27
DEVELOPMENT AND LEARNING

Aarre Laakso

Introduction

Questions about learning and development are questions about the possibilities and limits of human nature – about the origins and prospects of our most basic abilities and our most unique traits. In addition to their intrinsic interest, these questions have important policy implications. Our theories about development and learning, whether they are informed or naïve, guide social policy in many areas, including medicine and education. Insofar as we know the means of cultivating the potential of our offspring, we will surely want to do so to the greatest extent possible. Hence, our understanding of development and learning shapes our individual and group practices of healthcare, child rearing, education and personal growth. Our theories about development and learning also guide our individual behavior toward children (and adults as well) as parents, relatives, neighbors, childcare providers, teachers, spouses, friends, colleagues, managers and mentors.

There is not enough space here to discuss every facet of development and learning. Learning is studied in many fields, including psychology, philosophy, mathematics, statistics, computer science, linguistics, neurophysiology, molecular biology and artificial intelligence. Each has its own methods, vocabulary and theories. The goal of this essay is to give the reader a sense of the diversity of theoretical issues about learning and development that may be of interest to philosophically inclined psychologists and psychologically inclined philosophers. Because this volume has separate entries on "Nativism" (Chapter 19) and "Evolutionary Models in Psychology" (Chapter 26) discussion of those issues is limited in this piece to the way they are treated in the developmental literature.

An uncontroversial but vague way of defining ***development*** is "the pattern of continuity and change in an entity over time." Philosophers of psychology are particularly concerned with the development of biological organisms, particularly human beings (human development). The interest is not limited to human beings, however. It extends to other organisms, especially animals, and to non-biological systems, including computers, robots and abstractions. Among both philosophers and psychologists, the interest in human development is most intense with regard to child development, because this is the period of most rapid and obvious transformation. The psychological study of human development, particularly child development, is known as "developmental psychology." It has three main branches: cognitive development, emotional development and social development. In recent years, a broader, more interdisciplinary field called "developmental science" has coalesced around studies of development more generally, in all species and across all periods of the life span.

It is more difficult to define *learning*, because one's definition of learning depends a great deal on one's theory of development. The distinction between learning and development is intimately tied up with questions about nativism. For the extreme empiricist, *all* mental development is learning. For the extreme nativist, learning is impossible and so does not exist. Another difficulty is that philosophers tend to take a narrower view of learning than psychologists do. On a standard psychological definition, "learning" refers to a change in behavior that results from interaction with the environment, excluding behavioral changes attributable to genetic factors, physiological influences (such as illness, fatigue, drugs or "blows to the head"), and maturation (e.g., Lewis et al. 1969). By contrast, an important philosophical definition of learning is "the acquisition of a form of knowledge or ability through the use of experience" (Hamlyn 1995: 496). Note the emphasis on *knowledge* or *ability* in contrast to *behavior*. Indeed, Hamlyn specifically denies that conditioning or imprinting are forms of learning, writing that "For learning to take place experience has to be *used* in some way, so that what results is in a genuine sense knowledge or is dependent on knowledge" (496). For present purposes, we may say that the term "development" refers to changes in an organism that tend to be broader, longer term, more biologically constrained, and more physical in nature than changes described as "learning."

From antiquity, three assumptions have framed philosophical debates about the nature of learning and development. The first is that the unit of analysis is the person (or, more generally, the whole organism), not some higher or lower level unit (such as the family or the cell). The second assumption is factors contributing to learning and development can be classified as either "innate" or "environmental." The final assumption is that there are two types of learning: (a) explicit learning of declarative, propositional knowledge to which the learner has conscious access; and (b) implicit learning of behavioral skills, to which the learner has no conscious access. Over the last 50 years, developments in psychology and other fields have challenged all three assumptions. The third section reviews the evidence that the individual organism is not always the appropriate unit of analysis. The fourth section reviews the conceptual difficulties with apportioning the causes of learning and development to the "genome" or the "environment." The fifth section reviews some of the wide variety of types and modes of learning that psychologists have discovered. First, however, we make a brief foray into history.

A brief history

As with most scientific concepts, the history of the concepts of development and learning is one of increasing differentiation and sophistication, driven by the goals of simultaneously encompassing all of the phenomena of interest while explaining each quantitatively and mechanistically. Developmental science, like all of the sciences, has its roots in philosophy (see the entries on the "Rationalist Roots of Modern Psychology" (Chapter 1), the "Empiricist Roots of Modern Psychology" (Chapter 2), and the "Early Experimental Psychology" (Chapter 3)). However, it also has a unique history that includes emphasis on certain works in the early canon over others as well as recent figures with distinct philosophical pedigrees.

Philosophers, starting with Plato, have traditionally distinguished just two modes of learning – "learning that" (explicit learning of propositional or declarative facts) and "learning how" (skill-based or implicit learning). Although Ryle (1949) coined the terms, the distinction is already captured in Meno's question to Socrates ("whether virtue is acquired by teaching or by practice"). In general, philosophers have tended to privilege propositional learning. Aristotle, for example, argued that "theoretical knowledge" was superior to "practical" or "productive" knowledge. However, a minority, notably Hubert Dreyfus (e.g., Dreyfus and Dreyfus 1986), have emphasized the priority of skill learning over propositional learning. (See also the entry on "Action and Mind" [Chapter 38].)

