**Chapter 1**

 **The Science of the Mind**

The Science of the Mind

Historical Approaches to the Mind

Philosophy

 The Mind-Body Problem

 Perspectives on Consciousness

Insights from Philosophy: The Black and White Room

 The Scientific Method

Neuroscience

 Cutting Edge Brain Research: Normal People with Abnormal Brains

Psychology

Linguistics

Anthropology

Education

Artificial Intelligence

Innovations in Computing: A Worthy Competitor

Contemporary Cognitive Science

Social and Ethical Questions on the Horizon

Summary

 Review Questions

 Recommended Resources

Learning Objectives

* Discover the complexity identified in contemporary views of the mind
* Understand the historical roots of modern scholarly approaches to the mind
* Explain how cognitive science emerged as a field of study and evolved over the last 70 years
* Think critically about how thinking machines are affecting society and individuals and what future effects may occur

**The Science of the Mind**

In all of science, the mind is one of the most poorly understood phenomena. Scholars continue to debate what it is. If we were asked to define it, we might find it to be one of those familiar concepts that is difficult to pin down in words. In daily conversation, we routinely refer to the word, as in “something slipped my mind,” “I’ve made up my mind,” or our friends are “losing their minds.” We hear the word in popular music, as in Willie Nelson’s song *Always on My Mind,* in the Pixies’ song *Where is my mind,* and in Oasis’s lyric *slip inside the eye of your mind.* These uses of the word *mind* hint at the complexity that researchers explore formal academic studies of the human mind. The original meaning of the word mind was *memory* in Old English with an earlier root meaning *to think* (Mind, 2020)*.* The Oxford English Dictionary defines it as the aspect of people that enables them to think, to feel, to be aware of themselves, and to experience consciousness (OED, 2020). This definition aptly describes the range of abilities that scholars recognize as involving the human mind.

Cognitive science is the broad interdisciplinary field focused on understanding the mind. In the 1970s, scholars from six disciplines were recognized to be generating relevant research. These disciplines were philosophy, psychology, neuroscience, linguistics, anthropology, and computer science, specifically the subarea of **artificial intelligence** (AI), which aims to create machines with human-like abilities (Sloan, 1978). Education came to be included as a 7th core discipline in the field (Cognitive Science Society, n.d.). Each discipline approaches the study of the mind differently, working within the scholarly norms that have developed over time. Table 1.1 displays the seven disciplines with a brief description. In this book, we examine current and past research on the science of the mind, the interdisciplinary field of cognitive science, covering numerous topics, including, but not limited to, learning, knowledge, memory, language, decision-making, problem-solving, general intelligence, minds in non-human species, and the extent to which AI can simulate the full range of human abilities.

Over the last few decades, we have become increasingly aware of AI, as there has been an increase in the depictions of thinking machines in TV and film. At the same time, we have grown familiar with commercial products, such as the talking online assistants Siri, Alexa, and Google Assistant. Few would mistake these AIs for humans. These examples of AI are categorized as **weak AI**, as they simulate a narrow range of human-like behavior (Bermúdez, 2014). We find examples of **strong AI**, which refers to innovations that are capable of simulating the full range of human-like behavior in machine that also have a human-like mind and consciousness, only in works of fiction. The earliest known example is in the novel Erehwon, which was published in the 19th century by the English writer Samuel Butler (1872). The protagonist explores a foreign country, finding things to be different from 19th century England. Three chapters focus on machines, which have evolved to have consciousness through a process described as similar to Darwin’s natural selection. Since then, numerous other examples of strong AI have captured our imaginations. Table 1 lists selected examples from TV and film.

[Insert Table 1.1 about here]

Only after we understand the full nature of the mind in humans will it be possible to create artificial minds that are capable of the full range of human-like decision-making, behavior, and, perhaps even, conscious awareness. It may be surprising to learn that many science and technology leaders agree that it is just a matter of time before we have strong AIs existing among us*.* To them, it is not a matter of *if* strong AIs will be possible*,* but *when* will strong AIs be created*.* These science and technology leaders agree on a second point. They predict that strong AIs will pose an existential threat to humankind. Stephen Hawking (1942 - 2018) expressed this view before his death. His opinion may have been informed by his experience with the advances in technology that enabled him to continue to communicate with a visual keyboard operated with his eye movements after he had lost his ability to control most of his muscles due to the degenerative condition amyotrophic lateral sclerosis (ALS, view before his death, Rutschman, n.d.). Elon Musk shared the dire prediction about strong AI, with his view being informed by his experience with AI applications for his Tesla automobiles (Shead, 2020). Microsoft founder Bill Gates has also acknowledged the potential threats of AI, comparing it to nuclear energy as there are ways in which humankind can benefit while there are also potential dangers (Clifford, 2019). These perspectives on strong AI are similar to those expressed in 1950 by the father of artificial intelligence himself, Alan Turing (1912 - 1954) (Levesque, 2017). In recent years, hundreds of prominent scholars and IT professionals have begun to raise concerns and to call for governments to consider regulation on AI technologies (Etzioni & Etzioni, 2017).

If humankind is to survive, we must understand as completely as possible the nature of the human mind, the extent to which other species have a form of consciousness comparable to the human mind, and whether minds can exist in machines. This book aims to provide an introduction to the science of the mind that includes important contributions from all seven core disciplines. In this chapter, we begin with historical approaches to the mind, including the distinct histories of each of the seven disciplines of cognitive science. Second, we examine the formal beginnings of cognitive science and how it has evolved over the last 70 years. In the last section of this chapter, we consider some of the ways in which our increasing reliance on thinking machines are already affecting individuals and society and how these changes are expected to continue in the future.

**Historical Approaches to the Mind**

As we review the notable historical origins of the study of the mind, we begin with a review of contributions from philosophy, which can be regarded as the first scholarly tradition. We then examine some of the important events in the science of the mind, starting in the 19th century, when increasing levels of scholarly activity laid a foundation for the numerous academic disciplines that become institutionalized in colleges and universities in the 20th century. We begin our review of the 19th century events with those from neuroscience and psychology.

**Philosophy**

 Among the world’s oldest surviving documents are those written by the ancient philosophers. The word philosophy contains the Greek word *philosophia* meaning *love of wisdom* (OED, 2020). The origins of the academic discipline of philosophy are traditionally viewed as ancient Greek. The work of philosophers from ancient Greek provided the foundation for Western philosophy (Russell, 2009). The most familiar of these are Socrates, Plato, and Aristotle. However, in other parts of the world, there were other philosophers and other philosophical traditions. In this section, we will examine three major themes in the history of philosophy that are relevant to cognitive science. The first is the debate about whether the mind and the body are separate entities, which is traditionally referred to as **the mind-body problem**. The second is perspectives on consciousness and conscious experience. The last is the scientific method, which began as a philosophical approach to gaining knowledge and ultimately formed the foundation for future scientific disciplines.

