold) until it was pointed out to them. There is in fact no scientific evidence to support Key's claims, and virtually the whole of cognitive science and neuroscience to contradict them.

John O'Toole, president of the American Association of Advertisers, asked in his review of Key's most recent book, "Why is there a market for yet another re-run of this troubled man's paranoid fantasies?" (T. E. Moore, 1992). There may not be. In a pretrial hearing of the Judas Priest case (discussed subsequently), after listening at length to Key failing to give a straight answer to questions put to him, and hearing him claim that "science is pretty much what you can get away [with] at any particular point in history and you can get away with a great deal" (T. E. Moore, 1992), the judge decided he had had enough and refused to allow Key into his court room during the actual trial to serve as a witness for the plaintiffs. Perhaps in a broader sense it's time the rest of the world did the same.

29.5 Subliminal Auditory Self-Help Tapes

The story with subliminal auditory self-help tapes is much the same as that with subliminal advertising, except that the techniques have shifted from the visual to the auditory domain, and instead of the crass motives of advertisers, we have the apparently

more noble objective of psychotherapy (Moore, 1992). One further difference is that instead of one person (Key) making money on books and the lecture circuit, we now have many companies making quite substantial profits[11] on the selling of audiotapes that to the naive listener (me, at any rate) appear to contain from 15 minutes to over an hour of rather monotonous, new age "music" performed typically on pan pipes, harps, and flutes, although some use orchestrated classical pieces. You also get sometimes extensive fold-out notes with the tapes, containing more psychobabble than a week of Oprah Winfrey, Ann Landers, and Dr. Joyce Brothers rolled together. Sometimes the documentation also includes a long list of references to "scientific" proof that subliminal audiotapes work, most of which are either irrelevant (such as studies on priming in lexical decision), nonsense (such as Key's books), or simply wrong. Some of the tapes even go so far as to warn you about the many fraudulent tapes out there that use subliminal techniques that don't work, unlike, it is claimed, the "scientifically proven" process used on the current tape.

What you don't get, however, is anything that will result in dramatic improvements in mental and physical well-being, or even improved bowling-scores, beyond placebo effects. In fact, you may not even be getting the promised subliminal auditory messages. Merikle (1988), for example, subjected a collection of supposedly subliminal audiotapes to a sensitive spectrographic analysis, and found nothing on the tapes beyond the music. Now that's subliminal! Obviously, any auditory stimulus that is too subliminal to register on auditory equipment more sensitive than the human ear is too subliminal to have any effect at all.

The business of subliminal auditory tapes got its start with Hal Becker, an engineer who began experimenting with visual subliminal techniques in the 1950s. He produced a device in 1978 to insert subliminal messages in music audiotapes. According to *Time* magazine in 1979, 50 department stores had begun using the device to insert messages such as "I am honest; I will not steal" many times at a low ("subliminal") level in the background music of the stores in an attempt to discourage shoplifting. *Time* claims that the stores reported a significant reduction in theft, although no evidence is offered to substantiate the claim (McIver, 1988).

As T. E. Moore (1992) notes, there are two rather fundamental problems with the rationale for the effectiveness of subliminal audiotapes. The first has to do with the nature of physical signal or message itself. With visual subliminals, the subsequent masking with other visual material does not change the target message, it simply limits the length of exposure to the target—a procedure experimental psychologists call backward masking. In contrast, with auditory subliminals, the target stimulus is first reduced substantially in volume, and then is overlaid with a masking stimulus (i.e., the music) of much greater physical energy. Given that it is physical energy that affects the basilar membrane of the ear, how is the listener supposed to separate the physically drowned-out subliminal signal from the masking stimulus? It is analogous to trying to detect by weight alone which of two kilogram bags of sugar at the supermarket used more glue in the packaging.

The second problem T. E. Moore (1992) notes is the complete lack of a theoretical rationale for why such messages should have the therapeutic effects attributed to them.

[11] Some of these tapes sell for as much as $400 per set (T. E. Moore, 1992).

Why should the repeated sub-audible presentation of the messages "I am a nice person" or "I will not eat" be effective in promoting either self-esteem or weight-loss, when the superliminal presentation of them would produce nothing but boredom? To the extent that an explanation is offered, it is usually to attribute to "the unconscious" with no evidence whatsoever precisely the processes necessary for the imputed effects to occur. As Greenwald (1992) argued recently, there is neither theoretical necessity nor empirical support for the psychoanalytic unconscious.

There is also no empirical support for the alleged therapeutic effects of these tapes. None of the nine reported studies on the efficacy of these tapes has shown an effect consistent with the manufacturers' claims (T. E. Moore, 1992; Pratkanis, 1992). One of the more interesting of these studies is that of Pratkanis, Eskenazi, and Greenwald (1990), because in addition to demonstrating no effect of the subliminal tapes, it suggests why so many users of them are willing to provide testimonials to their effectiveness.

In the experiment, volunteers interested in the potential of subliminal tapes (and therefore most like those who would actually buy them) were recruited to participate in a study of the effectiveness of the tapes to promote either self-esteem or improved memory. According to the manufacturer, the self-esteem tape contained messages such as "I have high self-worth and high self-esteem", and the memory tape messages such as "My ability to remember and recall is increasing daily". Volunteers first filled out measures of self-esteem and memory, and then proceeded to use either the esteem or the memory tape according to the manufacturer's instructions every day for 5 weeks. Although one-half of the subjects received their tapes with the correct instructions, the other half received the tapes with the instructions for the other tape: memory instructions for the self-esteem tape, and vice versa. After the 5 weeks, subjects again filled out the esteem and memory measures, and also indicated whether they believed the tape they had received to have been effective.

The results were that the tapes produced no effect whatsoever on either self-esteem or memory, regardless whether the correct or incorrect instructions had been received. However, the subjects *believed* the tapes to have been effective, indicating that they thought their self-esteem had improved if they thought they had received a self-esteem tape (regardless of whether or not they had), and that their memory had improved, if that was the tape they thought they had received (again, regardless of whether or not they had). Pratkanis et al. (1990) called this effect on belief the "illusory placebo effect", and it demonstrates one reason why testimonial or anecdotal evidence for the efficacy of some product or treatment (e.g., "it worked for me!") is almost always worthless.

29.6 Subliminal Messages in Rock Music and Suicide Trials

In the Fall of 1982, a fundamentalist preacher from California, Gary Greenwald, arrived in Lethbridge to hold two days of public lectures on the evils of what he referred to as "backward masking" in rock music. These capacity-crowd lectures were followed by a mass rally and record-smashing spree. Similar performances by both Greenwald

and many others have occurred throughout Canada and the United States, numerous books of the same theme have been published, various state governments have considered legislation to control "backward masking" in rock music, or at least require warning stickers on the recordings, and at least one Canadian member of parliament and the Consumers Association of Quebec have supported government investigation of the issue (McIver, 1988; Vokey & Read, 1985).

Greenwald's use of the term "backward masking" is not what experimental psychologists mean by the term, mentioned earlier. Rather, it refers to the masking of the meaning of an auditory message by playing it backward. Unlike other claims of subliminal persuasion, the stimulus in auditory backward masking is usually quite audible, often as the forward or normal lyrics of the song. Greenwald's claim is that when these songs are played backward, intelligible messages can be heard that are not apparent when heard in the forward or normal direction. However, they are, he claims, perceived unconsciously by listeners as they hear the recordings in the forward direction. Greenwald and his colleagues further claim that these backward messages inevitably are evil in content and, because they are perceived unconsciously, cannot be resisted as they lead the unsuspecting youthful listener down a path of licentious sex and drugs and who knows what all.

As with Key and his many "examples" of subliminal advertising, to make their case, Greenwald and the others provide an ever-growing list of evil, "satanic", backward messages that they have found in rock music. You can order audiotapes from Greenwald in which he demonstrates many of the messages he has found in the recordings of such rock groups as Queen, Jefferson Starship, Led Zeppelin, The Beatles, and Pink Floyd, to name just a few. The initial question, of course, is whether these messages are really there; there is no question that Greenwald and his audiences believe they hear the messages that Greenwald plays for them, but are they actually there or are they, as with Key's examples, constructions and projected fantasies of the listener?

Reversed speech retains many of its speech-like qualities; in fact, in our own work on the issue (Vokey & Read, 1985), we opined that to us it sounds a lot like the Swedish Chef from the old *Muppets* television show.[12] Consequently, any speech played backward will occasionally produce something akin to a word or a phrase, particularly if the listener is told how to interpret the gibberish, as Greenwald does with his listeners. For example, playing the phrase "Jesus loves you" backward will sound something like "we smell sausage", particularly now that you've been told what to listen for, but it is strictly coincidence. As generations of cloud-watchers will attest, you can "detect" meaningful patterns in just about anything, particularly if you are told what to look for.

Poundstone (1983, 1986, 1993) has analysed many of these alleged backward messages and has found that most of them are merely coincidental reversals. Some, however, particularly in recordings made since the claims for such backward messages became a major media event in the early 1980s, clearly are engineered reversals—that is, reversals intentionally inserted in the recording—and clearly intended as jokes. Unlike the coincidental reversals, these engineered reversals are clear and unambiguous when played backward. One of my personal favourites is in the song "Goodbye Blue

[12] A claim, incidentally, that was misconstrued by one *Globe and Mail* reporter to be that we had found that all rock music was actually Swedish when played backward!

Sky" on Pink Floyd's, *The Wall*. Just as the song ends and before the next song begins (which, not so incidentally, is identified as "Empty Spaces" on the record label, but as "What shall we do now?" on the album cover), is a somewhat muted passage that sounds something like speech, but you could play the album through many times without noticing it. However, played backward a voice very clearly says: "Congratulations, you have just discovered the secret message. Please send your answer to old Pink, care of the funny farm..." (Poundstone, 1983).