The first evidence collected about development and learning was primarily based on introspection. Augustine (1955 [397]), for example, claimed to remember in some detail his own acquisition of language. Systematic observation did not develop until the nineteenth century. Charles Darwin (1877) was an early practitioner of the "diary method" of developmental studies, whereby a caretaker keeps a careful record of the behaviors of his or her own child over time for use as scientific evidence. J. M. Baldwin (1906 [1895]) was the first to publish a report on a scientific experiment with a child (not coincidentally, his own daughter). He gave the account of child development in which mature cognitive capacities were hypothesized to develop out of simple infant behaviors in a series of qualitatively distinct stages. Although largely ignored in North America for most of the twentieth century, Baldwin's works had tremendous influence on developmental theory through Vygotsky and Piaget. G. Stanley Hall (1904) and his student Arnold Gesell (1925) compiled detailed normative information about everything from infants' motor achievements to adolescents' dreams. Freud's (1949 [1905]) psychosexual theory emphasized changes in the locus of sexual impulses during child development. Although it drew on his clinical experience with adults, Freud's theory was not based on direct studies of children and has largely fallen out of favor in studies of development. Russian psychologist Lev Vygotsky (1986 [1934]) emphasized the role of sociocultural influences in cognitive development, viewing learning as a process of interaction in social contexts that affects the relationship between people and their environment. Piaget developed his classic cognitive-developmental theory of child development in Switzerland starting in the 1930s. Baldwin's theory was a major influence on Piaget (1999 [1954]), and Piaget's is still the best-known stage theory. Piaget was not well known in North America until the 1960s, however, due to the dominance of behaviorism. Inspired by Pavlov's (1927) discovery of classical conditioning through studies of animal learning, as well as by an urge to rid psychology of the influence of psychoanalytic and metaphysical theorizing, Watson and Skinner emphasized the role of the environment in development and learning. Based primarily on animal studies, Skinner's (1935) theory of operant conditioning suggested that the frequency of a child's behavior could be increased by reinforcing it and decreased by punishing it through the application of external stimuli. Piaget's work came to be recognized in North America as part of the cognitive turn in psychology that started in the 1950s, and it still pervades much theorizing in developmental science.

Leading contemporary theories of learning and development are diverse; they include neo-behaviorist theories, such as Albert Bandura's (1986) "social cognitive theory"; neo-Piagetian or "constructivist" theories (e.g., Case 1992); information-processing and connectionist theories (e.g., MacWhinney et al. 1989); ethological and evolutionary theories, such as John Bowlby's (1982 [1969]) "attachment theory"; neo-Vygotskian theories, such as Barbara Rogoff's (2003) "sociocultural" theory; ecological systems theories, such as Urie Bronfenbrenner's "bioecological view" (Bronfenbrenner and Morris 2006); behavioral genetics theories (e.g., Plomin 1986); developmental systems theories (e.g., Gottlieb 2007); dynamic systems theories (Thelen and Smith 1994); and many others.

Units of analysis and levels of explanation

The "unit of analysis" is the class of entity that the theorist believes to develop or learn. In both psychology and philosophy, it has usually been assumed that the unit of analysis is the person or, more generally, the biological organism. The "level of explanation" is the class of entities used in the explanation. In the canonical case (see "What Is Psychological Explanation?" [Chapter 8]), the level of explanation is a sub-personal mechanism, typically a neural one (see "The Interface between Psychology and Neuroscience" [Chapter 11] and "Levels of Mechanisms" [Chapter 24]). Many contemporary theories of learning and development, however, depart from one or both of the canonical positions.

One reason to believe that the person (or child or organism) is not always the proper unit of analysis is the existence of abstract, formal learning theories developed in linguistics, mathematics, statistics and computer science, as well as in psychology. There are many different such theories, but all of them treat learning (to a greater or lesser degree) independently of human or even animal learning. These theories are relevant to developmental psychologists and philosophers of psychology because they attempt to develop quantitative models of learning that may be applicable to biological organisms.

Mathematical learning theory, as it was developed in psychology (Estes 1950), aims to predict biological behavior quantitatively – for example, to predict the frequency of behavioral responses as a function of stimulus features. In another sense, mathematical learning theory refers to "learning from data." There are a variety of approaches, but generally speaking they concern themselves with two primary problems. One is optimization – adaptation of the learner to maximize some measure of the quality of behavior. The second is complexity – the amount of (physical, spatial or temporal) resources required to implement a solution to the optimization problem.

One important variant of mathematical learning theory is Bayesian learning theory. In general, "Bayesian learning" refers to learning that uses methods based on the Bayesian interpretation of statistics (i.e., that probabilities represent a subjective degree of certainty as opposed to an objective relative frequency). In that wide sense, Bayesian learning applies to a variety of mathematical formalisms and techniques, including Bayes' theorem, and machine learning algorithms, such as expectation maximization (Russell and Norvig 2003).

The field known as statistical learning theory has concerned itself primarily with the optimization problem. Historically, the emphasis has been on learning solutions to regression and classification problems that minimize the error between known data points and the estimated solution. More recently, the emphasis has broadened to encompass not only minimizing error on known data but also generalizing well to new data (Vapnik 1999).

Algorithmic learning theory (sometimes also known as formal learning theory or the theory of inductive inference) is concerned specifically with the learnability in principle of languages (or theories), without consideration of the feasibility of learning under time and space constraints (Gold 1967).

Computational learning theory – sometimes referred to as probably approximately correct (PAC) learning – starts from the idea that learning is the "phenomenon of knowledge acquisition in the absence of explicit programming" (Valiant 1984: 1134) that is, "the process of deducing a program for performing a task" (1142). Learning is typically viewed as the application of an algorithm that, given some input, transforms the internal state of a computer to better predict future input. Computational learning theory is the study of mathematical models of machine learning, with a particular emphasis on "learnability" – on what sorts of resources (time, memory, computer architecture) would be required to learn certain kinds of things and, conversely, what kinds of things can be learned with a given set of resources. It therefore emphasizes issues of computational efficiency. The idea is to identify the contexts in which learning is computationally feasible, meaning that it can be performed in polynomial time.

Machine learning (Mitchell 1997) is the study of computer algorithms that automatically improve their performance, often by some form of induction. It has developed out of a confluence between computer science, particularly artificial intelligence, and statistics. The tasks considered in machine learning theory are typically classification tasks, regression tasks, sequential decision-making tasks or clustering tasks. Classification and regression tasks typically assume that the learning is "supervised," in the sense that the system is provided with a set of exemplar inputs with the "correct" outputs (a set of labeled training exemplars), and expected to learn from them the ability to generalize to new inputs. Sequential decision-making tasks usually assume a semi-supervised learning regime, in which decisions are rewarded or punished, in some cases only remotely (for example, by the outcome of a competition such as a chess game). Clustering tasks typically assume that the learning is "unsupervised," that is, that the system receives no feedback.

Unlike formal learning, biological learning (particularly the sort of learning one sees in human development) is ongoing and cumulative. Each learning accomplishment builds on what has been learned previously and sets the stage for what will be learned next. Often, "what will be learned next" is qualitatively different from what has come before. Moreover, this cumulative learning occurs on different time scales over the course of the lifespan – learning to tie one's shoes and learning to play master-level chess are good examples. By contrast, machine learning algorithms are typically designed to deal with a single kind of learning problem, and to stop when they have achieved a degree of correctness or generalization. Machine learning algorithms also typically do not take into account such factors as motivation, emotional state, multimodal interactions, or limitations on memory.