**The Mind-Body Problem**

The question of whether the mind is an entity distinct from the body was among the questions considered by the ancient Greek philosophers. Plato’s view was that the mind (or soul) is separate from the body, able to exist following death and having existence prior to birth. It was a view shared by his teacher Socrates (Nails, 2020). The view that the mind and body are distinct entities is referred to as **dualism** (Robinson, 2020). Centuries later, French philosopher René Descartes (1596 - 1650) also argued for a form of dualism (often referred to as Cartesian dualism) that provides a foundation for modern sciences, including physics, but also was compatible with the religious view of the eternal soul (Dicker, 2013). Descartes claimed that the mind and body were distinct entities and that the mind was immaterial, not adhering to the laws of nature (i.e., not measurable). In accordance with this view, the mind is a metaphysical entity.

Dualism continues to be embraced in contemporary society by scholars as well as the general public. The most prominent dualist philosopher today may be the Australian philosopher David Chalmers (1996) who believes that mental states, such as those occurring as subjective experience of events, and the physical states of the body are different properties. When we consider the general public, we find millions, perhaps more, dualists, who believe that the soul exists after death, a view shared by numerous world religions. Recent surveys have found that many professionals working in technical fields, including the healthcare professions, hold dualist views (Demertzi et al., 2009; Reggia et al., 2015).

Many scholars, including some in ancient Greece, rejected dualism in favor of **monism**, the view that the mind is a product of the body or bodily processes. This view reflects the belief that all things in the universe are composed of one kind of substance (Schaffer, 2020). The two most notable monists during this era were Democritus (460 BC - 370 BC), a contemporary of Socrates, and Aristotle (384 – 322 BC). Democritus proposed that all things are composed of parts or atoms (Berryman, 2016). Democritus’s monism is similar to the modern view of **materialism**, which views the mind to be a product of the body (Lange, 2016). Aristotle described the mind-body relationship using the analogy of a piece of clay, which can be made to have different shapes (e.g., round, flat, bowl-shaped, etc.) (Gill, 1989). The changeable states of the mind reflect different physical states of the body. Contemporary monists, such as the American philosopher Daniel Dennett (1991), most scientists, and the scientifically oriented public, endorse **reductionism**, which is the view that phenomena can be understood through the analysis of their parts (van Riel & Van Gulick, 2019). The reductionist solution to the mind-body problem is that the mind and consciousness can be understood through understanding the body, the brain, and its functioning.

We can find both dualist and monist perspectives of the mind-body problem within the Eastern philosophies, which developed in different regions of Asia, particularly India, China, and Japan, some predating Western philosophy (Kalupahana, 1992). For example, Adi Shankarachara, who lived around the 8th century, advocated a monist view in the Advaita Vedanta branch of Hindu philosophy from India (Jayatilleke, 2013). In contrast, the Buddha (480 - 400 BC) who founded the Buddhist religion that has flourished throughout Asia described the mind as distinct from the body (Karunamuni, 2015). Furthermore, the mind was viewed as composed of five aspects: a) material form; b) sensory consciousness; c) perception; d) feelings, and e) volition (Karunamuni, 2015). The first factor, material form, refers to matter, which includes the matter making up the body as well as matter external to the body. The other four factors are subjective experiences. The mind is also described as changing moment-to-moment, as in the familiar metaphor *stream of consciousness.* One of the most intriguing aspects of the Buddhist view of the mind is the suggestion that the individual’s sense of self is illusory, not corresponding to anything real, but arising from the observation of the stream generated by the five factors of the mind. This belief is referred to as the doctrine of the *not-self*, which has been interpreted as claiming that there is not a permanent self. Scholars continue to debate whether the doctrine of the not-self was intended to suggest that the self does not exist or whether the not-self is an aspect of self (Harvey, 1995).

[Insert Photo 1.1 The Oldest Known Buddhist Statue here]

**Perspectives on Consciousness**

Growing out of the consideration of the mind-body problem is a large scholarly literature on the nature of human consciousness, which focuses on the nature of our sense of self from moment to moment in our waking experiences (see Liu & Perry, 2014, for review). In the 17th century, Descartes drew attention to the self with his quote *I think therefore I am* and referred to *pensée* (or thought) as the entirety of our conscious mind (Van Gulick, 2018). Since then, multiple aspects of consciousness have attracted interest, including a) wakefulness, which recognizes different levels of consciousness (e.g., being awake vs. being in a coma); b) sentience, which refers to sensing and interacting with the world; and c) self-consciousness, which emphasizes the fact that consciousness includes being aware that one is aware.

 In an influential paper titled *What it Feels Like to be a Bat*, Thomas Nagel (1974) suggested that organisms can be described as having consciousness if and only if there is “something that it is like to be that organism” (p. 166). He asserts that such a definition of consciousness means that consciousness exists in not only humans, but also other species. By describing the problem with the example of a bat, which navigates the world using a sophisticated navigational system of echolocation, he states that a human who experiences the mental state of a being a bat would not have the physical body of a bat or past experiences of being a bat; thus, the experience would differ from the mental state of a bat. Many have criticized Nagel’s (1974) assertions and conclusions (Dennett, 1991; Hacker, 2002). Nevertheless, the example continues to be useful as it makes clear that the monist view that all experience can be understood in terms of physical properties and processes requires understanding of both the physical and the subjective.

Chalmers (1995) distinguished the easy problem of consciousness, which refers to understanding our abilities to perceive what we see and our abilities to comprehend what others say, from the hard problem of consciousness, which refers to understanding why there is a subjective feeling of experience. The term **qualia** is used to refer to the sensations occurring during direct experience (Tye, 2018). Dennett (1991) has described qualia as being inherently private, unable to be compared with the qualia of others, and impossible to describe. Some examples of experiences in which the qualia occur include the misery of a painful headache, the delight of eating a delicious strawberry, and the perception of a color. Box 1.2 describes Jackson’s (1982) compelling problem exploring the extent to which qualia are a special type of knowledge is referred to as *Mary in the Black and White Room.*

[Insert Box 1.2 Insights from Philosophy: *Mary in the Black and White Room* about here]

 Dennett (1991) has argued for the monist position that consciousness does not exist as something distinct from physical processes. Rather, consciousness is a “bag of tricks,” carried out by physical processes of the brain resulting in the subjective experience of consciousness, which seems to provide a complete and error-free reflection of reality (Dennett, 2003); however, this is an illusion. There are now numerous psychological studies of perception that demonstrate the reality that we experience in consciousness is not always what has occurred in the physical world. For example, numerous laboratory and field experiments have shown that people can be completely unaware when there are large changes made to what they are looking at (pictures as well as physical surroundings). The phenomenon is referred to as **change blindness** (Rensink, 2002; 2018).