The more important question is whether the meaning of backward messages could affect listeners. The only evidence that Greenwald and other fundamentalists offer is first an equation of backward messages with studies of subliminal advertising such as Key's that, in their view, have demonstrated powerful effects. Second, they assert that it is the young people who preferentially listen to rock music who have the greater incidence of sexual promiscuity, drug abuse, and other criminal behaviour. Aside from the fact that both sets of behaviours are a function of youth, citing them as evidence of the effectiveness of backward messages in rock music is circular reasoning, because backward masking was advanced as the explanation for these behaviours in the first place. If that isn't clear, consider the following: "A and B are found together because A causes B." How do you know A causes B? "Easy, because A and B are found together." In fact, there is no evidence that the meaning of messages heard backward has any affect on people. In the only published research on this question, in a series of experiments, we could find no effect of the meaning of engineered, backward messages on listeners' behaviour, either consciously or unconsciously (Vokey & Read, 1985).

The seriousness with which people take these claims is apparent by the fact that they have been the basis of at least two wrongful death lawsuits filed against recording companies and artists by the parents of teens who committed suicide, allegedly as a consequence of listening to the rock music recordings.[13] In the first of these, a father filed suit against heavy-metal rocker Ozzy Osbourne and his record company after the man's 19-year-old son committed suicide in 1984 after a night of heavy drinking and listening to Osbourne's *Blizzard of Ozz* album. The father contended that the content of the album song "Suicide Solution", particularly a 27-second, *forward* subliminal message, had driven his son to suicide. However, as the subsequent legal proceedings made clear, the song is actually intended to be anti-suicide. Furthermore, the alleged subliminal message is in fact quite perceptible (or at least as perceptible as any of the lyrics are in heavy metal music), not subliminal at all (Poundstone, 1993). The California District Court of Appeals dismissed the suit in 1988, citing the "free speech" amendment of the American constitution.

The second case involved two teenage boys who, two days before Christmas in 1985, spent the afternoon drinking beer, smoking marijuana and listening to Judas Priest's *Stained Class* album. They then took a shotgun into a playground of a nearby church and shot themselves, one of them dying instantly, and the other surviving with half his face blown off. The boys' families filed a multi-million dollar lawsuit not, as Poundstone (1993) notes, against the beer company, drug dealers, or the gun shop from which the weapon had come, but against Judas Priest and CBS records. The survivor

[13] In addition to the two cases discussed here, similar suits have been filed in at least five other states in the U. S. (Poundstone, 1993).

initially offered to the police the explanation for the suicide attempt that “life sucks”. Later, he cited the Judas Priest song “Beyond the Realms of Death”, claiming to have been “mesmerized” by the suicidal themes of the lyrics. Possibly because of the earlier court ruling in the Osbourne case, by the time the suit came to trial in Reno, Nevada in 1990, the plaintiffs argued instead that the cause of the boys’ behaviour was a series of backward subliminal messages on the album, and a forward subliminal “Do it” in the cut “Better by You, Better than Me”.

It was because of the alleged backward messages that Don Read and I were asked to serve as expert witnesses and to assist the defence in this case. Our evidence was presented to the court much as outlined here. However, it soon became clear that in the opinion of the plaintiffs and the judge, the crux of the case concerned the existence and likelihood of effect of the forward subliminal, “Do It”. Unlike the earlier Osbourne ruling (which concerned consciously perceptible messages), the judge in this case ruled that subliminal messages are not protected by the First Amendment of the U. S. constitution, principally because, being subliminal, it is not possible for persons to be aware and therefore prevent themselves from unconsciously “hearing” and thereby rejecting them, as they can with aware speech. Despite that, however, the judge found in favour of the defendants—not because in his view the whole idea of subliminal messages inducing people to suicide is nonsense, but because there was not sufficient evidence to prove that the subliminal message “Do It” had been intentionally placed there by Judas Priest (Vance and Belknap v. Judas Priest and CBS Records, 1990).

29.7 Conclusion

Despite apparent use of subliminal techniques in different media, there is as we have seen simply no evidence for effective subliminal persuasion in film or video, advertising, self-help audiotapes, or rock music, and there is certainly no theoretical basis to expect it. Perhaps the next time somebody tells you in whispered tones about the infamous “Popcorn and Coke” study as support for how some subliminal audiotape taught “this man” calculus while he was sleeping, you can set them straight.

Chapter 30

Implicit Learning
John R. Vokey

Is it possible to learn something without being aware that you've learned it? Can you gain knowledge about structural relationships, concepts, categories, and contingencies without knowing you've learned it, even when you are able to use it successfully in some task? Some scientists have suggested not only that it is possible, but that such "learning without awareness" or "implicit learning" is the usual way we acquire, retain, and use much of our knowledge about the world around us. They argue that it is the system of learning that underlies "intuition", and that it functions independently of the conscious or explicit learning system.

30.1 A Demonstration

Try the following simple demonstration on one of your friends. While your friend is telling you a long story, respond as you usually do in such social situations, for example, nodding your head and saying "uh-huh" or "umhmm" to express interest, except do so for all and only the plural nouns in your friend's story. If you keep track of the number of plural nouns your friend uses, you should find that the rate with which your friend uses them increases, sometimes quite substantially. It can be even more effective if a group of people, such as students in a classroom, reinforce some behaviour of the story-teller, say, an instructor.[1] Yet, if you then ask your friend (or instructor) why he or she was using so many plural nouns, he or she will often have no idea, dismissing your claim with such perfunctory statements as "That's just the way I talk" or "The story (or lecture) just required that many of them".

[1] Don't try this demonstration on psychology professors; they are a dour, humourless bunch, who, being omniscient, will go out of their way to thwart you. Do feel free, however, to try it on professors in your other courses.

30.2 Implicit Operant Conditioning

This simple demonstration was the core of a series of experiments in the 1950s on learning without awareness (e.g., Adams, 1957; Eriksen, 1960), although research on the issue began much earlier (e.g., Thorndike & Rock, 1934). The intuition informing much of this work is the simple observation that non-human animals are capable of the same form of what is called "instrumental learning" or "operant conditioning". Operant conditioning is what you use to train your dog, for example. It refers to the fact that behaviours (such as the production of plural nouns) followed by rewards or reinforcers (your "uh-huhs" in the demonstration) increase in frequency, while unreinforced behaviours drop off. Although capable of instrumental learning, the belief is that animals obviously aren't doing it consciously; thus, it is not implausible that humans could also evince similar learning without awareness. If you have trouble buying the argument, you are probably thinking of some anthropomorphic, high-level mammal such as dog, cat, or dolphin. Adjust the level of animal downward until you reach one that you are sure is not capable of consciousness. For me, worms and insects do the trick. Rest assured that these animals are also capable of learning.

30.2.1 Two problems

Despite its plausibility and apparent empirical support, the idea has always had its critics. They point to two major problems: (1) the sensitivity of the measure of awareness, and (2) the difficulty of establishing just what has been learned. For example, in the demonstration case the conclusion is that the person has learned a specific contingency (that plural nouns lead to reward), and that he or she has done so without awareness of that contingency. But most studies of this type assess awareness only well after the learning has been established. What if the person had been aware of the contingency while learning, but has simply forgotten it by the time he or she is asked about it? Think of the myriad skills you have learned, such as driving a car, operating your computer, etc., all involving many specific contingencies. Many or all of these contingencies were learned explicitly (and sometimes painfully), but have since become quite automatic, often to the point that we are no longer able to describe explicitly what we have learned. But, although they are now good examples of implicit memory or implicit retrieval, they aren't examples of learning without awareness because we were aware of them as we learned them. Along the same line, perhaps more probing or more precisely focused questions would reveal that the person was still aware of the contingency in some form or another.

Even more troublesome is the possibility that the person was explicitly learning about some contingency, just not the one nominally being taught.[2] If this contingency were correlated with the "correct" one, then it would appear that the person had learned the "correct" contingency without awareness when in fact what he or she had learned was something else of which they were fully aware. In an elegant experiment, Dulaney

[2]There is also the possibility that the person was not learning any contingency at all, at least not in the usual sense of some strengthening of an abstract association, but rather some other form of knowledge that produces similar behavioural effects. This possibility is considered in the subsequent discussion on the implicit learning of complex structure.

(1961) demonstrated that this was precisely what was going on in the instrumental learning studies of implicit learning. He showed that his subjects were consciously learning such heuristics or rules as "produce an item from the same semantic category as the item just reinforced". Although this rule is incorrect, it is correlated with the correct one of producing plural nouns.

For example, if the person had just been rewarded for saying "diamonds" then staying with the same semantic category would make it likely that he or she would produce another plural noun ("rubies"), which would be rewarded, than if the person changed categories, even though the rule is not correct (Shanks & St. John, 1994). Imagine the person's surprise when he or she is then informed about the correct contingency. From the learner's perspective, it would appear that he or she talked about jewels for a while, then jewellery, then fashion accessories, then clothes, etc., but nothing that would appear to require the use of plural nouns. And at any rate, the person certainly was not thinking about plural nouns while doing it. As the rule the person was using appears on the surface to bear no relationship to the correct one, the only conclusion the person can come to in order to account for the apparent learning of the correct rule is that it must have occurred unconsciously. Numerous studies of this type have been done since and none has shown unequivocal evidence for implicit learning, leading Brewer (1974) in a review of the literature to conclude that operant learning in adult humans does not occur without awareness.