However, studies in machine learning and studies in biological learning are converging. Although some types of neural networks (e.g., connectionist networks trained by backpropagation) are physiologically implausible, others approximate the actual behavior of biological neurons much better. Some of the formalisms and data structures developed in machine learning may have abstract correspondences to the activity of biological neurons, even if their mechanisms are quite different.

Besides the fact that learning is often studied in the abstract, there are more concrete indications that the individual organism is not always the most appropriate unit of analysis. There are many levels below the level of the organism that have been taken either as levels of explanation or units of analysis in theories of development and learning. These include the molecular level (DNA, RNA, polypeptides, proteins, neurotransmitters), the chemical level (methylation and acetylation, among others), the cellular level (cells, neurons), the inter-cellular level (e.g., synapses, neurotransmitters), and several distinct intra-cellular levels (in general, tissues and organs, and specifically in the nervous system: networks, circuits, functionally individuated volumes, anatomically individuated volumes, the neocortex, the brain as a whole, and the central nervous system in its entirety).

In neuroscience, learning is typically viewed at the synaptic level, as a biochemical process that has the effect of increasing or decreasing the firing rate of the downstream neuron. (See "Cellular and Subcellular Neuroscience" [Chapter 25].) Many neuroscientists agree that pairs of neurons connected by a synapse are capable of learning in a process similar to the one that Hebb (1949) suggested. Inter-cellular learning is typically explained in terms of long-term potentiation or long-term depression. In long-term potentiation (LTP), a synapse is altered by a stimulus (or series of stimuli) in such a way that subsequent activation at the presynaptic cell is more likely to elicit an action potential in the postsynaptic cell. In long-term depression (LTD), the synapse is altered in such a way that subsequent activation at the presynaptic cell is less likely to elicit an action potential in the postsynaptic cell. The exact mechanisms of LTP and LTD have been subject to intense scrutiny for some time but are not well understood and are non-uniform, differing by the organism, the brain region, the type of synapse, or the age (developmental maturity) of the organism (Malenka and Bear 2004).

Justifications for the claim that the appropriate level of explanation is sub-personal are complex and typically involve a combination of naive realism and naturalism. Both the process of learning and the result (the trace of what is learned) are taken to be situated in the mind (and therefore the brain) of the individual learner. That is, learning is a change of a property of a person, specifically a change in beliefs (the contents of an individual human mind). Jerry Fodor (1980) has defended this position on the grounds of what he calls "methodological solipsism," the view that the contents of thoughts are determined by the physical states of the individual having the thoughts. It is also a logical consequence of "psychophysical supervenience" (Kim 1982), the thesis that any psychological difference must have a corresponding physical difference. Cognitive psychologists, and many developmental scientists, often take this position, usually tacitly but sometimes explicitly.

Internal and external factors

The long debate about the respective roles of "nature" and "nurture" in development can be regarded as a debate about the respective roles of "internal" (sub-personal, genetic) and "external" (supra-personal, environmental) factors in development. In the contemporary debate, a common assumption is that variance across a population can be separated into independent genetic and environmental contributions to individual differences, plus statistical interactions of the two. As a consequence of this assumption, linear models (such as analysis of variance) have typically been used in attempts to demonstrate interactions between heredity and environment. However, pervasive evidence of gene-environment correlations and nonlinear interactions between genetic and environmental factors demonstrates that linear models are inadequate. The notion that genetic and environmental influences on development are strictly additive has given way recently to the recognition that development is shaped by frequent synergistic, non-additive interactions.

It is useful to distinguish two forms of interdependence between genotype and environment – genotype × environment interactions and genotype-environment correlations (Plomin et al. 1977). Genotype × environment interactions are the effects of environmental variation on gene expression, the reason that individuals with homogeneous genotypes respond differently to different environments. Genotype-environment correlations are genetic influences on individual variations in organisms' exposure to kinds of environments – the fact that different genotypes are selectively exposed to different environments. Parents create environments for their children that are influenced by their own heredity. For example, a parent who is a talented athlete may expose her own children to an "athletic environment" that is rich with sports gear and athletic competition. Although the child does not initially participate in creating this environment, he or she may have inherited some portion of the parent's athletic abilities. Hence, there may be a "passive correlation" between the environment and heredity. In addition, the child's behavior (which is influenced by the child's heredity) may evoke a behavioral response from the parent that in turn reinforces the child's behavior. These are known as "evocative correlations" between heredity and environment. Later in life, children can and do actively seek environments that complement their genetic tendencies, a process that has been called "niche picking" (Scarr and McCartney 1983).

What one considers "internal" or "external" depends upon what one takes the unit of analysis to be. Molecular biologists take the perspective that the genome is the appropriate unit of analysis. Relative to the genome, all else (including intra-cellular properties) constitutes the "environment." Even within the cell, many "external" factors regulate gene expression. Nearly every cell in the human body contains an identical copy of the person's DNA, which controls cellular construction and behavior. Yet, different cells are constructed and behave very differently depending on precisely where they are (what part of what tissue), what stage of development the person is at, and what is going on in the environment around the person. This is possible because the expression of protein-coding genes is modulated in response to external factors, during development and in differentiated cells.

It has become evident that causal role supposedly played by the so-called "gene" is not localizable to a stretch of DNA, but is shared by the DNA and a variety of other molecular resources. Because DNA is inert, genes depend on intracellular signals for initiating and terminating activity. Not only *when* the DNA activated but the *way* in which it is expressed depends on the nature of these intracellular signals, which vary across different cells and different phases of development. Some are naturally triggered by mechanisms outside the cell (for example, by hormones), often ultimately by influences outside the organism (for example, through grooming and other nurturing behaviors). A wide variety of normally occurring environmental and behavioral influences on gene activity have been documented in a range of species from nematodes to humans. The known environmental effects are implicated in normal variations, not just abnormalities, and collectively, they are convincing evidence that gene × environment interaction is not only possible but necessary for normal development. The causes of epigenesis are distributed among sequences of DNA, including regulatory sequences distributed throughout the genome, molecular factors extrinsic to the chromosome, their potential distal causes in the extra-cellular environment, and the contingent history of the cell. The sum of these ongoing, stochastic, bidirectional exchanges between external factors at many levels and heredity during the course of development has been called "probabilistic epigenesis" (Gottlieb 2007).