**The Scientific Method**

One of the world changing innovations to come from the philosophical tradition is the scientific method (Andersen & Hepburn, 2020). The scientific method is a process of proposing possible explanations of phenomenon and testing predictions against observations (Carey, 2020). The roots of scientific method date back to approximately the 1500s in Western Europe as a way of establishing new knowledge, going beyond other ways of knowing, such as relying on authority figures or intuition (Okasha, 2016). Over the next 250 years, the method attracted attention and came to be adopted by more and more practitioners. Some may find it surprising that the scientific method is simply a process that involves a series of steps used to gain knowledge about a phenomenon of interest. The first step is observing the phenomenon of interest. Through observation, one may develop a possible explanation for why it occurs as it does. The second step is to state the explanation precisely enough that one can generate a testable prediction. Third, one would make observations of the phenomenon that enable one to determine whether the specific prediction was confirmed or disconfirmed. If confirmed, then the explanatory description could be regarded as accurate (i.e., true about the phenomenon). Most importantly, if the prediction was disconfirmed, then one would go back to the second step, revise the explanation, generate a new prediction, and then make new observations to determine if the new prediction occurred. The prediction that is tested is referred to a **research hypothesis**. In contrast, a **research theory** is a comprehensive description of a phenomenon that is developed through the testing of multiple research hypotheses.

Some phenomena are more easily studied using the scientific method than others. First, the phenomenon of interest must be repeatable, ideally occurring multiple times in the future and providing opportunities for predictions about the phenomenon to be tested. Second, the phenomenon of interest should be measurable, ideally being able to be quantified such that low numbers reflect less of the phenomenon or lower frequency of the phenomenon and higher numbers reflect more of it or greater frequency of it. Lastly, multiple people must be able to reach a consensus about whether the phenomenon occurred as described and about whether the measurements of the phenomenon are sound. Phenomena that are most often considered impossible to study with the scientific method are those that pose challenges for one or more of these characteristics. For example, historical events are difficult to study scientifically, as they occur just once. It is true that one can overcome this challenge by using computer-based models of an event to test predictions about how events might have unfolded under different sets of conditions. An example is the sinking of the Titanic. In her dissertation Jennifer Hooper McCarty (See McCarty & Foecke, 2008) used a computer model to test and to confirm the idea that the use of low quality rivets contributed to the rapid sinking of the ship. With stronger rivets, the ship would have taken longer to sink.

A second example of a phenomenon typically considered impossible to study using the scientific method is clairvoyance or the ability to have knowledge about people, objects or events without direct sensory contact (also called extrasensory perception or ESP). The phenomenon is repeated, as there are many people who claim the power of clairvoyance and may perform feats of clairvoyance for a fee. The challenge of studying clairvoyance using the scientific method is finding a way to measure it such that most researchers would agree that the assumptions and measurements are sound. In fact, the scientific testing of clairvoyance began in the early part of the 20th century with the invention of Zener cards (Rhine et al., 1940), a deck of 25 cards. Cards depicted one of five symbols (i.e., circle, square, wavy lines, cross, or star). When tested with a well-shuffled deck of Zener cards, a person who scores significantly above chance (i.e., 1/5 on each draw, assuming cards are replaced, equally 20%) consistently would be described as passing the test, earning the label of a clairvoyant.

When we evaluate the six disciplines of cognitive science in terms of how often the scientific method is used, we find differences. Some contemporary philosophers employ methods from experimental philosophy in which research participants are asked to read vignettes and to make judgments about them (Knobe & Nichols, 2017). Participants’ judgments tend to be described as intuitions. Most philosophers continue to employ traditional methods, which involve formulating compelling questions and proposing arguments to answer those questions (Akiba, 2020). Thought experiments, such as the Black and White Room (Jackson, 1982), are routinely proposed and answers critically analyzed. In neuroscience and psychology, the majority of research in the field utilizes the scientific method and involves the testing of hypotheses using empirical methods and statistics to evaluate how likely the sample is to have been determined by randomness or due to variable(s) of interest. In the fields of linguistics and anthropology, the scientific method is used to test theories, but data are not always subjected to statistical evaluation. In philosophy and computer science, the scientific method is used least frequently. In computer science, the majority of the research is of an applied nature, aimed at creating computer applications as well as hardware of different types, achieving different goals (e.g., performing faster and more efficiently). Some computer scientists employ methods of experimental computer science that involve developing and testing theories of performance of computational models (Dubois, 2018).

Popular impressions of different disciplines are often misleading. For example, it is common to hear some disciplines described as “hard” sciences (e.g., physics, chemistry, and biology) and others described as “soft sciences (e.g., psychology, sociology, and anthropology). This distinction may lead one to believe that they use different scientific methods, one that is superior to the other. Some might assume that the quality of the research process and/or the research discoveries may differ with “hard sciences” being superior. In fact, the different scholarly disciplines use the same scientific method. The only thing that differs across the disciplines is the phenomena that the discipline attempts to explain. Within each discipline there are scholars working on different types of research questions. Research that is focused on addressing a practical problem is categorized as **applied** **research**. Examples of applied research include what the best way to reduce depression and anxiety symptoms in college students, which police line-up testing procedures lead to the most reliable eye-witness identifications of suspects, and do online shoppers prefer human customer service employees over AI applications? Other research that investigates the fundamental nature of phenomena (e.g., its composition or functioning) is categorized as **basic** **research**. Examples of basic research include how do people how is color perceived, what happens in the brain when one performs a math problem, and for bilinguals, how are their two languages stored in memory?

In each of the remaining chapters of this book, we will learn about important contributions from philosophy, which are related to the topic of the chapter. These will be presented as a text box feature. Many of these examples are famous thought experiments considered in philosophy, such as the example of the Black and White room in this chapter. Some chapters, including the following chapter, we will learn about ideas from philosophy that have influenced psychological approaches to the mind.