30.3 Implicit Classical Conditioning

To circumvent the problem of correlated rules, the experimenter has to gain complete control of the contingency relationship so that none other but the "correct" contingency exists, something that is not possible with instrumental learning tasks. One way to achieve this goal is to use a form of training known as classical or Pavlovian conditioning (after Ivan Pavlov, who identified it). This form of training has the further advantage that it is considered by some to be even more primitive than instrumental learning, so it is clear that it doesn't require awareness to be learned, in that even the most primitive of animals seem capable of it. The question, then, is can it occur without awareness in adult humans. In such implicit classical conditioning experiments, one stimulus, say a red light, is paired with electric shock, and another stimulus, say a blue light, is paired with no shock, and the experimenter records the galvanic skin response (GSR), a measurable change in the electrical potential of the skin that occurs in anticipation of pain or other noxious events.[3] Initially, of course, both lights produce exactly the same GSR, but after a few pairings, the GSR to the red light is markedly different from that to the blue, which remains essentially unchanged, indicating that the subject has learned the association of the red light with shock.

One obvious problem is that very few people would fail immediately to pick up explicitly on the contingency, even if their being aware of it wasn't the source of the conditioned GSR. Consequently, experimenters have come up with all sorts of creative ways of hiding the correct contingency in a flurry of other events that ultimately are

[3]Other unconditioned stimuli and responses have been used, but shock and the GSR have been the most common; the pattern of results is the same, regardless.

not correlated with the correct contingency. However, as long as the experimenters are careful to use appropriately sensitive measures of awareness, then no matter how creative the experimenter has been, the results are always the same: if subjects show conditioning of the GSR, then they also show awareness of the correct contingency; and subjects who show no awareness of the contingency, also fail to show conditioning of the GSR. Hence, as with instrumental learning, Brewer (1974) concluded that classical conditioning in adult humans also does not occur without awareness. No experiment has occurred since to change this conclusion (Shanks & St. John, 1994).

30.4 Implicit Learning of Complex Structure

It's possible that these simple learning tasks are just that: too simple; possibly basing our intuitions for what should be a likely place to find implicit learning in humans on what non-human animals can do unconsciously is misleading. Perhaps a more complex learning problem, something uniquely human, would be a more promising candidate situation. Arthur Reber has been a major proponent of this view; in fact, he is credited with naming it "implicit learning" in an article in 1967. He describes implicit learning as a process by which knowledge of the structure of a complex stimulus environment is acquired largely independently of conscious awareness of specific components of that environment (Reber, 1989). He argues that the implicit learning system is used not for tasks and situations that are easy to handle explicitly, such as the learning of simple contingencies, but for those highly complex structures that are difficult or impossible to learn explicitly in a reasonable amount of time, particularly given the often impoverished and ambiguous contexts in which we encounter them.

The quintessential human trait is language, by which is meant (by linguists, anyway) a formal, syntactical structure, or generative grammar. We don't encounter the grammar of our language in a simple form, such as a list of rules or explicit diagrams ("grammar school" training, notwithstanding) that would facilitate the explicit learning of it.[4] Rather, we encounter the grammar of our language only by way of exemplars of its use in the speech around us. These are hardly optimum or even adequate conditions for the explicit learning of the rules; imagine attempting to learn the far simpler rules of baseball or even cribbage from simply watching them being played: I doubt you'd ever pick up on the "infield-fly rule" or the point gain for "his nobs", for example, or be able to discriminate the formal rules of the game from the heuristics and strategies used by the players. In fact, even professional linguists have elucidated probably only a small part of the broader structure underlying, say, spoken English from studying examples of its use—and that's taken many lifetimes of some of our most brilliant intellectuals to do. Yet, most children master this incredibly complex structure in just a few short years with little or no explicit training, and most of us can speak grammatically and correctly recognize grammatical speech, effortlessly and with little or no explicit thought. Here then would appear to be a clear example of implicit learning: not only are most of us currently unaware of the grammatical rules we use to communicate, we also apparently weren't aware as we learned them as young children.

[4] Furthermore, if you recall your grammar school days or have ever taken a course in linguistics, you will remember how difficult it was to learn and use the few rules you were taught explicitly.

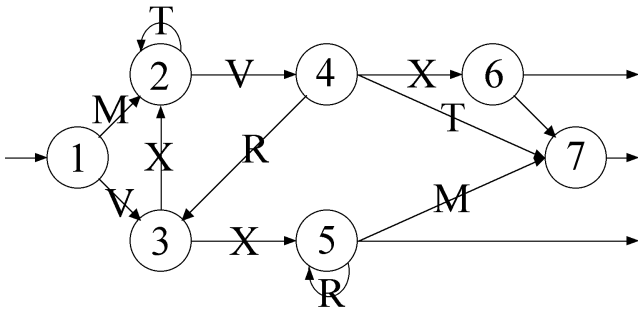


Figure 30.1: An example of an artificial grammar (from Vokey & Brooks, 1992).

Using the example of language learning as a guide, Reber constructed much simpler artificial grammars to study implicit learning. An example of one is shown in Figure 30.1. Grammatical items are formed by following the arrows between the numbered nodes, picking up the letters in passing. For example, the item "MVRXM" is grammatical according to this grammar, whereas the item "MVXRT" is not because it cannot be formed in this way. A sample of grammatical items from this grammar is shown in Table 30.1.

These artificial grammars are inordinately many times simpler than the grammars of natural languages, or even that of the orthography or "spelling rules" of English, which they superficially resemble. These "finite state grammars", as they are known, are not just simpler than natural language grammars, they are fundamentally different; no linguist currently seriously considers them to be even representative of the grammars underlying natural language, although they were considered promising candidates for study in the early 1950s. Despite their simplicity, these grammars retain many of the learning characteristics of their natural analogues. For example, even a grammar as simple as the one in Figure 30.1 requires at least an afternoon of study to learn when studied in its formal, diagrammatic form (Reber, 1989), and asking subjects to attempt explicitly to deduce the rules from a sample of exemplars (which is all one ever gets with natural grammars), such as the eight items shown in Table 30.1, is virtually impossible (Brooks, 1978; Reber & Allen, 1978).

However, we aren't able to say much explicitly about our natural grammar, either; but we are able to use our implicit knowledge of it to guide our behaviour. One grammar task that most adults and linguistically-competent children find simple to do is to discriminate grammatical productions from ungrammatical ones, even when the examples make no sense. In fact, making sense appears to be completely irrelevant in making these judgments, exactly as a linguist would predict from an independent knowledge-base of pure syntax. For example, Noam Chomsky's famous phrase "Colourless green ideas sleep furiously" clearly makes no sense, but is just as clearly grammatically correct, whereas the phrase, "I have a jacket with green two sleeves", is understandable, despite being ungrammatical. Yet, if you ask people how they know that the first phrase is grammatical and the second not, they have extreme difficulty explicitly citing the appropriate rules, despite having little difficulty apparently in using them. In fact, what

VXM	MVRXR
VXTVX	MTTTTVT
MTTVRXM	VXRRRRR
VXTTTVT	MVRXTVX

Table 30.1: Items generated from the artificial grammar shown in Figure 30.1.

most people say is that the judgment is intuitive, the grammatical items just look or "feel" right.

Can we get a similar dissociation between performance and explicit knowledge of the rules with artificial grammars? Remarkably, the answer is "yes", and with surprisingly little training. As a simple demonstration, ignore the diagram of the grammar, and study the items in Table 30.1 for a short while until you are satisfied that you could remember them if you were to be given a later memory test.

Study Them Now!

Now, look at the list of items shown in Table 30.2. Some of these items, as with the eight study items, were generated by the grammar; the remainder were constructed to violate the grammar in one way or another. Without looking at the study items or the grammar, go through the list, placing a check-mark next to the items you "feel" are probably grammatical—that is, probably follow the rules of the grammar in Figure 30.1, leaving the other, ungrammatical items blank.

Try it now!

In formal experiments similar to what you've been asked to do, participants are first asked to study in some fashion or another a sample of grammatical items, and then are told for the first time about the existence of the grammar, but not the rules themselves, and then are asked to sort a list of new items for grammaticality. If you are like the many subjects who have been given this task in numerous experiments by Reber and others, you should have found it relatively effortless, even if you weren't completely sure, as they often are not, about any of your choices.

Don't Read Any Further until You've Tried the Grammar Task!

Subjects given this task, although usually with many more test items, typically have accuracy scores anywhere between 60 and 80 percent correct, depending on the

MVRXTVT	VXVRXRX	VXTTTVX	VXTVRXT
VXRRRRM	MTVRXVR	MTTVRXR	MTVRXRX
MTTTTVX	MVRXRRT	VXTVT	VXRRT
MVRXM	MTTVM	MVR	MVRXTVR

Table 30.2: Some of the sample items shown here were generated by the grammar shown in Figure 30.1; the remainder were constructed to violate the grammar.

precise instructions and training tasks given them. To determine your percent accuracy score in discriminating grammatical from ungrammatical items, add the number of odd-numbered items (i.e., items in columns 1 and 3 of Table 30.2) that you checked to the number of even-numbered items (i.e., items in columns 2 and 4 of Table 30.2) you left blank, divide by 16, and multiply by 100. How did you do?[5] If you are like most subjects, you will have no idea how you did as well as you did; that is, your explicit knowledge of the artificial grammar will lag well-behind your ability to sort the items with respect to grammaticality, just as it does for the natural grammar of your native language. Yet, unlike your many years of experience with your native grammar, you probably spent no more than a few minutes learning this artificial one.

Reber (1989) has taken evidence such as this as strong support for his contention that not only do people learn (at least some of) the rules of the artificial grammar, but that they learn it implicitly, by engaging a learning system separate from the explicit learning system. Reinforcing this conclusion is the finding that asking the subjects to search explicitly for the rules rather than simply study the initial list of items results in few if any useful, explicit rules and performance little better than chance when they are asked to sort the items for grammaticality (Brooks, 1978; Reber & Allen, 1978).