Many levels *above* the level of the organism have also been taken either as units of analysis or as levels of explanation in theories of development and learning. In roughly ascending order, these levels include the mother-child dyad; the home and nuclear family; the neighborhood (playground, friends and neighbors); the childcare, school or workplace; the community (including local services such as healthcare and children's services); the culture (including customs and values); the society (including laws and institutions); the climate; the historical timeframe; the environment; and the geological period.

There are a number of contemporary theories of learning and development that accept the classical assumption that the individual person (or organism) is the appropriate unit of analysis but reject the assumption that the appropriate level of explanation is exclusively sub-personal. On Bronfenbrenner's "bioecological model" (Bronfenbrenner and Morris 2006), for example, "proximal processes" (interactions with the immediate environment) are the "engines of development," and developmental outcomes (including learning) are joint functions of proximal processes, the characteristics of the person (in the relevant case, the learner), the environmental context both narrow and wide, and a time interval.

On a variety of "holistic" or "contextualist" theories of learning and development, the surrounding context is either or both: (a) necessary for understanding learning; or (b) part of the nature of learning and development itself. For example, Vygotsky argued that mind cannot be understood in isolation from society. Indeed, the view that the surrounding context is part of the nature of learning itself is a logical consequence of content externalism, as advocated, for example, by Hilary Putnam (1975). On these views, learning is learning a proposition or a practice (skill), either of which has meaning only in a rich socio-cultural context.

Entities larger than the individual also have lifetimes over which they may learn. The "situated cognition" movement has done productive research based on the assumption that entities such as ships and airplanes also learn (Hutchins 1995). Even if one does not want to go that far, it is common to acknowledge that learning is often distributed among group members, and it is the group that learns, not the individual.

In summary, development and learning usually result from nonlinear stochastic interactions between causes at multiple levels, both higher and lower. In particular, outcomes at the personal or organism level almost always result from a wide variety of sub-personal factors, including hereditary ones, interacting with a particular supra-personal milieu. As a consequence, outcomes are liable to change dramatically if either sub-personal or supra-personal factors change even slightly. The discovery of these facts has driven calls from various quarters for a range of changes to theory and practice. The most radical claim that sub-personal and supra-personal factors are "inseparable" and therefore urge abandoning all distinctions and embracing "postmodernism" wholeheartedly (e.g., Overton 2006). More reasonable proposals typically involve calls for a shift toward the practice of routinely using nonlinear statistical – or "systems" – models in developmental studies. As a result of these considerations, a consensus is converging that multilevel, nonlinear analyses and systems explanations are required in the developmental sciences (see the section, "Dynamic theories," below).

Modes and types of learning

As discussed in the second section, above, philosophers have traditionally emphasized just two modes of learning: explicit propositional or declarative learning ("knowing that"), and implicit practical, or skill learning ("knowing how"). The distinction between *implicit learning* and *explicit learning* has to do with whether the learner has conscious, reflective access to what has been learned. Knowledge that results from implicit learning is *tacit knowledge*, knowledge to which the learner does not have self-conscious access. They have also tended to privilege propositional learning over skill learning. Psychologists, on the other hand, distinguish many dozens of different forms of memory (Roediger et al. 2002), each one of which presumably involves a different kind of learning. Also, psychologists tend, if anything, to privilege implicit learning over explicit learning, as may be seen by the many studies of *priming* in the psychological literature.

The idea that learning is a change in behavior resulting from experience emphasizes the product of learning (the outcome, the change) as opposed to the process by which the product arises, chalked up blithely to "experience." Reliance on this conception of learning leads to a passive view of learning – that it is something that can be done for you or done to you, by a teacher for example. As many teachers know, this "commercial" view of learning (you can just go out and buy learning – it thereby becomes your rightful possession) is widespread among students and their parents. Another way of looking at learning is as a process, or better, an activity, something that a person *does* in order to understand the real world and bring meaning to life, as opposed to something that a person *has*. This view emphasizes the actions that a person takes, and the consequences they have, both internal and external, when learning takes place.

Habituation is a form of learning that results in a decrease of response to a repeated or continued stimulation. Habituation may be measured by looking time (decrease in looking time as a stimulus becomes more familiar), head turn (decrease in head orientation response as a stimulus becomes more familiar) or even heart rate (decrease in transitory heart-rate deceleration upon presentation as a stimulus becomes more familiar). Heart-rate measurements have even been used to measure habituation *in utero*.

Associative learning, broadly defined, is simply the learning of associations between one set of things and another. One of the reasons that associative learning is so popular in psychology is that, in certain forms, it can be described in strictly behavioral terms. Another is that it is consistent with the known neural mechanisms of learning. A third reason is that associative learning seems to be ubiquitous among animals.

Classical (or Pavlovian) conditioning starts with an "unconditioned" (previously associated) stimulus-response pairing, such as food and salivation. If a "conditioned" (not previously associated) stimulus, such as ringing a bell, is regularly paired with the unconditioned stimulus for a sufficient length of time, then the conditioned stimulus will become associated with the unconditioned response. The association (learning) may be demonstrated by presenting the conditioned stimulus alone (without the unconditioned stimulus) and observing that it elicits the same response as the unconditioned stimulus. Operant conditioning (or instrumental learning), by contrast, is concerned with associating new consequences (new postconditions) with preexisting, voluntary behavior in order to modify the frequency of the behavior. Rewards are, by definition, consequences that tend to increase the frequency of a behavior. Punishments are, likewise, consequences that tend to decrease the frequency of a behavior. The association (learning) is demonstrated in the change of frequency of the behavior.

In psychology, the term "statistical learning" may refer to learning based on statistical theory (including both associationism and Bayesian learning) or specifically to associative learning of the statistical relationships among items in experience. Usually, these are stimuli, but other items active in working memory (e.g., concepts retrieved from long-term memory) may also participate in statistical learning. Statistical learning has been demonstrated in nonhuman primates (cotton-top tamarins) and human adults, children and infants for a variety of types of stimuli, from sequential auditory patterns composed of either speech sounds or tones presented in rapid succession to shapes in a visual display (e.g., Saffran et al. 1999). Statistical learning of *nonadjacent* temporal dependencies (for example, among speech sounds generated by an artificial grammar, and among acoustic correlates of English phrases) has also been demonstrated in infants as young as eight months old, and on that basis, it has been hypothesized that statistical learning might play a critical role in acquiring syntax (Newport and Aslin 2004).