**Neuroscience**

 The discipline of neuroscience approaches the understanding of the mind from the tradition of monism, assuming that achieving a complete understanding of the body and bodily processes would ultimately lead to an understanding of all mental experiences. Modern neuroscience shares much of its early history and prehistory with the medical sciences (Finger, 2004). References to neurological disorders can be found in the Edwin Smith Papyrus, an Egyptian document that is believed to have been used to train physicians dated to the 17th century BC (Feldman & Goodrich, 1999). One case describes a man with an injury to the temporal bone that affected his ability to speak. It was described as a condition that was not curable. Much was learned about the body through the examination and dissection of the dead, a practice known to have occurred in ancient Greece as part of medical training (Ghosh, 2015) continuing in the following centuries to today. By the 1700s, researchers had discovered a role for electricity in the body. During the Roman Empire, electric rays were used to treat chronic headaches (Parent, 2004). They were applied directly to the head allowing the electricity to stimulate the brain. Experiments showed that electricity also was involved in the enervation of muscles. In a famous albeit morbid demonstration, Giovinni Adlini (1762 - 1834) demonstrated how electricity, when applied directly to muscles, could move parts of the body (e.g., open the jaws, open eyes, move limbs, etc.). The body was that of George Foster, who was executed publicly hours before for the murders of his wife and child.

In the 19th century, there were also published reports of cases of brain damage that appeared to be linked to specific behavioral deficits. The methodology is now referred to as the case study approach. Researchers can use this method to provide insight into the possible associations between deficits in functioning and the locations in that brain that have been damaged. The most famous case study of the 19th century is that of Phineas Gage (1823 - 1860), a railroad worker who sustained a brain injury when a metal rod passed through his head (Macmillan, 2002). He not only survived the accident, but did not lose consciousness. The most notable aspect of his case was that in the 12 years that he lived following his injury, he exhibited changes in behavior and personality. Prior to the accident, friends and family described him as a modest, religious man rarely in a bad mood. After the accident, he had outbursts of temper that struck others as out of character for him. Following the publication of the case study, researchers developed the theory that his behavioral change was due to his frontal lobe being damaged. The case study suggested that the frontal lobe may play a role in controlling emotional responses.

The most influential case studies from the 19th century were those of Paul Broca (1824 - 1880) and Carl Wernicke (1848 - 1905). Working independently in hospitals in different countries, they documented two different areas in the brain that when damaged were associated with different types of language deficits (Rutten, 2017). We now use the term aphasia to refer to language problems arising from brain damage. Broca, working in a hospital in France, described a form of aphasia caused by damage to a region in the frontal area of the left hemisphere. The patients with Broca’s aphasia had difficulty producing speech, but the speech that patients could produce was meaningful. Wernicke, working in a hospital in Germany, identified a different region in the left hemisphere, located behind the ear, which was damaged in patients who showed reduced ability to comprehend others’ speech. Although their speech was produced fluently, the content of the speech was nonsensical. In their studies, Broca and Wernicke each had accumulated data from multiple cases in which they had observed the language deficit in the patient and only later documented the location of the brain damage in the patient’s brain following the death of the patient in a post-mortem examination or autopsy.

The case studies of Phineas Gage and the work of Broca and Wernicke added support for the view that specific locations of the brain were believed to be responsible for specific functions (Kennison, 2020). This view is referred to as the **localizationist view.** The localizationist view originated with Franz Gall (1758 - 1828) who proposed that specific regions of the brain were responsible for specific personal characteristics and abilities (e.g., generosity, musical ability, secretiveness, and others) (Davies & Davies, 1955). The belief came to be called phrenology and those who advocated it phrenologists. Phrenologists claimed to be able to discern an individual’s personal characteristics by examining the pattern of bumps in one’s skull. They assumed that a bump in the skull was the result of an enlarged brain region below. When a brain region was particularly large, the person was believed to be endowed with more of that the trait or ability associated with that brain region. Although phrenology has been proven to be false and one of the most frequent examples of **pseudoscience** (i.e., a set of beliefs falsely portrayed as based scientific principles), it will forever be recognized as the origin of one of the most important early ideas about the brain.

As the 19th century came to a close, the groundbreaking discovery of the neuron was made by Santiago Ramón Cajal (1852 -1934). In 1894, he discovered that the brain was composed of numerous individual cells called neurons (Finger, 2004). His discovery would not have been possible without a method of staining tissue developed by Camillo Golgi (1843 - 1926). Prior to Cajal’s discovery, the common belief was that all the brain was composed of connected cells. Cajal was able to see the spaces between the stained neurons (or synapses) under the microscope. For his discovery, Cajal received the Nobel Prize in medicine in 1906. He is also known for his intricate and beautiful drawings of neurons, which allowed others to see what he saw using a microscope (Swanson et al., 2017).

[Insert Figure 1.1 about here]

In the early 20th century, Wilder Penfield (1891 - 1976) and colleagues began exploring what happened when weak electrical signals were applied directly to brain tissue, stimulating small groups of neurons (Penfield & Roberts, 2014). They were brain surgeons at McGill University in Montreal Canada, performing medically necessary procedures on patients. Prior to the surgeries, after patients had been prepared for open brain surgery, the surgeons electrically stimulated different brain regions and noted patients’ responses. The technique is referred to electrical brain stimulation (EBS) and continues to be used today to enable surgeons to map areas controlling functions needed for patients to maintain the highest quality of life (e.g., language, movement, music, etc.) so that the regions may be preserved as much as possible during surgery (Guillory & Bujarski, 2014). Patients do not experience discomfort during the pre-surgical EBS, as there are no pain receptors in the brain. The pioneering work by Penfield and colleagues (Penfield & Roberts, 2014) documented the functioning of large brain regions, including the language areas in the left hemisphere and the areas controlling movement (i.e., the motor cortex). However, when they stimulated relatively small groups of neurons, patients’ responses differed across stimulations, suggesting that the linkages between brain locations and functions/behaviors was clearest for large, rather than, small areas of brain tissue.

By the end of the 20th century, multiple technologies had been developed (Kolb et al., 2019). These are now routinely used to image the brain in a non-invasive manner for the purposes of medical diagnosis as well as research on how the brain is involved a wide range of mental processes. The later use of brain imaging techniques involves the testing of healthy research participants. Three of the most widely used in human research are electroencephalography (EEG), which involves measuring electricity produced within the brain and measurable via electrodes placed on the scalp, positron emission tomography (PET), which involves injecting a radioactive isotope into the bloodstream and measuring where the isotope travels via the bloodstream within the brain, and functional magnetic resonance imagining (fMRI), which involves measuring changes in water molecules in the brain when the body is placed in a strong magnetic field. fMRI, the most powerful tool, was invented in 1991, advancing existing MRI technology, which was developed in 1977. The excitement about the promise of new discoveries about brain functioning led the National Institutes of Health and the Library of Congress in the United States to declare the 1990s the decade of the brain (Library of Congress, 1990). In Chapter x, we will learn more about how these methodologies have been used to understand mental processes.