As with the claims for the implicit learning of simple associations discussed earlier, Reber's claims have not been without their critics; in fact, the last few years has seen a plethora of such challenges, although there have been criticisms of his conclusions all along (e.g., Brooks, 1978). These criticisms are similar to those levelled at the other work discussed previously, particularly the possibility that what the subjects are actually learning in the grammar tasks is correlated with but otherwise not the same as the abstract rules of the artificial grammar.

Brooks (1978), for example, argued that what subjects are using in the learning of artificial grammars is the similarity between test items and specific, remembered study items. Because items that obey the rules will tend to look more alike one another than will ungrammatical items, this simple strategy will tend to sort the items correctly. If you compare each of the items you selected in Table 30.2 to the study items in Table 30.1, you will probably find that you preferred those items that were highly similar to one or more of the study items, and rejected those items that had no close matches. My own work on these tasks has served to extend and refine this simple idea (Brooks & Vokey, 1991; Vokey & Brooks, 1992; Vokey & Brooks, 1994).

A related idea is that what the subjects are doing is accumulating not whole instances as in Brooks' and my theorizing, but simply small parts of the items, such as bigrams and trigrams. Test items containing recognizable parts of the study items are then judged to be grammatical by subjects, whereas items containing unfamiliar parts are rejected. Again, this strategy will tend to sort the items correctly even though the subjects have learned nothing directly about the underlying grammar (e.g., Dulaney, Carlson, & Dewey, 1984).

[5] Don't be concerned if your score was at or even below chance (50%). With so few test items, it would be easy to get most or all of them wrong simply by chance alone.

30.5 Conclusion

The possibility that subjects in implicit learning experiments are learning and responding to something other than the nominal basis of the task as set by the experimenter has plagued the interpretation of the results of these experiments. There is no question that our performance in many tasks is influenced by knowledge of which we are unaware (Higham & Vokey, 2000; Higham, Vokey, & Pritchard, 2000; Vokey & Higham, 1999). In a recent review of the literature on implicit learning, however, Shanks and St. John (1994) concluded that despite 60 years of attempted demonstrations, there is still no unequivocal evidence that the acquisition of such knowledge occurs implicitly.

References

Abel, E. (1980). *Marihuana: The first twelve thousand years.* New York: Plenum Press.

Adams, J. K. (1957). Laboratory studies of behavior without awareness. *Psychological Bulletin*, *54*, 383–405.

Ainsworth, M., Blehar, M., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Hillsdale, NJ: Erlbaum.

Allen, S. W., Norman, G. R., & Brooks, L. R. (1992). Experimental studies of learning dermatologic diagnosis: The impact of examples. *Teaching and Learning in Medicine*, *4*, 35–44.

Alzheimer, A. (1907/1987). About a peculiar disease of the cerebral cortex (translated by L. Jarvick and H. Greenson). *Alzheimer's disease and associated disorders*, *1*, 7–8.

Arendt, J. (1998). Biological rhythms: the science of chronobiology. *Journal of the Royal College of Physicians London*, *32*, 27–35.

Arnold, W., & Trillmich, F. (1985). Time budget in galapagos fur seal pups: The influence of the mother's presence and absence on pup activity and play. *Behaviour*, *92*, 302–321.

Bagemihl, B. (1999). *Biological exuberance: Animal homosexuality and natural diversity*. New York: St. Martin's Press.

Bagshaw, M., & Benzies, S. (1968). Multiple measures of the orienting reaction and their dissociation after amygdalectomy in monkeys. *Experimental Neurology*, *20*, 175–187.

Barber, N. (1991). Play and energy regulation in mammals. *Quarterly Review of Biology*, *66*, 129–147.

Bartholomew, K. (1990). Avoidance of intimacy: An attachment perspective. *Journal of Social and Personal Relationships*, *7*, 147-178.

Bartlett, N., Vasey, P., & Bukowski, W. (2000). Is gender identity disorder in children a mental disorder? *Sex Roles*, *43*, 1-33.

Bartus, R. T., Dean, R. L., Beer, B., & Lippa, A. S. (1982). The cholinergic hypothesis of geriatic memory dysfunction. *Science*, *217*, 408–417.

Batshaw, M. (1999). *Children with disabilities.* (Fourth ed.). New York: Paul H. Brooks Co.

Bem, S. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clincial Psychology*, *42*, 155-162.

Benoit, D., & Parker, K. (1994). Stability and transmission of attachment across three generations. *Child Development*, *67*, 1816-1835.

Berndt, T. (1997). *Child development*. (Second ed.). Dubuque: Brown and Benchmark.

Biklen, D. (1990). Communication unbound; autism and praxis. *Harvard Educational Review*, *3*, 291–314.

Biklen, D. (1992). Typing to talk: Facilitated communication. *American Journal of Speech and Language Pathology*, *1*, 15–17.

Biklen, D., Morton, M. W., Gold, D., Berrigan, C., & Swaminathan, S. (1992). Facilitated communication: Implications for individuals with autism. *Topics in Language Disorders*, *12*, 1–28.

Biklen, D., Morton, M. W., Saha, S. N., Duncan, J., Gold, D., Hardardottir, M., et al. (1991). "I AMN NOT A UTISTIVC ON THJE TYP" ("I'm not autistic on the typewriter"). *Disability, Handicap, & Society*, *6*, 161–180.

Bolin, A. (1996). Transcending and transgendering: Male-to-female transsexuals, dichotomy and diversity. In G. Herdt (Ed.), *Third sex, third gender: Beyond sexual dimorphism in culture and history* (p. 447-486). New York: Zone Books.

Bouchard, T. J. (1994). Genes, environment, and personality. *Science*, *264*, 1700-1701.

Bouchard, T. J., & McGue, M. (1981). Familial studies of intelligence: A review. *Science*, *212*, 1055-1059.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment*. New York: Basic Books.

Bracken, M. (2002). Steroids for acute spinal cord injury. *Cochrane Database of Systematic Reviews*, CD001046.

Brecher, E. M. (1972). *Licit and illicit drugs*. Boston: Little, Brown.

Brennan, J. F. (1986). *History and systems of psychology* (Second ed.). New Jersey: Prentice-Hall.

Brewer, W. F. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes.* New York: Erlbaum.

Brodhun, M., Bauer, R., & Patt, S. (2004). Potential stem cell therapy and application in neurotrauma. *Experimental and Toxicologic Pathology*, *56*, 103-112.

Brooks, L. R. (1978). Non-analytic concept formation and memory for instances. In E. Rosch & B. Lloyd (Eds.), *Cognition and concepts* (pp. 169–211). Hillsdale, NJ: Erlbaum.

Brooks, L. R. (1987). Decentralized control of categorization: The role of prior processing episodes. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization.* Cambridge: Cambridge University Press.

Brooks, L. R., Norman, G. R., & Allen, S. W. (1991). Role of specific similarity in a medical diagnostic task. *Journal of Experimental Psychology: General*, *120*, 278–287.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General*, *120*, 316–323.

Bruner, J. S., Goodnow, J. J., & Austin, G. (1956). *A study of thinking*. New York: John Wiley.

Bruton, C. J. (1988). *The neuropathology of temporal lobe epilepsy*. New York: Oxford University Press.

Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review*, *100*, 204–232.

Cador, M., Robbins, T. W., & Everitt, B. J. (1989). Involvement of the amygdala in stimulus-reward associations: Interaction with the ventral striatum. *Neuroscience*, *30*, 77-86.

Calculator, S. N., & Singer, K. M. (1992). Letter to the editor: Preliminary validation of facilitated communication. *Topics in Language Disorders*, *12*, ix–xvi.

Cannon, W. B. (1932). *The wisdom of the body*. New York: W. W. Norton.

Carskadon, M. A., & Dement, W. C. (1994). Normal human sleep: An overview. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), *Principles and practice of sleep medicine* (pp. 16–25). Toronto: Saunders.

Cartmill, M. (1990). Human uniqueness and theoretical content in paleoanthropology. *International Journal of Primatology*, *11*, 173–192.

Cheney, D., & Seyfarth, R. (1990). *How monkeys see the world*. Chicago: University of Chicago Press.

Cheney, D., Seyfarth, R., & Silk, J. (1995). The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: Evidence for causal reasoning. *Journal of Comparative Psychology*, *109*.

Choi, J., & Silverman, I. (1996). Sexual dimorphism in spatial behaviours: Applications to route-learning. *Evolution & Cognition*, *2*, 165-171.

Clark, C. M., Trojanowski, J. Q., & Lee, V. M. Y. (1997). Neurofibrillary tangles in Alzheimer's disease: Clinical and pathological implications. In J. D. Brioni & M. W. Decker (Eds.), *Pharmacological treatment of alzheimer's disease: Molecular and neurobiological foundations* (pp. 217–238). Wiley-Liss.

Cohen, D. B. (1979). *Sleep and dreaming: origins, nature, and function*. New York: Pergamon Press.

Corsellis, J. A. N., Bruton, C. J., & Freeman-Browne, D. (1983). Alzheimer's disease: a disorder of cortical cholinergic innervation. *Science*, *219*, 1184–1189.

Cousins, N. (1957). Smudging the subconscious. *Saturday Review*, *October 5*, 20.

Coyle, J. T., Price, D. L., & DeLong, M. R. (1983). Alzheimer's disease: a disorder of cortical cholinergic innervation. *Science*, *219*, 1184–1189.