There is an important connection between a form of learning known in psychology as "perceptual learning" and the Churchlands' eliminative materialism. Eliminative materialism depends on the "plasticity of mind" because folk psychology can only be eliminated by inter-theoretic reduction if (1) it is a theory; (2) there is a better theory on offer; (3) the better theory on offer is reducible to a lower level theory; and (4) we can live without it. Because folk psychology is such a well-embedded part of our everyday experience, we can only live without it if we can change the way we experience everyday life. We can change the way that we experience everyday life only if the mind is sufficiently plastic that, in virtue of adopting a new theory, we can come to experience, or perceive, things differently. The Churchlands have argued forcefully from thought experiments and historical developments in science that the mind is sufficiently plastic that it is possible for us to come to perceive things differently. Perceptual learning just is the phenomenon of learning that changes the nature of our perceptions, and there is an impressive array of empirical evidence for perceptual learning (Goldstone 1998).

There are many other forms of learning. Clark Glymour, Alison Gopnik and their colleagues (Gopnik et al. 2004) have suggested that human learning of causal relationships (including the relationships between one's own actions and effects in the world) is a form of "Bayesian learning," that is, might be based on a mechanism that is capable of forming representations isomorphic to those formed in the "Bayes net" formalism for causal maps. Developmentalists agree that infants can learn by imitation (for example, to imitate facial expressions), although exactly how early they can do so remains a subject of debate. Humans often seem to learn a lot from a single example, particularly one accompanied by an explanation. This form of learning is often referred to as "learning by example." "Learning by analogy" is learning by using a mapping from one (known) domain to another (initially unknown) domain. "Imprinting" (for example in the way in which ducklings imprint on their mothers) is also often considered a form of learning in the psychological literature.

Methodological challenges

In addition to the general methodological issues in psychology at large (see "Conceptual Problems in Statistics, Testing and Experimentation" [Chapter 14]), there are several issues that are unique to developmental studies. There are many ethical and practical constraints on the study of learning and development, particularly in young human children. Because of these constraints, our knowledge about development and learning in children is on much shakier epistemological ground than is our knowledge about learning in adults, or even (in some respects) in nonhuman animals.

Although introspection is no longer widely used in psychology, self-reports (usually structured to some degree as clinical interviews or questionnaires) continue to be an important source of evidence in developmental science, particularly for older children and adolescents. Diary studies continue to be particularly important in the study of language acquisition, although they raise many important epistemological issues including investigator bias, sampling rate, and information about the context.

Systematic observation has also been widely used to study development and learning. Many developmentalists with a social orientation place particular emphasis on the importance of naturalistic observation. Thus, developmental studies become a kind of ethnography, sometimes to the degree that it is suggested that the only meaningful observations of development can happen when the child is completely unaware of the researcher. This is an interesting case of an argument against the use of experimental methods in science. One issue in any observational study is the choice of sampling strategy – whether to sample by time, or sample by event. It is virtually impossible to sample the child's environment continually, although new technologies are pushing the limits of what it is feasible (see below). The modality of sampling studies has changed with the times, from the handwritten records kept by Darwin and others through audio and video recording, with a transition to digital formats underway. Recordings are usually from the observer's point-of-view or some other ("third-person") point-of-view. Wireless lavaliere microphones are sometimes used to record a child's speech sounds more accurately.

Observational studies of all kinds are fraught with observer bias and observer influence. Both are exacerbated by the fact that the researcher is also the participant's mother or primary caretaker. If not, the caretaker is usually nearby and, naturally, wants to get involved or, intentionally or not, becomes the focus of attention. Because many observational studies also involve later coding or transcription from recorded media, a further opportunity for observer bias arises. This is usually quantified by statistical measures of agreement (e.g., Cohen's *kappa*), which can only measure the degree to which the two codings agree, not tell how much they might both be biased. In an ideal observation study, the data collection (the observations) and any further processing of the data, including coding and transcription, would all be performed by persons naïve to the purpose or hypothesis of the study. Such studies are rare, however.

Two main kinds of correlational studies are also widely used in research on development. One is cohort studies, in which groups that have something in common (for example, being in the same school district) at one point in time are studied longitudinally, and measures taken at later dates are correlated with measures taken at earlier dates. Among these, twin studies, which consist of studies of twins separated at birth, have been particularly important in attempts to untangle innate and environmental influences on development. The second kind of widely used correlational study is the corpus study. Particularly in the study of language acquisition, large corpora of observational data from many researchers have been collected and systematized (MacWhinney 2000). These corpora can then be searched for correlations, for example, between earlier parental behavior and later child behavior.

A wide variety of psychophysiological methods have been used in studies of development. Autonomic indicators such as heart rate, blood pressure, respiration, sucking rate, pupil dilation and electrical conductance of the skin are particularly useful in the youngest infants, because measurements are minimally invasive or disruptive. Indeed, heart rate is commonly used as a dependent measure in studies of *prenatal* development.

Measurements of brain activity are difficult with infants and young children, both because they find the procedures and equipment frightening and because they lack the capacity to voluntarily control their movement. Event-related potential (ERP) studies have been conducted even with neonates. However, the spatial resolution in these studies is generally lower than in adult studies. A state-of-the-art adult ERP system contains 256 electrodes, while those in common use contain 128 electrodes. Those used in infant studies typically contain a dozen or so electrodes, although infants as young as 5 months have been fitted with caps containing 128 electrodes (Reynolds and Richards 2005).

Functional magnetic resonance imaging (fMRI) is particularly challenging, because it requires the participant to remain motionless in a narrow, enclosed space for a long period, despite – among other challenges – an almost deafening roar from the equipment itself. This has limited the use of fMRI in younger children. Although sedation or general anesthesia is sometimes used for medical purposes, they are typically not justified for research and make behavioral experiments difficult or impossible. Researchers have worked around these problems by preparing children in a simulated scanner using a cover story (e.g., "we're going to ride on a spaceship") and behavioral training to reduce body motion (Kotsoni et al. 2006).

When interpreting the results of studies using psychophysiological methods, we must keep in mind that many factors influence physiological responses or their measurements, including hunger, boredom, arousal, fatigue and movement. Interpreting a physiological response requires inference. In particular, the fact that a stimulus produces a temporally or spatiotemporally consistent pattern of brain activation does not tell us how either the brain or the person processes the stimulus.