Despite the rise in the availability of brain imaging technologies over the last few decades, it remains the case that the majority of people in the world have not had a brain scan of some type (e.g., prenatal ultrasound, CT scan, MRI scan, or fMRI scan). Thus, our understanding of how the brain controls behavior and cognition remains limited. Typical brain imaging research studies test small numbers of individuals (Yeung, 2018). Small sample sizes limit researchers’ ability to understand the full range of human neural diversity. The reliance on small samples may contribute to the results of brain imaging studies being harder to replicate (Turner et al., 2018). Thus, our knowledge of brain and behavior remains limited. An example of this is the fact that there are now multiple reports of individuals who have received a brain scan in the course of medical evaluation and/or treatment and learned that their brains were extremely atypical in some way. Despite the very atypical brain, they have led fairly normal lives before the brain scan. Box 1.1 describes a few of the most compelling cases.

[Insert Box 1.1 Cutting Edge Brain Research: Normal People with Abnormal Brains about here]

 Within the field of cognitive science, the interest in neuroscience is more limited than one may presume. Neuroscience is viewed as providing just one level of analysis, how mental processes are carried out by the brain and body. Other levels of analysis have been described. For example Marr (1982) described three levels of analysis in understanding vision: computational, which refers to identifying the particular task being performed (e.g., picking up a cup), the algorithmic, which refers to specifying how the task is performed in terms of a series of stages of processing (e.g., see cup, generate a plan to plan to extend arm, generate plan to adjust hand position to grip the cup, execute the plans to move arm and hand, etc.), and the implementational, which refers to the details of how the stages of the process physically occur (e.g., in the most precise biological terms detailing the lowest level of cellular communication, signals generated from movement plans travel from brain via the spinal cord to enervate muscles, etc.). Others have proposed different terms for the different possible levels of analysis in the study of the mind. For example, Pylyshyn (1984) proposed the levels to be semantic, the syntactic, and physical. Neuroscience provides knowledge about the implementational or physical level of analysis. Historically, in cognitive science, there has been greater emphasis placed on understanding the computational and algorithmic levels of analysis for various aspects of human ability over the implementational level of analysis (Sloan, 1978). In each of the remaining chapters of this book, we will learn about cutting edge brain research that is related to the topic of the chapter. The information will be featured in a text box. We will also learn about what research on the brain has revealed about the nature of consciousness in Chapter 3.

**Psychology**

 The discipline of psychology focuses on understanding mind and behavior using scientific methods (Goodwin, 2015). Since its official beginning in 1879, marked by the founding of Wilhelm Wundt’s (1832 – 1920) laboratory for the scientific study of the mind at the University of Leipzig in Leipzig, Germany, it has become one of the largest academic disciplines. Wundt completed medical school training in 1856 and later studied physiology with Hermann von Helmholtz, a notable physiologist of the time who mentored Wundt in understanding the biological processes involved in sensation and perception (Rieber, 2013). Wundt was advocate of voluntarism, the view that the mind is controlled through will power (Goodwin, 2015). He studied the mind by using the **method of introspection**, which involves participants observing and describing their internal experiences.

Wundt trained numerous students to use introspection to understand mental phenomena. One of his most influential students was Edward Titchener (1867 - 1927) who established his research laboratory at Cornell University in Ithaca, New York. Titchener is credited with the founding of **Structuralism**, as approach to psychology that focused on understanding the structures of the mind, which included sensations, images, and affections. The method involved having participants experience a particular sensation, image or affection in a laboratory setting and then have them describe it following a standardized set of instructions (Titchener, 1928). At the time, modern statistical tools did not exist. They could not determine whether an observed result was due to change or having such a low likelihood as to be statistically meaningful. Consequently, the researchers routinely repeated their studies in order to observe that similar results could be obtained again and again. Even today, it is common for researchers to repeat experiments in order to determine whether similar results can be observed. Following the founding of Wundt’s laboratory, numerous students trained by Wundt spread the focus and practices of structuralism to other countries, including the United States (Goodwin, 2015).

As structuralism was gaining momentum in different parts of the world, the most famous psychologist of all time, Sigmund Freud (1856 – 1939) was developing his *talking cure* in his medical practice in Vienna, Austria*.* His approach to healing the mind is referred to as psychodynamic psychology, which emphasized the role of the unconscious mind in influencing human behavior. He also suggested that we have three minds, rather than one. The three minds were the id, the ego, and the superego. The id referred to our unconscious mind, including our basest desires including sexual impulses. The super ego referred to beliefs about expected behavior from the perspective of authority figures and society. The ego referred to conscious processes involving thoughts and feelings by which we purposefully act (Lear, 2005). Freud’s view of the three-part mind appears is recycled frequently in popular culture depicted as a person faced with a decision with a devil on one shoulder (i.e., the id) and an angel on the other shoulder (i.e., the superego), each encouraging different courses of action.

A third approach to psychology emerged in the late 19th century, flourishing in the United States. The approach was referred to as functionalism and focused on the functions, rather than, the structures of the mind (Goodwin, 2015). Among the prominent American functionalists was Williams James (1842 – 1910), the father of American Psychology. He rejected laboratory methods in the study of the mind. Nevertheless, he made major advancements in understanding psychological topics in his many writings and a wide range of topics, including the stream of consciousness, the difficulty in breaking bad habits, and different types of memory, one that lasts temporarily and one that seems permanent. Other functionalists included Edward Thorndike (1874 -1949) whose laboratory experiments with cats, documenting the law of effect, which refers to the fact that successful behaviors are repeated and unsuccessful behaviors are unlikely to be was an inspiration for later research on learning.