Creed, T. T. (1987). Subliminal deception: Pseudoscience of the college lecture circuit. *Skeptical Inquirer*, *11*, 358–366.

Crossley, R. (1992). Getting the words out: Case studies in facilitated communication training. *Topics in Language Disorders*, *12*, 46–59.

Crossley, R., & Remington-Gurney, J. (1992). Getting the words out: Facilitated communication training. *Topics in Language Disorders*, *12*, 29–45.

Cummins, R. A., & Prior, M. P. (1992). Autism and assisted communication: A response to Biklen. *Harvard Educational Review*, *62*, 228–241.

Darwin, C. (1859). *The origin of species*. New York: The Modern Library.

Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London: John Murray.

Dasser, V. (1988). A social concept in Java monkeys. *Animal Behaviour*, *36*, 225–230.

Datlow Smith, M., & Belcher, R. G. (1993). Brief report: Facilitated communication

with adults with autism. *Journal of Autism and Developmental Disorders*, *23*, 175–183.

Davis, K. L., Davis, B. M., Greenwald, B. S., Mohs, R. C., Mathe, A. A., Johns, C. H., et al. (1986). Cortisol and alzheimer's disease. I: Basal studies. *American Journal of Psychiatry*, *143*, 300.

de la Torre, J. C. (1994). Impaired brain microcirculation may trigger Alzheimer's disease. *Neuroscience and Biobehavioral Reviews*, *18*, 397–401.

Deitrich, R., Dunwiddie, T., Harris, R., & Erwin, V. G. (1989). Mechanism of action of ethanol: initial central nervous system actions. *Pharmacological Reviews*, *41*, 489–537.

Deleon, M. J., Golomb, J., George, A. E., Convit, A., Tarshish, C. Y., McRae, T., et al. (1993). The radiologic prediction of Alzheimer's disease: The atrophic hippocampal formation. *American Journal of Neuroradiology*, *14*, 897.

Dement, W., & Kleitman, N. (1957). Cyclic variations in eeg during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalography and Clinical Neurophysiology*, *9*, 673–90.

Dement, W. C. (1978). *Some must watch while some must sleep*. New York: W. W. Norton.

Demonet, J. F., Thierry, G., & Cardebat, D. (2005). Renewal of the neurophysiology of language: functional neuroimaging. *Physiological Reviews*, *85*, 49–95.

Descartes, R. (1637). *Discourse on method*.

Descartes, R. (1641). *Meditations*.

Devore, H. (1989). *Gender blending: Confronting the limits of duality*. Bloomington: Indiana University Press.

Devore, H. (1998). *FTM: Female-to-male transexuals in society*. Bloomington: Indiana University Press.

Dietz, V., & Colombo, G. (2004). Recovery from spinal cord injury: Underlying mechanisms and efficacy of rehabilitation. *Acta Neurochirurgica Supplementum*, *89*, 95-100.

Donnellan, A. M., Sabin, L. A., & Majure, L. A. (1992). Facilitated communication: Beyond the quandary to the questions. *Topics in Language Disorders*, *12*, 69–82.

Dreger, A. (1998a). *Hermaphrodites and the medical invention of sex*. Cambridge, MA: Harvard University Press.

Dreger, A. (1998b). Ambiguous sex—or ambivalent medicine? Ethical issues in the treatment of intersexuality. *Hastings Center Report*, *May-June*, 24–35.

Dulaney, D. E. (1961). Hypotheses and habits in verbal "operant conditioning". *Journal of Abnormal and Social Psychology*, *63*, 251–563.

Dulaney, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology : General*, *113*, 541–555.

Dunn, L., & Dunn, L. (1981). *Manual for forms L and M of the Peabody Picture Vocabulary Test—Revised*. Circle Pines, MN: American Guidance Service.

Durmer, J. S., & Dinges, D. F. (2005). Neurocognitive consequences of sleep deprivation. *Seminars in Neurology*, *25*, 117–29.

Eberlin, M., McConnachie, G., Ibel, S., & Volpe, L. (1993). Facilitated communica-

tion: A failure to replicate the phenomenon. *Journal of Autism and Developmental Disorders*, *23*, 507–530.

Eimer, M. (2000). Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clinical Neurophysiology*, *111*, 694–705.

Eriksen, C. W. (1960). Discrimination and learning without awareness: A methodological survey and evaluation. *Psychological Review*, *67*, 279–300.

Esiri, M. M., Wilcock, G. K., & Morris, J. H. (1997). Neuropathological assessment of the lesions of significances in vascular dementia. *Journal of Neurology, Neurosurgery, and Psychiatry*, *63*, 749–753.

Everitt, B. J., Dickinson, A., & Robbins, T. W. (2001). The neuropsychological basis of addictive behavior. *Brain Research Reviews*, *36*, 129–138.

Fausto-Sterling, A. (2000). *Sexing the body: Gender politics and the construction of sexuality*. New York: Basic Books.

Ferriero, D., & Dempsey, D. (1999). Impact of addictive and harmful substances on fetal brain development. *Current Opinion in Neurobiology*, *12*, 161–166.

Fodor, J. (1975). *Language of thought*. New York: Crowell.

Frisch, K. von. (1976). *The dance language and orientation of bees*. Cambridge, MA: Harvard University Press.

Fukuda, T. (1984). *Statokinetic reflexes in equilibrium and movement*. Tokyo: University of Tokyo Press.

Fuster, J. M. (1989). *The prefrontal cortex* (2nd ed.). New York: Raven.

Gaffan, D. (1974). Recognition impaired and association intact in the memory of monkeys after transection of the fornix. *Journal of Comparative and Physiological Psychology*, *86*, 1100–1109.

Gallagher, M., & Pellymounter, M. (1988). Spatial learning deficits in old rats: A model for memory decline in the aged. *Neurobiology of Aging*, *9*, 549–556.

Gambaryan, P. P. (1974). *How animals run*. New York: Wiley.

George, C., Kaplan, N., & Main, M. (1985). *An adult attachment interview* (ed.). Berkeley, CA: University of California, Berkeley.

Glantz, M., & Pickens, R. (1992). *Vulnerability to drug abuse*. Washington, D.C: American Psychological Association.

Goedert, M. (1996). Tau protein and the neurofibrillary pathology of Alzheimer's disease. *Annual of the New York Academy of Sciences*, *777*, 121-131.

Goldberg, R. (2003). *Drugs across the spectrum*. Toronto: Thomson Wadsworth.

Goozen, S. H. M. van, Cohen-Kettenis, P. T., Gooren, L. J. G., Frijda, N. H., & Poll, N. E. Van de. (1995). Gender differences in behavior: Activating effects of cross-sex hormones. *Psychoneuroendocrinology*, *20*, 343-363.

Grady, C. L., Haxby, J. V., Horwitz, B., Sundaram, M., Friedland, R., & Rapopart, S. I. (1988). Longitudinal study of the early neuropsychological and cerebral metabolic changes in dementia of the Alzheimer's type. *Journal of Clinical and Experimental Neuropsychology*, *10*, 576–596.

Greenough, W., Black, J., & Wallace, C. (1987). Experience and brain development. *Child Development*, *58*, 539–559.

Greenwald, A. G. (1992). Unconscious cognition reclaimed. *American Psychologist*, *47*, 766–779.

Gruendal, A. D., & Arnold, W. J. (1974). Influence of preadolescent experiential

factors on the development of sexual behavior in albino rats. *Journal of Comparative and Physiological Psychology*, *86*, 172–178.

Gwynne, D. (1991). Sexual competition among females. *Trends in Evolution & Ecology*, *6*, 118–121.

Gyger, M., Marler, P., & Pickert, R. (1978). Semantics of an avian alarm call system: The male domestic fowl, *Gallus domesticus*. *Behaviour*, *102*, 15–40.

Haase, E. (1977). Diseases presenting as dementia. In C. Wells (Ed.), *Dementia*. Philadelphia: Davis.

Hamilton, W. (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology*, *7*, 1–52.

Hampson, E., & Kimura, D. (1992). Sex differences and hormonal influences on cognitive function in humans. In J. B. Becker, S. M. Breedlove, & D. Crews (Eds.), *Behavioral endocrinology* (p. 357-398). Cambridge, MA: MIT Press.

Hampson, E., Rovet, J. F., & Altmann, D. (1998). Spatial reasoning in children with congential adrenal hyperplasia due to 21-hydrozylase deficiency. *Developmental Neuropsychology*, *14*, 299-320.

Harris, L. J. (1978). Sex differences in spatial ability: Possible environmental, genetic and neurlogical factors. In M. Kinsbourne (Ed.), *Asymmetrical function of the brain* (p. 405-522). London: Cambridge University Press.

Hastings, M. H. (1997). Central clocking. *Trends in Neuroscience*, *20*, 459–464.

Hauser, M., & Marler, P. (1993). Food-associated calls in rhesus macaques (*Macaca mulatta*): I. Socioecological factors. *Behavioural Ecology*, *4*, 194–205.

Hay, W., Barlow, D., & Hay, L. (1981). Treatment of stereotypic cross-gender motor behavior using covert modeling in a boy with gender identity confusion. *Journal of Clinical Psychiatry*, *49*, 388-394.

Hazan, C., & Zeifman, D. (1994). Sex and the psychological tether. *Advances in Personal Relationships*, *5*, 151-177.

Hebb, D. (1949). *The organization of behaviour.* New York: McGraw-Hill.

Hepburn, C. G., Loughlin, C. A., & Barling, J. (1997). Coping with chronic work stress. In B. H. Gottlieb (Ed.), *Coping with chronic stress* (p. 343-366). New York: Plenum Press.