It is widely acknowledged that causal inferences are strongest when based on experimental studies, that is, on deliberate manipulations applied to individuals selected at random from a population. However, many experimental paradigms that have contributed to our current understanding of development and learning can only be used with animals. In some cases, this is because the manipulations are already ethically questionable when performed on immature nonhuman animals and would be abhorrent if performed on human children. These include a number of procedures that have yielded invaluable scientific information about learning and development. Among them are single-cell recording from awake, behaving animals (which typically requires open-scalp surgery) and a variety of environmental manipulations (including induced genetic abnormalities or complete deprivation of mobility or perceptual experience in one or more modalities until adulthood). In these cases, we must rely on inferences from the animal studies.

There are, of course, many experimental manipulations that can be performed at minimal risk to the well-being of the child. The younger the child is, however, the more difficult it is to design an experimental procedure, because the less ability infants have to communicate their experiences through coordinated movement. Early experimental studies of development, such as Piaget's, often relied on overt behavioral skills that required fine motor coordination (such as controlled reaching and searching). Many more recent procedures depend on measurements of looking time – the amount of time that an infant spends looking at one thing (or in one direction) rather than another. These include habituation and familiarization procedures, the "violation of expectancy" procedure, the "head-turn preference procedure" and the "intermodal preferential looking paradigm."

Habituation or familiarization to a repeated or continued stimulus may be measured by looking time, by head turn, by sucking rate, or even by heart rate. When a significant difference in one of these measures can be demonstrated, it is logical to conclude that the infant can discriminate the habituated stimulus from the novel one. On the other hand, when a significant difference in one of these measures is not demonstrated, it is not logical to conclude that the infant cannot discriminate the stimuli, for all the same reasons that we cannot generally draw conclusions from an experimental failure to find a significant effect. However, that conclusion is often drawn in the literature, and some important findings, such as estimates of infants' visual acuity, depend upon it.

Moreover, looking-time and head-turn measures are often interpreted as a "preference" for the stimulus toward which the eyes (or head) are directed longer or more frequently. Looking-time studies usually show a "preference" for novel stimuli. On the other hand, contingent sucking patterns (e.g., when an infant's sucking pattern determines which of two sounds an infant will hear) have often shown a "preference" for familiar stimuli, that is, a tendency to suck in the pattern that causes the familiar stimulus to continue.

The measurement in the violation of expectancy procedure is looking time. When an effect is found, it is a statistically significant difference between looking times in one condition (the "expectancy" condition) and another (the "violation" condition). In the head-turn preference procedure, an infant's attention to an auditory stimulus is examined by measuring the amount of time the infant's head orients toward the direction of the stimulus. This depends on the fact that infants tend to orient visually (and thus turn their heads) in the direction of a sound source. By being permitted to hear the sound only as long as the head turn is maintained, infants may be trained to maintain the response, increasing the sensitivity of the measure. In the intermodal preferential looking procedure, infants are simultaneously presented with a linguistic (or other auditory) stimulus and two video events, only one of which matches the linguistic stimulus.

In all of these procedures, interpreting a difference in looking time as something else (even as a violation of expectancy, a surprise or a preference, but especially as an indication of an inference, a belief or a theory) requires a sound argument connecting looking time with the language used in its interpretation. It is rare to see such an argument, but it is easy to slip into interpretations of looking-time differences as evidence for rich cognitive structures. Many published papers make such errors. Although a bias toward rich interpretations is enshrined in the very names of many of the procedures, it is very difficult to know whether a head-turn or looking-time measure is an indication of a "preference" for familiarity, a preference for novelty, a violation of expectancy, or any other mental state. Recently, Schöner and Thelen (2006) proposed a dynamic field theory (DFT) model that qualitatively fits many of the looking-time patterns reported in the literature based strictly on perceptual representations, without assuming any higher level cognitive states or processes. The model makes many simplifying assumptions and also leads to a number of empirical predictions that have yet to be tested, but it is suggestive.

Dynamic theories

Related to issues about units of analysis and levels of explanation are issues about the dynamics of learning, including time scales, continuity and stability. Focusing on the personal level, we tend to consider learning changes that occur on time scales from minutes to weeks and developmental changes that occur over the lifetime or within a life period measured in years. When we expand our focus to supra-personal levels, then we tend to consider longer time periods (decades for cultural and community developments, centuries for historical developments, and longer periods for evolutionary, climate and geological developments). Similarly, when we narrow our focus to sub-personal levels, then we tend to consider smaller time scales, from behaviors (head turns, looks) that may be measured in tenths of a second to broad changes in neural activity, which may be measured in 100ths of a second, down to milliseconds (action potentials).

One of the key theoretical issues in the study of development has been whether development is best regarded as a continuous process of quantitative change or whether it consists of discontinuous states that are qualitatively different. The classic stage theories posit sequences of qualitatively different periods in development. However, these theories have always been plagued by findings of vast individual differences. This is another force that has led toward the convergence on dynamical models of development – in a nonlinear system, qualitatively different behavior can arise from incremental, qualitative change.

A related issue is if it is appropriate to "discretize" time and other quantities in the study of development, a question that leads to debates about the nature of explanation and representation, levels of analysis and time scales. Information-processing theories of development and learning, including connectionist models, implicitly assume that there is no harm in considering only discrete intervals of time and other quantities (such as synaptic weighting and activation in connectionist models). Dynamical theories instead assume that time and other quantities are continuous, in the sense that an infinitesimally small difference in the state of the system could make a difference to its future behavior. This view receives some support from neuroscience, in that (continuous) spike-timing models of the neuron have been shown to be empirically more adequate than (discrete) firing-rate models, including in learning (spike-timing-dependent synaptic plasticity; Roberts and Bell 2002). However, it seems unlikely to make a difference in practice, given that digital computers built on top of analog circuits are used to model both discrete and continuous models of development and learning.

Another issue is whether developmental systems are inherently stable or dynamic. Information-processing models, including connectionism, take the natural state of the system to be "at rest." Input to the system is not destabilizing, but merely brings it directly to another stable state; nor is the lack of input destabilizing, but merely allows the system to continue in the current state. By contrast, dynamical models take the natural state of the system to be continuous change. Any apparent stability is hard won and precarious. Input to the system destabilizes it, and settling into a new equilibrium is an achievement. Information processing theorists generally interpret their models as abstractions of critical processes that are embedded in real organisms. They do not model the dynamics because they do not see the necessity of doing so for explaining the phenomena of interest. Detailed dynamical models of developmental phenomena (such as those built in the DFT framework), by contrast, model the dynamics themselves because the theorists who use them believe that they are relevant to the phenomena of interest. This may be part of the reason that DFT models have primarily been used for modeling motor behavior (such as reaching and perseveration) rather than for modeling higher level cognitive processes (such as decision-making and learning).