Following the publication of James Broadus Watson’s (1878 – 1958) *Behaviorist Manifesto* (Watson, 1913) in which he argued that the only worthy topics for study were those that could be directly observed, researchers in psychology increasingly turned away from topics related to unobservable mental processes. The approach became known as **behaviorism**. Watson was raised in South Carolina and received his Ph.D. from the University of Chicago, having been trained by researchers who embraced the functionalist approach to psychology. Watson emphasized the type of learning referred to as classical conditioning, which occurs when a stimulus triggers a behavioral response in an organism despite the fact that the stimulus never triggers that response in nature (Goodwin, 2015). This type of learning was also described by Ivan Pavlov (1849 – 1936) whose example has become widely cited. A dog salivates when presented with food, but after several occasions of the dog experiencing the ringing of bell when the food is presented, the ringing of the bell alone will lead the dog to salivate. For his work on the digestive processes in dogs, Pavlov won the Nobel Prize for Physiology or Medicine in 1904 (Todes, 2014). In one of the most well-known experiments in psychology, Watson and his graduate student Rosalie Rayner demonstrated that humans can learn emotional responses through classical conditioning (Watson & Rayner, 1920). The study has come to be known as the Little Albert study. A toddler who initially showed no fear to a white rat came to show fear consistently to the white rat after experiencing several episodes in which appearance of the white rat coincided with an unexpected loud noise (i.e., hammer striking metal pipe behind the child). The child later showed fear responses to other objects that were white and furry, including a white rabbit and a Santa clause beard. The study serves as an example of a classic experiment that has not been replicated because its procedures would violate modern guidelines on the ethical treatment of human research participants.

A powerhouse of behaviorism in the 20th century was B.F. Skinner (1904 - 1990) who described a second type of learning (Goodwin, 2015). He called it operant conditioning. His work on operant conditioning was influenced by Thorndike’s law of effect. Skinner referred to any behavior that an organism produced (e.g., pressing a button, raising a paw, or smiling) as an operant. In numerous laboratory experiments with animals, often pigeons, Skinner showed that interventions could influence the frequency of specific behaviors. Anything that increased the frequency of an operant was referred to as a reinforcement (e.g., food pellet given when a lever was pressed). Anything that decreased the frequency of an operant was a punishment (e.g., a mild electric shock when an animal moved to a specific location). He was able to condition animals to produce an amazing array of types of behaviors, and others could as well. Modern techniques of animal training rely on operant conditioning (McGreevy & Boakes, 2011).

Behaviorists including Skinner viewed the production of behaviors in humans and other species as involving only stimulus and response, without an intervening mental state in which the organism *decides* to produce the response (Goodwin, 2015). In the period between Watson’s (1913) manifesto and the end of the 1950s, behaviorist researchers churned out numerous studies on how observable behaviors were produced and modulated. Nevertheless, there was growing interest in the mind and related topics, a shift in psychology that is referred to as the Cognitive Revolution, and excitement about behaviorism was waning. Even B.F. Skinner turned his attention toward a topic that many would view as involving copious amounts of mental activity; however, Skinner continued to view all behavior as produced via S-R learning that did not involve intervening mental states. He published the book *Verbal Behavior* (Skinner, 1957), in which he claimed that children acquire language through the same procedures as they learn all other skills, through general learning principles, such as classical conditioning and operant conditioning. His account of language acquisition involved a complex system in which each observable language behavior was learned through a situation in which reinforcements and/or punishments were applied. The book and the behaviorist approach to language acquisition was not only unsatisfying as a theory, but also incorrect, as pointed out in an influential review by Noam Chomsky, a linguistics professor at MIT.

**Linguistics**

In his review of Skinner’s (1957) book, Chomsky (1959) pointed out that knowing a human language involves knowing a set of rules, an observation made by Ferdinand de Saussure (1857 - 1913), the Swiss linguist regarded as the father of modern linguistics, decades earlier. The study of languages has a long history, dating back to ancient Greece (Allan, 2013; Robins, 2013). Early work in linguistics was focused understanding ancient texts, comparing the characteristics of ancient languages and finding similarities between ancient languages and languages still spoken. By the 19th century, linguistics had recognized that language changed over time with some ancient languages being ancestor languages to languages still spoken. For example, Latin in the ancestor language for the romance languages (i.e., Spanish, French, Italian, Portuguese, and Romanian). Saussure was the first to claim that all human languages, regardless of how where in the world they were spoken, are complex and systematic, having rules that the speakers follow when producing the language (de Saussure, 1916/1977). He provided the analogy that language was similar to chess, which had rules for the movement of different game pieces. The sentences of language were constructed through the application of language-specific rules. He was the first to suggest that the languages spoken by *primitive* peoples were comparably sophisticated as the scholarly languages spoken in Western Europe.

Chomsky’s (1959) review emphasized that language rules allow for generativity of sentences, such that from a small set of rules any speaker can create so many unique utterances that will not have been created before and may not be created again. It is possible to put this claim to the test by using an Internet search engine. Create a sentence with 10 or more words and search for it (i.e., using quotes around it so that the search is for the verbatim sentence). Chomsky pointed out that most of the sentences that children are able to comprehend, even at the first years of life, are sentences that they have not experienced before. We will learn more about language acquisition and linguistic contributions in cognitive science in Chapter 5.

**Anthropology**

 Around the time of the cognitive revolution, scholars began to appreciate the existence of cultural knowledge shared among members of communities and passed down from generation to generation (Sloan, 1978). Language was an example of such knowledge, as were general practices related to daily living (e.g., cooking, agricultural, tool-making, etc.) as well as beliefs and customs, religion (Welsch & Vivanco, 2015). Anthropology offered the promise of being able to shed light on the minds of modern humans by illuminating the minds of prehistoric humans. An important topic is how complex behavior, such as tool-making, emerged in humans and how it related to other changes in human societies. The earliest known examples of tools used by humans were stone implements, dating back to approximately 2.6 million years ago. The stone objects were modified to have sharp edges and likely used for cutting. The tool-making abilities of humans exceeds that observed in other species (Van Schaik et al., 1999) and is evidence of how much human groups have been able to modify most aspects of the natural environment to provide more reliable supplies of food, more protective and more aesthetically appealing types of shelter, and a wide variety of objects that enhance daily living (e.g., vehicles, medicines, and art objects). A second important cultural innovation was language, which is estimated to have emerged in humans more than 100,000 years ago (Reboul, 2017). Anthropologists as well as those interested in evolutionary processes recognized that as human societies changed over time, there were also changes in bodies and brain. We will learn more about the evolutionary perspectives on the mind in Chapter 8.

**Computer Science**

The origins of modern computer science can be found in the 18th century with the work of Charles Babbage, who is regarded as father of the field (Hyman, 1985). The computer is a tool, which was initially innovated to solve a particular problem. At the time, navigational tables were published and used by ship captains to sail around the world. Errors in navigational tables could lead to shipwrecks, loss of expensive cargo, and also the loss of life. In his ambitious design of his difference engine, he aimed to create a machine that prevented mistakes in mathematical calculations, which were commonly made by human calculators (Dasgupta, 2014). He began the design in 1821 and completed it in 1830, but never completed the construction of the machine. He embarked on a second ambitious project that he called the analytical engine, which improved upon the design of the difference engine by incorporating punch cards that enabled the machine to have a rudimentary form of information storage. By the time of Babbage’s death in 1871, the machine existed only on paper.