Herdt, G. (1996). Mistaken sex: Culture, biology and the third sex in New Guinea. In G. Herdt (Ed.), *Third sex, third gender: Beyond sexual dimorphism in culture and history* (p. 419-446). New York: Zone Books.

Heutink, P., Stevens, M., Rizzu, P., Bakker, E., Kros, J., Tibben, A., et al. (1997). Hereditary frontotemporal dementia is linked to chromosome 17q21-q22: a genetic and clinicopathological study of three Dutch families. *Annals of Neurology*, *41*, 150–159.

Hicks, A., Martin, K., Ditor, D., Latimer, A., Craven, C., Bugaresti, J., et al. (2003). Long-term exercise training in persons with spinal cord injury: Effects on strength, arm ergometry performance and psychological well-being. *Spinal Cord*, *41*, 34-43.

Higham, P. A., & Vokey, J. R. (2000). The controlled application of a strategy can still produce automatic effects: Reply to Redington (2000). *Journal of Experimental Psychology: General*, *129*, 476–480.

Higham, P. A., Vokey, J. R., & Pritchard, J. L. (2000). Beyond dissociation logic:

Evidence for controlled and automatic influences in artificial grammar learning. *Journal of Experimental Psychology: General.*, *129*, 457–470.

Hirsh, R. (1974). The hippocampus and contextual retrieval of information from memory: A theory. *Behavioral Biology*, *12*, 421–444.

Hirsh, R., & Krajden, J. (1982). The hippocampus and expression of knowledge. In R. L. Isaacson & N. E. Spear (Eds.), *The expression of knowledge* (pp. 213–241). New York: Plenum.

Hobson, J. A. (1989). *Sleep*. New York: W. H. Freeman & Co.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences*, *9*, 1–66.

Homa, D., Sterling, S., & L., T. (1981). Limitations of exemplar-based generalization and the abstraction of categorical information. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 418–439.

Horne, J. (1988). *Why we sleep: the functions of sleep in humans and other mammals*. New York: Oxford University Press.

Howlett, A. C., Breivogel, C. S., Childers, S. R., Deadwyler, S. A., Hampson, R. E., & Porrino, L. J. (2004). Cannabinoid physiology and pharmacology: 30 years of progress. *Neuropharmacology*, *47*, 345–358.

Hruska, R. E., & Silbergeld, E. K. (1979). Abnormal locomotion in rats after bilateral intrastriatal injection of kainic acid. *Life Sciences*, *25*, 181–194.

Hudson, A., Melita, B., & Arnold, N. (1993). Brief report: A case study assessing the validity of facilitated communication. *Journal of Autism and Developmental Disorders*, *23*, 165–173.

Hukkanen, J., Jacob 3rd, P., & Benowitz, N. (2005). Metabolism and disposition kinetics of nicotine. *Pharmacological Reviews*, *57*, 79-115.

Hutt, C. (1966). Exploration and play in children. *Symposia of the Zoological Society of London*, *18*, 61–81.

Hyman, B. T., Van Hoesen, G. W., Damasio, A. R., & Barnes, C. L. (1984). Alzheimer's disease: cell-specific pathology isolates the hippocampal formation. *Science*, *225*, 1168–1170.

Iversen, L. L. (2000). *The science of marijuana*. New York: Oxford University Press.

Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., & Rosenthal, R. A. (1964). *Role stress: Studies in role conflict and ambiguity*. New York: Wiley.

Kalat, J. W. (2004). *Biological psychology*. Toronto: Thomson Wadsworth.

Kaler, S., & Freeman, B. (1994). Analysis of environmental deprivation: (c)ognitive and social development in romanian orphans. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *35*, 769-782.

Kapp, B. S., Wilson, A., Pascoe, J., Supple, W., & Whalen, P. J. (1990). A neuroanatomical systems analysis of conditioned bradycardia in the rabbit. In M. Gabriel & J. Moore (Eds.), *Learning and computational neuroscience: Foundations of adaptive networks* (pp. 53–90). Cambridge, MA: MIT Press.

Katzman, R., & Kawas, C. (1994). The epidemiology of dementia and alzheimer's disease. In R. Terry, R. Katzman, & K. Bick (Eds.), *Alzheimer's disease* (pp. 105–122). New York: Raven Press.

Kavanau, J. L. (1994). Sleep and dynamic stabilization of neural circuits: a review and synthesis. *Behavioral Brain Research*, *63*, 111–126.

Kessler, S. (1998). *Lessons from the intersexed*. New Brunswick, NJ: Rutgers University Press.

Key, W. B. (1973). *Subliminal seduction.* Englewood Cliffs, New Jersey: Signet.

Key, W. B. (1976). *Media sexploitation*. Englewood Cliffs, New Jersey: Prentice-Hall.

Key, W. B. (1980). *The clam-plate orgy*. Englewood Cliffs, New Jersey: Prentice-Hall.

Key, W. B. (1990). *The age of manipulation: The con in confidence, the sin in sincere*. Englewood Cliffs, New Jersey: Prentice-Hall.

Kleitman, N. (1963). *Sleep and wakefulness*. Chicago: University of Chicago Press.

Kolb, B. (1990). Animal models for human pfc-related disorders. *Progressive Brain Research*, *85*, 501–520.

Kolb, B. (1995). *Brain plasticity and behaviour.* New Jersey: Lawrence Erlbaum Associates.

Kolb, B., & Gibb, R. (1991). Environmental enrichment and cortical injury: Behavioral and anatomical consequences of frontal cortex lesions. *Cerebral Cortex*, *1*, 189–198.

Kolb, B., Gibb, R., & van der Kooy, D. (1994). Neonatal frontal cortical lesions in rats alter cortical structure and connectivity. *Brain Research*, *645*, 85-97.

Kolb, B., & Whishaw, I. (1994). *Fundamentals of human neuropsychology.* (Fourth ed.). New York: W.H. Freeman.

Lakke, J. P. W. F. (1985). Axial apraxia in Parkinson's disease. *Journal of Neurological Sciences*, *69*, 37–46.

Lalumière, M. L., Quinsey, V. L., & Craig, W. M. (1996). Why children from the same family are so different from one another: A Darwinian note. *Human Nature*, *7*, 281-290.

Lancaster, J., & Lancaster, C. (1987). The watershed: (c)hange in parental-investment and family-formation strategies in the course of human evolution. In J. Lancaster, J. Altmann, A. Rossi, & L. Sherrod (Eds.), *Parenting across the life span: (b)iosocial dimension* (p. 187-205). New York: Aldine de Gruyter.

Landfield, P. W., Baskin, R. K., & Pitler, T. A. (1981). Brain aging correlates: retardation by hormonal-pharmacological treatments. *Science*, *214*, 581–584.

Landfield, P. W., Waymire, J., & Lynch, G. (1978). Hippocampal aging and adrenocorticoids: Quantitative correlations. *Science*, *202*, 098–1102.

Lazarus, R., & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.

Lieberman, P. (1991). *Uniquely human: The evolution of speech, thought and selfless behaviour*. Cambridge: MIT Press.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, *56*, 1479-1498.

Lupien, S. J., de Leon, M., de Santi, S., Convit, A., Tarshish, C., Nair, N. P., et al. (1998). Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nature Neuroscience*, *1*, 69–73.

Lupien, S. J., Gaudreau, S., Tchiteya, B. M., Maheu, F., Sharma, S., Nair, N. P. V., et al. (1997). Stress-induced declarative memory impairments in healthy elderly subjects: Relationships with cortisol reactivity. *Journal of Clinical and Endocrinological Metabolism*, *82*, 2070–2075.

Luria, A. R. (1968). *The mind of a mnemonist: a little book about a vast memory* (Translated from the Russian by Lynn Solotaroff ed.). New York: Basic Books.

Macedonia, J. (1990). What is communicated in the anti-predator calls of lemurs: Evidence from playback experiments with ring-tailed and ruffed lemurs. *Ethology*, *86*, 177–190.

Magnus, R. (1924). *Körperstellung*. Berlin: Springer.

Main, M., & Solomon, J. (1990). Procedures for identifying infants as disorganized/disoriented during the (a)insworth strange situation. In M. Greenberg, D. Cicchetti, & E. Cummings (Eds.), *Attachment in preschool years: (t)heory, research, and intervention* (p. 121-160). Chicago: University of Chicago.

Maisto, S. A., Galizio, M., & Connors, G. J. (2004). *Drug use and abuse*. Toronto: Thomson Wadsworth.

Martin, J. P. (1967). *The basal ganglia and posture*. London: Pitman Medical Publishing.

Martin, P., & Caro, T. M. (1985). On the functions of play and its role in behavioral development. *Advances in the Study of Behavior*, *15*, 59–103.

Martin, R. (1990). *Primate origins and evolution: A phylogenetic reconstruction*. Princeton, NJ: Princeton University Press.

Maynard Smith, J. (1955). Fertility, mating behaviour and sexual selection in *Drosophila subobscura*. *Genetics*, *54*, 261–279.

McEwen, B. S. (2000). The neurobiology of stress: From serendipity to clinical relevance. *Brain Research*, *886*, 172-189.

McGlone, J. (1980). Sex differences in human brain asymmetry: A critical survey. *Behavioral and Brain Sciences*, *3*, 215-227.

McIntosh, C., & Chick, J. (2004). Alcohol and the nervous system. *Journal of Neurology Neurosurgery and Psychiatry*, *75*(Supplement 3), 16–21.

McIver, T. (1988). Backward masking, and other backward thoughts about music. *Skeptical Inquirer*, *13*, 50–63.

Meaney, M. J., Stewart, J., & Beatty, W. W. (1985). Sex differences in social play: The socialization of sex roles. *Advances in the Study of Behavior*, *15*, 1–58.