Embodiment

Does the body play a necessary role in learning, or merely an accidental one? The fact that there are so many formal theories of learning suggests that embodiment is not necessary in principle. However, abstract theories of learning may be parasitic on actual human learning. Indeed, the argument is well-known in philosophy that meaning depends on supra-personal states of affairs (Putnam 1975). It follows that learning the meaning of anything depends on those same states of affairs. One might argue that the sociocultural aspects of meaning depend on the capacity to participate in a society or culture, which in turn requires embodiment. If that were true, then learning the meaning of anything would require embodiment. It is not clear why that would have to be true, however. Certainly, there are many disembodied societies, including those involved in "massively multiplayer" online role-playing games and those involved in the bargaining negotiations conducted by software agents, although these may also be parasitic on real societies.

Even if embodiment is not necessary for learning in general, it may still be critical for actual human learning, in that learning mechanisms may be specialized for operating "within" a body. In a certain sense, that is trivially true, because the mechanisms that mediate learning among animals include nervous systems, which are parts of bodies. It may also be true in the more interesting sense that the neural mechanisms that mediate learning in different organisms may be specialized to the kinds of bodies they have. This is again true in a trivial sense – a nervous system is (directly) connected to sensory receptors and motor effectors in the body that it controls, not in any other body. But this latter trivial sense of embodiment also entails a deeper one – the modifications to the nervous system that comprise learning in any organism are specific to the environmental inputs and behavioral outputs that body is capable of receiving and performing. We know that we can adapt to changes to our bodies, including changes that happen during development, paralysis and the loss of limbs, as well as virtual reality environments and such things as remote piloting (Clark 2003). This may suggest that learning is disembodied – that what body you have doesn't make all that much of a difference – but it also suggests that the opportunities we have for learning are determined by the kinds of bodies we have, biological or otherwise.

The debate about embodiment has pragmatic implications. Distance learning, computer-based training, and "open education" in general depend upon the possibility of disembodied learning. The more effective forms tend to reduce the perceived distance between the learner's body and the educator's. This can happen by bringing the educator's body to the learner (for example, through multi-modal multimedia), by bringing the learner's body to the educator (for example, by providing the means for the learner to "raise his hand" and ask a question virtually), by turning the computer into the educator (for example, through interactive learning sessions that involve pressing keys or typing), or by turning the learner into the educator (for example, by motivating the learner to educate his or her self by finding meaning in what he or she learns). The most effective practices generally combine all four methods.

Looking forward

Because development and learning involve so many different phenomena, at different levels of explanation and different physical and temporal scales, with reciprocal influences within and between levels, interdisciplinary research is essential and is becoming more and more common. Both philosophical and psychological studies of development and learning suffer from an excessive reliance on research involving participants sampled from a small, culturally and socioeconomically homogenous, portion of the world's population. Broadening the pool of research participants to include a representative cross sample of the diverse world population is a major challenge that psychologists have only begun to tackle.

Cognitive developmental neuroscience (Nelson et al. 2006) is an emerging sub-discipline that uses primarily physiological methods to study cognitive development from before birth through adulthood. This field will continue to push the spatial and temporal limits of what we know about development and learning.

Adequately modeling nonlinear relationships, particularly among nested levels with reciprocal causation, requires sophisticated statistical techniques. To some extent, the techniques have already been developed (Goldstein 2003). However, many statistical software packages do not provide easy means of using the full range of hierarchical nonlinear models. Partially as a consequence of the limited availability of appropriate tools (but also partly as a consequence of limitations of training or habit), researchers tend to use linear models even in circumstances where they are clearly not appropriate. Researchers need to learn and teach their students the principles of multilevel nonlinear modeling, demand that providers of statistical software make them easy to use, and incorporate this knowledge and practice into the writing and review of research papers.

New technology is also enabling new forms of research. The recent widespread availability of high-quality digital audio- and video-recorders has vastly increased the amount of and dramatically changed the character of observational data that can be collected. Methods have been developed to perform relatively accurate eye-tracking, even with young infants (Aslin and McMurray 2004). A technique has recently been developed to mount a miniature video camera on a young child's head, providing a first-person point of view. And a study is underway at MIT (Roy et al. 2006) in which every room of a young child's home is video- and audio-recorded 16 hours a day!

References

Aslin, R. N., and McMurray, B. (2004) "Automated Corneal-Reflection Eye Tracking in Infancy: Methodological Developments and Applications to Cognition," *Infancy* 6: 155–63.

Augustine (1955 [397]) *Confessions*, Philadelphia: Westminster Press.

Baldwin, J. M. (1906 [1895]) *Mental Development in the Child and in the Race*, New York: Macmillan.

Bandura, A. (1986) *Social Foundations of Thought and Action*, Englewood Cliffs, NJ: Prentice-Hall.

Bowlby, J. (1982 [1969]) *Attachment and Loss*, vol. 1: *Attachment*, New York: Basic Books.

Bronfenbrenner, U., and Morris, P. A. (2006) "The Bioecological Model of Human Development," in R. M. Lerner (ed.), *Handbook of Child Psychology*, 6th edn, Hoboken, NJ: Wiley, pp. 793–828.

Case, R. (1992) *The Mind's Staircase*, Hillsdale, NJ: Erlbaum.

Clark, A. (2003) *Natural-Born Cyborgs*, Oxford: Oxford University Press.

Darwin, C. (1877) "A Biographical Sketch of an Infant," *Mind* 2: 285–94.

Dreyfus, H. L., and Dreyfus, S. E. (1986) *Mind over Machine*, New York: Free Press.

Estes, W. K. (1950) "Toward a Statistical Theory of Learning," *Psychological Review* 57: 94–107.

Fodor, J. A. (1980) "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology," *Behavioral and Brain Sciences* 3: 63–109.

Freud, S. (1949 [1905]) *Three Essays on the Theory of Sexuality*, London: Imago.