Sixty years after Babbage’s death, a young Alan Turing was studying mathematics at the University of Cambridge and found inspiration in Babbage’s computing machines (Bernhardt, 2017). During World War II, Turing and thousands of others worked to create a code-breaking machine to decipher the daily communications of the Nazis (Erskine & Smith, 2001). Following the war, Turing published some of his most influential papers, which garnered him the title of the father of artificial intelligence. Turing (1950) first suggested that computers could one day process human language. In his paper *Computing Machines and Intelligence*, he described how researchers could test future examples of artificial intelligence. The original example describes a testing situation in which an interrogator is in one room and two others A, a man, and B, a woman, are in another. The interrogator aims to determine the sex of A and B, submitting written questions to each (e.g., how long is your hair?) and evaluating the responses. Turing (1950) further elaborates the thought experiment by suggesting that it may one day be possible to substitute a machine in place of A or B and change the task of the interrogator to that of determining whether both responders are human. In some, AI competitions have a panel of interrogators that evaluate participants (some human and some machines), and after (how much time), the panelists submit evaluations for each, distinguishing which were human and which were machines (AISB, n.d.). Recently, some competitions have invited the public to be interrogators of the machines.

Modern scholars recognize the official beginning of the field of artificial intelligence as 1956, during a six week summer workshop at Dartmouth College attended by 47 leading scholars in computer science and related fields (Moor, 2006; Solomonoff, 1985). Some of these individuals include Marvin Minsky, Allen Newell, Herbert Simon, and Ray Solomonoff. Solomonoff (1985) proposed six milestones for the new field. The first of these was the Dartmouth workshop, which served as the field’s beginning. The remaining milestones were challenges for the future. Among these were: a) the development of a general theory of problem solving; b) the creation of computers that can improve themselves; c) the creation of computers that can read text and incorporate the knowledge in the text into their own knowledge, as humans do. Although there is still much work to be done, some progress has been made these goals. There is a much that remains to be done, if computers will be able to simulate humans completely in examples of strong AI. In the 20th century, one program of AI research that has yielded success is the creation of machines that can play games against human opponents. Box 1.3 describes this success. We will learn more about 20th century innovations in artificial intelligence in Chapter 10.

[Insert Box 1.3 Innovations in Computing: A Worthy Competitor about here]

**Education**

 The academic discipline of education, as philosophy, likely predates recorded history, as all human societies are known to engage in methods of passing knowledge from one generation to the next (Marrou & Marrou, 1982). Notable events in the history of education include the invention of writing, which is estimated to have happened independently in different societies (Powell, 2012). The invention of paper in China in 200 BC (Barrett, 2008) and papyrus earlier in Egypt in 3000 BC (Houston, 2016) allowed for increased numbers of people to have access to books and other reading materials. The invention of the printing press in 1440 enabled books and other reading materials to be produced quickly, further increasing the availability of books and the knowledge contained therein to the general public (Kennison, 2019).

In the 19th century, governments in Europe and the United States increased efforts to educate their populations by implementing mass elementary education (Zinkina et al., 2016). The training of teachers began to occur in schools referred to as normal school. By the beginning of the 21st century, groundbreaking research in cognitive science provided evidence-based strategies for optimal learning (Brown et al., 2014). One strategy was spaced learning, which involved repeated learning sessions spaced over time (Cepeda et al., 2006). Spaced learning was shown to lead to superior long-term retention of information as evidenced by exam performance as compared to short duration studying just prior to the exam (i.e., cramming). The use of technology in teaching has led to changes in how educators view the processes of teaching and learning (Kentnor, 2015). Individual courses as well as undergraduate and graduate degrees are now being delivered via the Internet. Some anticipate that the future of education will involve most students completing the majority of their work via online platforms. Education’s contributions to cognitive science has increased rapidly as technology has enabled teaching and learning to occur more and more via the Internet. In the remaining chapters of this book, we will learn about topics core to academic learning, such as Chapter 2 (knowledge), Chapter 4 (memory), and Chapter 5 (language).

**Contemporary Cognitive Science**

Cognitive science emerged as a scholarly area of inquiry in the 1950s (Bechtel et al., 2001) buoyed by the excitement surrounding the cognitive revolution. The psychologist George Miller (1920 - 2012) suggested that the field was launched at a symposium on information theory that was held at MIT on September 10th and 11th in 1956. Attendees of the symposium included computer scientists Herbert Simon (1916 – 2001) and Alan Newall (1927 - 1992) and linguist Noam Chomsky (1928 - ). The name for the new area of inquiry was coined in 1972 by Christopher Longuet-Higgins, a computer science professor at University of Edinburgh (Longuet-Higgins, 1972). Two decades would pass before the founding of the Journal of Cognitive Science in 1977 and creation of the Cognitive Science Society in 1979.

The first meeting of the Cognitive Science Society was held in 1979 in La Jolla California, organized by Roger Shank, Allan Collins, Donald Norman as well as others. The program included sessions on developments in computer science, belief, memory, categorization, mental models, language processing, human development, neurological perspectives on knowledge and complex behavior, and cognitive science implications for education (Norman, 1979). The majority of speakers were computer scientists, including Marvin Minsky, Alan Newall, and Herbert Simon, and psychologists, including John Anderson, Gordon Bower, Jean Mandler, and Donald Norman, whose Ph.D. was in mathematical psychology and whose research centered on cognitive engineering. The program also featured the philosopher John Searle, the linguist George Lakoff, and the neuroscientist Norman Geschwind. Analysis of the first 20 years of the Cognitive Science journal revealed that the majority of articles were authored by psychologists and computer scientists (Schunn et al., 1998). In the same period, the majority of presentations at the annual meeting of the Cognitive Science Society were also made by psychologists and computer scientists. Undergraduate training in cognitive science first became available in 1972 at Hampshire College in central Massachusetts, but it was not until 1982 that the first undergraduate degree in cognitive science was established at Vassar College in New York (Vassar, n.d.). The first department of cognitive science was established at the University of California at San Diego (UCSD) in 1986 (UCSD, n.d.). In the 70 years since, there are now established Ph.D. training programs in cognitive science at many of the leading world universities (Cognitive Science, n.d.).