Medin, D. L., & Ross, B. H. (1989). The specific character of abstract thought: Categorization, problem solving, and induction. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 5, pp. 189–223). New York: Academic Press.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology*, *35*, 113–138.

Merikle, P. M. (1988). Subliminal auditory tapes: An evaluation. *Psychology & Marketing*, *46*, 355–372.

Mistlberger, R. E. (1994). Circadian food-anticipatory activity: formal models and physiological mechanisms. *Neuroscience & Biobehavioral Reviews*, *18*, 171–195.

Mitani, J. C. (1985). Mating behaviour of male orangutans in the kutai reserve. *Animal Behavior*, *33*, 392-402.

Moore, C. (1985). Development of mammalian sexual behavior. In E. S. Gollin (Ed.), *The comparative development of adaptive skills: Evolutionary implications.* New Jersey: Erlbaum.

Moore, K., & Persaud, T. (1998). *The developing human: Clinically oriented embryology* (Sixth ed.). Philadelphia: W.B. Saunders.

Moore, S., Donovan, B., & Hudson, A. (1993). Brief report: Facilitator-suggested conversational evaluation of facilitated communication. *Journal of Autism and Developmental Disorders, 23*, 541–552.

Moore, S., Donovan, B., Hudson, A., Dykstra, J., & Lawrence, J. (1993). Brief report: Evaluation of eight case studies of facilitated communication. *Journal of Autism and Developmental Disorders*, *23*, 531–539.

Moore, T. E. (1992). Subliminal perception: Facts and fallacies. *Skeptical Inquirer*, *16*, 273–281.

Moos, R., & Billings, A. (1982). Conceptualizing and measuring coping resources and processes. In L. Goldberger & S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects.* New York: Free Press.

Morrison, A. R., Sanford, L. D., Ball, W. A., Mann, G. L., & Ross, R. J. (1995). Stimulus-elicited behavior in rapid eye movement sleep without atonia. *Behavioral Neuroscience*, *109*, 972–979.

Mortimer, J. A., French, L. R., Hutton, J. T., & Schuman, L. M. (1985). Head injury as a risk factor for alzheimer's disease. *Neurology*, *35*, 264–267.

Mortimer, J. A., van Duijn, C. M., Chandra, V., Fratiglioni, L., Graves, A. B., Heyman, A., et al. (1991). Head trauma as a risk factor for alzheimer's disease: a collaborative re-analysis of case-control studies. *International Journal of Epidemiology*, *20*, 28–35.

Moscovitch, M. (1994). Memory and working with memory: evaluation of a component process model and comparisons with other models. In D. L. Schacter & E. Tulving (Eds.), *Memory systems* (pp. 269–310). Cambridge, MA: MIT Press.

Mukhametov, L. M., Supin, A. Y., & Polyakova, I. G. (1977). Interhemispheric asymmetry of the electroencephalographic sleep patterns in dolphins. *Brain Research*, *134*, 581–584.

Mumby, D. G., & Pinel, J. P. J. (1994). Rhinal cortex lesions and object recognition in rats. *Behavioral Neuroscience*, *108*, 11–18.

Murphy, L. R. (1992). Workplace interventions for stress reduction and prevention. In C. L. Cooper & R. Payne (Eds.), *Causes, coping, and consequences of stress at work* (p. 301-339). Chichester, England: Wiley.

Nanda, S. (2000). *Gender diversity: Crosscultural variations*. Prospect Heights, IL: Waveland Press.

Neave, N., Hamilton, C. J., Hutton, L., & Pickering, A. T. (In press). Female advantage in location memory using ecologically valid measures. *Proceedings of the British Psychological Society*.

O'Connor, P. J., & Youngstedt, S. D. (1995). Influence of exercise on human sleep. *Exercise and Sports Science Reviews*, *23*, 105–134.

O'Keefe, J. A., & Nadel, L. (1978). *The hippocampus as a cognitive map*. London: Oxford University Press.

Orgeur, P., & Signoret, J. (1984). Sexual play and its functional significance in the domestic sheep (*Ovis aries L.*). *Physiology & Behavior*, *33*, 111–118.

Packard, M. G., Hirsh, R., & White, N. M. (1989). Differential effects of fornix and

caudate nucleus lesions on two radial maze tasks: Evidence for multiple memory systems. *Journal of Neuroscience, 9*, 1465–1472.

Packer, C. (1977). Reciprocal altruism in *Papio anubis*. *Nature, 265*, 441–443.

Pagel, M., & Harvey, P. (1992). Evolution of the juvenile period in mammals. In M. Pereira & L. Fairbanks (Eds.), *Juvenile primates.* Oxford: Oxford University Press.

Pavlov, I. P. (1960). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex.* New York: Dover.

Pellis, S. M. (1993). Sex and the evolution of play fighting: A review and model based on the behavior of muroid rodents. *The Journal of Play Theory and Research, 1*, 55–75.

Pellis, S. M., & Pellis, V. C. (1993). Influence of dominance on the development of play fighting in pairs of male Syrian golden hamsters (*Mesocricatus auratus*). *Aggressive Behavior, 19*, 293–302.

Pellis, S. M., & Pellis, V. C. (1994). The development of righting when falling from a bipedal standing posture: Evidence for the dissociation of dynamic and static righting reflexes in rats. *Physiology & Behavior, 56*, 659–663.

Pellis, S. M., Pellis, V. C., Chen, Y.-c., Barzci, S., & Teitelbaum, P. (1989). Recovery from axial apraxia in the lateral hypothalamic labyrinthectomized rat reveals three elements of contact righting: Cephalocaudal dominance, axial rotation and distal limb action. *Behavioural Brain Research, 45*, 241–251.

Pellis, S. M., Pellis, V. C., Chesire, R. M., Rowland, N., & Teitelbaum, P. (1987). Abnormal gait sequence in the locomotion released by atropine in catecholamine deficient rats. *Proceedings of the National Academy of Sciences, U.S.A., 84*, 8750–8753.

Pellis, S. M., Pellis, V. C., & McKenna, M. M. (1993). Some subordinates are more equal than others: Play fighting amongst adult subordinate males. *Aggressive Behavior, 19*, 385–393.

Pellis, S. M., Pellis, V. C., & Nelson, J. E. (1992). The development of righting reflexes in the pouch young of the marsupial dasyurus hallucatus. *Developmental Psychobiology, 25*, 105–125.

Pellis, S. M., Pellis, V. C., & Teitelbaum, P. (1991). A descriptive analysis of the postnatal development of contact-righting in rats (*Rattus norvegicus*). *Developmental Psychobiology, 24*, 237–263.

Pellis, S. M., Pellis, V. C., & Whishaw, I. Q. (1992). The role of the cortex in play fighting by rats: Developmental and evolutionary implications. *Brain, Behavior and Evolution, 39*, 270–284.

Pericak-Vance, M. A., & Haines, J. L. (1995). Genetic susceptibility to Alzheimer's disease. *Trends in Genetics, 11*, 504–508.

Perry, E. K., Perry, R. H., Blessed, G., & Tomlinson, B. E. (1977). Necropsy evidence of central cholinergic deficits in senile dementia. *Lancet, i*, 189.

Phillips, R. D. (1985). Whistling in the dark? A review of play therapy research. *Psychotherapy, 22*, 752–760.

Pierce, C. S., & Jastrow, J. (1884). On small differences in sensation. *Memoirs of the National Academy of Science, 3*, 73–83.

Pinker, S. (1994). *The language instinct*. New York: W. Morrow.

Plaut, S., & Friedman, S. (1982). Stress, coping behavior and resistance to disease. *Psychotherapy and Psychosomatics*, *38*, 274-283.

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences*, *10*, 1-60.

Poundstone, W. (1983). *Big secrets*. New York: Quill.

Poundstone, W. (1986). *Bigger secrets*. Boston: Houghton Mifflin Company.

Poundstone, W. (1993). *Biggest secrets*. New York: William Morrow and Company, Inc.

Pratkanis, A. R. (1992). The cargo-cult science of subliminal persuasion. *Skeptical Inquirer*, *16*, 260–271.

Pratkanis, A. R., Eskenazi, J., & Greenwald, A. G. (1990, April). *What you expect is what you believe (but not necessarily what you get): On the effectiveness of subliminal self-help audiotapes.* Paper presented at the meeting of the Western Psychological Association, Los Angeles, California.

Premkumar, D. R., Cohen, D. L., Hedera, P., Friedland, R. P., & Kalaria, R. N. (1996). Apolipoprotein e-4 alleles in cerebral amyloid angiopathy and cerebrovascular pathology associated with Alzheimer's disease. *American Journal of Pathology*, *148*, 2083–2095.

Raskind, M., Peskind, E., Rivard, M., Veith, R., & Barnes, R. (1982). Dexamethasone suppression test and cortisol circadian rhythm in primary degenerative dementia. *American Journal of Psychiatry*, *139*, 1468.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219–235.

Reber, A. S., & Allen, R. (1978). Analogy and abstraction strategies in synthetic grammar learning: A functionalist interpretation. *Cognition*, *6*, 189–221.

Rechtschaffen, A. (1998). Current perspectives on the function of sleep. *Perspectives in Biology and Medicine*, *41*, 359–390.

Reimund, E. (1994). The free radical flux theory of sleep. *Medical Hypotheses*, *43*, 231–233.

Rendall, D., Rodman, P., & Emond, R. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour*, *51*, 1007–1015.

Rendall, D., Seyfarth, R., Cheney, D., & Owren, M. (1999). The meaning and function of grunt variants in baboons. *Animal Behaviour*, *57*, 583–592.