Gesell, A. (1925) *The Mental Growth of the Preschool Child*, New York: Macmillan.

Gold, E. M. (1967) "Language Identification in the Limit," *Information and Control* 10: 447–74.

Goldstein, H. (2003) *Multilevel Statistical Methods*, London: Hodder Arnold.

Goldstone, R. L. (1998) "Perceptual Learning," *Annual Review of Psychology* 49: 585–612.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., and Danks, D. (2004) "A Theory of Causal Learning in Children: Causal Maps and Bayes Nets," *Psychological Review* 111: 3–32.

Gottlieb, G. (2007) "Probabilistic Epigenesis," *Developmental Science* 10: 1–11.

Hall, G. S. (1904) *Adolescence*, New York: Appleton.

Hamlyn, D. W. (1995) "Learning," in T. Honderich (ed.), *The Oxford Companion to Philosophy*, New York: Oxford University Press, 496.

Hebb, D. O. (1949) *The Organization of Behavior: A Neuropsychological Theory*, New York: Wiley.

Hutchins, E. (1995) *Cognition in the Wild*, Cambridge, MA: MIT Press.

Kim, J. (1982) "Psychophysical Supervenience," *Philosophical Studies* 41: 51–70.

Kotsoni, E., Byrd, D., and Casey, B. J. (2006) "Special Considerations for Functional Magnetic Resonance Imaging of Pediatric Populations," *Journal of Magnetic Resonance Imaging* 23: 877–86.

Lerner, R. M. (ed.) (2006) *Handbook of Child Psychology*, vol. 1: *Theoretical Models of Human Development*, Hoboken, NJ: John Wiley & Sons.

Lewis, M., Goldberg, S., and Cambell, H. (1969) "A Developmental Study of Information Processing within the First Three Years of Life: Response Decrement to a Redundant Signal," *Monographs of the Society for Research in Child Development* 34, no. 9: 1–41.

MacWhinney, B. (2000) *The CHILDES Project: Tools for Analyzing Talk*, Mahwah, NJ: Erlbaum.

MacWhinney, B., Leinbach, J., Taraban, R. M., and McDonald, J. L. (1989) "Language Learning: Cues or Rules?" *Journal of Memory and Language* 28: 255–77.

Malenka, R. C., and Bear, M. F. (2004) "LTP and LTD: An Embarrassment of Riches," *Neuron* 44: 5–21.

Mitchell, T. M. (1997) *Machine Learning,* New York: McGraw-Hill.

Nelson, C. A., De Haan, M., and Thomas, K. M. (2006) *Neuroscience of Cognitive Development,* Hoboken, NJ: Wiley.

Newport, E. L., and Aslin, R. N. (2004) "Learning at a Distance," pt 1: "Statistical Learning of Non-Adjacent Dependencies," *Cognitive Psychology* 48: 127–62.

Overton, W. F. (2006) "Developmental Psychology: Philosophy, Concepts, Methodology," in R. M. Lerner (ed.), *Handbook of Child Psychology*, 6th edn, Hoboken, NJ: Wiley, 18–88.

Pavlov, I. P. (1927) *Conditioned Reflexes*, London: Oxford University Press.

Piaget, J. (1999 [1954]) *The Construction of Reality in the Child,* London: Routledge.

Plomin, R. (1986) *Development, Genetics, and Psychology,* Hillsdale, NJ: Erlbaum.

Plomin, R., Defries, J. C., and Loehlin, J. C. (1977) "Genotype-Environment Interaction and Correlation in the Analysis of Human Behavior," *Psychological Bulletin* 84: 309–22.

Putnam, H. (1975) "The Meaning of 'Meaning'," in K. Gunderson (ed.), *Language, Mind, and Knowledge,* Minnesota Studies in Philosophy of Science, vol. 7. Minneapolis: University of Minnesota Press, pp. 131–93.

Reynolds, G. D., and Richards, J. E. (2005) "Familiarization, Attention and Recognition Memory in Infancy: An Event-Related Potential and Cortical Source Localization Study," *Developmental Psychology* 41: 598–615.

Roberts, P. D., and Bell, C. C. (2002) "Spike-Timing Dependent Synaptic Plasticity in Biological Systems," *Biological Cybernetics* 87: 392–403.

Roediger, H. L., Marsh, E. J., and Lee, S. C. (2002) "Kinds of Memory," in H. Pashler and D. Medin (eds), *Steven's Handbook of Experimental Psychology,* 3rd edn, Hoboken, NJ: Wiley, pp. 1–41.

Rogoff, B. (2003) *The Cultural Nature of Human Development,* New York: Oxford University Press.

Roy, D., Patel, R., Decamp, P., Kubat, R., Fleischman, M., Roy, B., Mavridis, N., Tellex, S., Salata, A., Gunness, J., Levit, M., and Gorniak, P. (2006) "The Human Speechome Project," in R. Sun and N. Miyake (eds), *Proceedings of the Twenty-Eighth Annual Meeting of the Cognitive Science Society,* Mahwah, NJ: Erlbaum, pp. 192–6.

Russell, S., and Norvig, P. (2003) *Artificial Intelligence: A Modern Approach,* Englewood Cliffs, NJ: Prentice Hall.

Ryle, G. (1949) *The Concept of Mind,* New York: Barnes & Noble.

Saffran, J. R., Johnson, E. K., Aslin, R. N., and Newport, E. L. (1999) "Statistical Learning of Tone Sequences by Human Infants and Adults," *Cognition,* 70: 27–52.

Scarr, S., and McCartney, K. (1983) "How People Make Their Own Environments: A Theory of Genotype→Environment Effects," *Child Development* 54: 424–35.

Schöner, G., and Thelen, E. (2006) "Using Dynamic Field Theory to Rethink Infant Habituation," *Psychological Review* 113: 273–99.

Skinner, B. F. (1935) "Two Types of Conditioned Reflex and a Pseudo-Type," *Journal of General Psychology* 12: 66–77.

Thelen, E., and Smith, L. B. (1994) *A Dynamic Systems Approach to the Development of Cognition and Action*, Cambridge, MA: MIT Press.
Valiant, L. G. (1984) "A Theory of the Learnable," *Communications of the* ACM 27: 1134–42.
Vapnik, V. N. (1999) *The Nature of Statistical Learning Theory*, New York: Springer.
Vygotsky, L. S. (1986 [1934]) *Thought and Language*, Cambridge, MA: MIT Press.