Richter, C. P. (1971). Inborn nature of the rat's 24-hour clock. *Journal of Comparative and Physiological Psychology*, *75*, 1-4.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization.* Hillsdale, NJ: Erlbaum.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*, 573–605.

Ross, B. H. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology*, *16*, 371–416.

Ross, B. H., & Kennedy, P. T. (1990). Generalizing from the use of earlier examples in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 42–55.

Sacks, O. (1970). *The man who mistook his wife for a hat and other clinical tales*. New York: Touchstone.

Sandstrom, N. J., Kaufman, J., & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task. *Cognitive Brain Research, 6*, 351-360.

Sapolsky, R. M. (1987). Stress and glucocorticoids in aging. *Endocrinology and Metabolism Clinics of North America, 16*, 965-980.

Sapolsky, R. M. (1992). *Stress, the aging brain, and the mechanisms of neuron death*. Cambridge, MA: MIT Press.

Sapolsky, R. M. (1994). *Why zebras dont get ulcers: An updated guide to stress, stress-related diseases, and coping*. New York: W. H. Freeman.

Sapolsky, R. M. (1997). McEwen-induced modulation of endocrine history: A partial review. *Stress, 2*, 1-12.

Sauter, S. L., Murphy, L. R., & Hurrell, J. J. (1990). Prevention of work-related psychologial disorders: A national strategy proposed by the National Institute for Occupational Safety and Health (NIOSH). *American Psychologist, 45*, 1146-1158.

Scarr, S., & McCartney, K. (1983). How people make their own environments: a theory of genotype-environment effects. *Child Development, 54*, 424–435.

Scholinski, D. (1998). *The last time i wore a dress*. New York: Riverhead Books.

Schwab, M., & Bartholdi, D. (1996). Degeneration and regeneration of axons in the lesioned spinal cord. *Physiological Reviews, 76*, 319-370.

Scott, T., Peckham, P., & Keith, M. (1995). Upper extremity neuroprostheses using functional electrical stimulation. *Baillieres Clinical Neurology, 4*, 57-75.

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

Selye, H. (1950). *The physiology and pathology of exposure to stress*. Montreal: Acta Medical Publishers.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences, 17*.

Shreiner Institute for Health Policy. (2001). *Substance abuse: The nation's number one health problem*. Waltham, MA: Heller Graduate School, Brandeis University.

Sidis, B. (1898). *The psychology of sensation*. New York: Appleton.

Silverman, I., Choi, J., Mackewn, A., Fisher, M., Moro, J., & Olshansky, E. (2000). Evolved mechanisms underlying wayfinding: Further studies on the hunter-gatherer theory of spatial sex differences. *Evolution & Human Behaviour, 21*, 201-213.

Silverman, I., & Eals, M. (1992). Sex differences in spatial abilities: Evolutionary theory and data. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (p. 531-549). New York: Oxford Press.

Silverman, I., & Phillips, K. (1993). Effects of estrogen changes during the menstrual cycle on spatial performance. *Ethology & Sociobiology, 14*, 257-270.

Simpson, J., & Rholes, W. (1998). Attachment in adulthood. In J. Simpson & W. Rholes (Eds.), *Attachment theory and close relationships* (p. 3-21). New York: The Guilford Press.

Skinner, B. (1938). *The behaviour of organisms.* New York: D. Appleton-Century.

Smith, C. (1985). Sleep states and learning: A review of the literature. *Neuroscience and Biobehavioral Reviews*, *9*, 157–168.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Smuts, B. B., & Smuts, R. W. (1993). Male aggression and sexual coercion of females in nonhuman primates and other mammals: Evidence and theoretical implications. *Advances in the Study of Behaviour*, *22*, 1–63.

Snowdon, D. A., Greiner, L. H., Mortimer, J. A., Riley, K. P., Greiner, P. A., & Markesbery, W. R. (1997). Brain infarction and the clinical expression of Alzheimer's disease. the nun study. *Journal of the American Medical Association*, *277*, 813–817.

Sorenson, A. P. (1984). Linguistic exogamy and personal choice in the northwest amazon. In K. M. Kensinger (Ed.), *Marriage practices in lowland south america* (pp. 180–193). Urbana: University of Illinois Press.

Squire, L. R., & Knowlton, B. J. (1994). Memory, hippocampus and brain. In M. Gazzaniga (Ed.), *Cognitive neurosciences* (pp. 825–837). Cambridge, MA: MIT Press.

Suter, P. M., Schutz, Y., & Jequier, E. (1992). The effect of ethanol on fat storage in healthy subjects. *New England Journal of Medicine*, *326*, 983–987.

Sutherland, R. J., & Rudy, J. W. (1989). Configural association theory: the role of the hippocampal formation in learning, memory, and amnesia. *Psychobiology*, *17*, 129–144.

Swaab, D. F., Fliers, E., & Partiman, T. S. (1985). The suprachiasmatic nucleus of the human brain in relation to sex, age and senile dementia. *Brain Research*, *342*, 37–44.

Szempruch, J., & Jacobson, J. W. (1993). Evaluating facilitated communications of people with developmental disabilities. *Research in Developmental Disabilities*, *14*, 253–264.

Tator, C. H. (1972). Acute spinal cord injury: A review of recent studies of treatment and pathophysiology. *Canadian Medical Association Journal*, *107*, 143-145.

Thorndike, E. L., & Rock, R. T. (1934). Learning without awareness of what is being learned or intent to learn it. *Journal of Experimental Psychology*, *17*, 1–19.

Tomasello, M., & Call, J. (1997). *Primate cognition*. Oxford: Oxford University Press.

Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (p. 19-136). New York: Oxford Press.

Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35–57.

Trivers, R. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the decent of man, 1871-1971* (pp. 136–179). Chicago: Aldine.

Troiani, O., Petrosini, L., & Passani, F. (1981). Trigeminal contribution to the head righting reflex. *Physiology & Behavior*, *27*, 157–160.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). San Diego, CA: Academic Press.

Van Someren, E. J. W., Hagebeuk, E. E. O., Lijenga, C., Scheltens, P., de Rooij,

S. E., Jonker, C., et al. (1996). Circadian rest-activity rhythm disturbances in Alzheimer's disease. *Biological Psychiatry*, *40*, 259–270.

Vance and Belknap v. Judas Priest and CBS Records. (1990, August 24). *86-5844/86-3939.* Second District Court of Nevada.

Vasey, P. L. (1998). Female choice and inter-sexual competition for female sexual partners in japanese macaques. *Behavior*, *135*, 579–597.

Vokey, J. R., & Brooks, L. (1994). Fragmentary knowledge and the processing-specific control of structural sensitivity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 328–344.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 328–344.

Vokey, J. R., & Higham, P. A. (1999). Implicit knowledge as automatic, latent knowledge. *Behavioral and Brain Sciences*, *22*, 787–788.

Vokey, J. R., & Read, J. D. (1985). Subliminal messages: Between the devil and the media. *American Psychologist*, *40*, 1231–1239.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*, 250-270.

Walker, P., & Cook, D. (1998). Gender and sex: Vive la difference. *American Journal of Physical Anthropology*, *106*, 255-259.

Watson, N. V., & Kimura, D. (1991). Nontrivial sex differences in throwing and intercepting: Relation to psychometrically-defined spatial functions. *Personality and Individual Differences*, *12*, 375-385.

Wever. (1979). *The circadian system of man*. New York: Springer-Verlag.

Wheeler, D. L., Jacobson, J. W., Paglieri, R. A., & Schwartz, A. A. (1993). An experimental assessment of facilitated communication. *Mental Retardation*, *31*, 49–60.

Whishaw, I., Gorny, B., Tran-Nguyen, L. T., Castañeda, E., Miklyaeva, E., & Pellis, S. (1994). Making two movements at once: Impairments of movement, posture, and their integration underlie the adult skilled reaching deficit of neonatally dopamine-depleted rats. *Behavioural Brain Research*, *61*, 65–77.

Whishaw, I., & Pellis, S. (1990). The structure of skilled forelimb reaching in the rat: A proximally driven movement with a single distal rotatory component. *Behavioural Brain Research*, *41*, 49–59.

Whishaw, I. Q., & Kolb, B. (1985). The mating movements of male decorticate rats: Evidence for subcortically generated movements by the male but regulation of approaches by the female. *Behavioural Brain Research*, *17*, 171–191.

Whittlesea, B. W. A. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *1*, 3–17.

Whorf, B. (1956). *Language, thought and reality*. Cambridge, MA: MIT Press.

Wiese, J. G., Shlipak, M. G., & Browner, W. S. (2000). The alcohol hangover. *Annals of Internal Medicine*, *132*, 897–902.

Wikan, U. (1977). Man becomes woman: Transsexualism in Oman as a key to gender roles. *Man*, *12*, 304-319.

Winson, J. (1993). The biology and function of rapid eye movement sleep. *Current Opinion in Neurobiology*, *3*, 243-248.

Wirz, M., Zemon, D., Rupp, R., Scheel, A., Colombo, G., Dietz, V., et al. (2005). Effectiveness of automated locomotor training in patients with chronic incomplete spinal cord injury: A multicenter trial. *Archives of Physical Medicine and Rehabilitation*, *86*, 672-680.

Wolpert, L. (1992). *The unnatural nature of science*. London: Faber and Faber.

Wonnacott, S., Sidhpura, N., & Balfour, D. J. (2005). Nicotine: from molecular mechanisms to behaviour. *Current Opinions in Pharmacology*, *5*, 53–59.

Zeifman, D., & Hazan, C. (1997). Attachment: (t)he bond in pair-bonds. In J. Simpson & D. Kenrick (Eds.), *Evolutionary social psychology* (p. 237-263). Mahwah, NJ: Erlbaum.

Index

Author Index