Some have argued that the field of cognitive science has not lived up to its promise. Nunez et al. (2019; 2020) asked the provocative question *What Happened to Cognitive Science?* The paper suggested that the field has failed, based on four indicators. Two indicators related to the affiliations of authors of published research articles in the journal *Cognitive Science* and also the patterns of citations. Both measures indicated a decrease, rather than an increase, in interdisciplinary research. Two additional indicators related to the extent to which cognitive science has been institutionalized in the form of Ph.D. training programs and core course work. The data indicated there are still few departments offering Ph.D. programs in cognitive science, and only about 15% of research universities in the U.S. offer an undergraduate degree. An analysis of course work for these undergraduate programs reveals little in the way of common required coursework. Among cognitive scientists, the Nunez study received attention, leading to an issue of the journal *Topics in Cognitive Science*, which published 11 replies and a follow up commentary by Nunez et al. (2020). Many of these replies agreed with these harsh assessments of the field, at least in part. Just three of the replies suggested that there is no reason for concern; the field is doing fine (French, 2019; Gentner, 2019; McShane et al., 2019). Gentner (2019) pointed out that cognitive science was never envisioned by anyone as a field that should pursue a “single theoretical framework” (p. 887). She described the field using the metaphor of a group of people assembled together to solve a problem. Each person in the group speaks a different language. She states that it is unlikely that the group’s working together is not expected to lead to the innovation of a new language to be used by everyone; rather, each person is likely to continue using their native language while attempting to learn one or more of the other language used in the group. She describes early and contemporary cognitive science as involving a diversity of viewpoints, which is a positive for the field. Others reject the notion that the field has failed simply because the idealized notion of cognitive science, which originated out of an AI perspective of the 1970s, has not come fully to fruition (French, 2019; McShane et al., 2019). Valuable research continues to be conducted that has benefited from the cross pollination of multidisciplinary ideas.

The question not addressed by these analyses of the field of cognitive science is: why has the progress of the field (if any has been made at all) been so slow. One possible contributing factor is the fact that doing interdisciplinary work of any type is hard (Boon & Van Baalen, 2019), as it involves the coordination of activities of individuals with different core interests, training, and institutional incentives for advancement. The institutional indicators noted by Nunez et al. (2020) may reflect the fact the success of disciplines is highly constrained by institutions, such as the universities in which scholars are employed and also grant funding agencies, which select priorities for awarding grants for research projects. The administrative structures of colleges and universities often create incentives for individuals and departments to compete rather than to collaborate. Over the last several decades, the economic pressures on higher education have increased due to sharp declines in public funding and increases in the direct costs to students (Zemsky et al., 2020).

During the sample period, interest in the creation and use of AIs by corporations has increased (West & Allen, 2018). Increasingly, work in businesses is being carried out by interdisciplinary teams (Tang, 2019). Unlike discoveries produced by academic scholars, which are described in publications, discoveries in companies are commercial intellectual property protected by patents (Poticha & Duncan, 2019). Knowledge about the science behind the products may not be available to the public, including students. The educational and research landscape in the area of cognitive science is likely to continue to be affected by economic pressures. It may be the case that location of the cutting edge research on the topic has shifted from universities to technology corporations.

**Social and Ethical Questions on the Horizon**

Advances in cognitive science generally, but those in AI specifically, have already begun to change our world. In the span of several decades, technology has rapidly changed society with the innovation of the Internet in the 1960s (Gillies et al., 2000), mobile cellular telephones in 1970s (Agar, 2013), and personal computers in the 1980s (Cringely, 1984). Technology has changed life for most people and promises to continue to change most aspects of society. The future of the field promises to deliver not only advancements in theory and in applications, but also social and ethical questions. Three of the most discussed areas are a) employment; b) privacy; and c) legal liability.

**Employment**

Innovations in AI and other technological advances are likely to eliminate many jobs and lead to changes in the nature of the work that humans perform across numerous employment sectors (West, 2018). Technological advancements in manufacturing in the 20th century changed how products were assembled. Increasingly factories incorporated **automation**, which involves the use of robotic equipment to perform tasks that were previously carried out by humans. For example, in factory that builds vehicles, workers may specialize in installing one part of the vehicle, which they do over and over each day. Automation is most successful when the same sequence of movements is performed repeatedly by the worker. Increasingly, businesses are using AI applications to perform more tasks, which previously were carried out by human employees. In the coming decades, humans are expected to be increasingly replaced by AIs in a variety of jobs, including transportation/driving, food service, food preparation, sales, legal research, journalism, and many others.

**Privacy**

Of all the ways in which technology promises to change society, its threat to personal privacy may be the greatest (Véliz, 2020). The word privacy is defined as the state of being free from intrusion by others (OED, n.d.). In a legal sense, privacy may also be extended to the state of being free from monitoring by others. Anyone who uses a cellular phone or accesses the Internet is vulnerable to having their movements, browsing histories, and purchasing patterns stored and accessed by business entities, governments, and individuals with a high level of computer science skills. The ability of corporations, governments, and others to collect and to analyze the personal data of others will likely increase in the coming decades. Despite the public’s increasing privacy concerns, there have been relatively few safeguards established to ensure that individuals will not be adversely affected by the unrestricted use of personal data.

In 2018, the European Union passed legislation known as the General Data Protection Regulation requiring that individuals not only consent to having their data collected, but also that they know and understand the nature of the data that are collected (European Union, n.d.). In 2019, California passed a similar law, which was named the California Consumer Privacy Act (State of California, n.d.). The law ensures that citizens have the right to know what information businesses collect from their devices about them as well as their children. In addition, the law permits individuals to sue businesses if their personal data is involved in a security breaches. In the coming decades, these laws and others like them will be tested in courtrooms as well as the court of public opinion. Among those with whom I have discussed this issue, there are some who like the fact that advertisers can use personal data to target them with ads for products that they may want to purchase.

**Legal Liability**

Legal liability generally becomes a question when something has gone awry, such has when there has been a car accident, a botched survey, or a missed cancer diagnosis during the reading of medical scans or biopsy slides (Pagallo, 2013). When car accident involves one or more self-driving vehicles, it may be challenging to determine where the legal liability should be assigned. The self-driving vehicles have a human operator, a corporation that owns the vehicle, another business entity that manufactured the vehicle, and another business entity that created the software that controls the vehicle, and possibly another business entity that created the original design of the vehicle and its operating software. Examples already exist (e.g., Wakabayashi, 2018). Legal questions related to technology, including the Internet, fall under the relatively new practice area referred to as cyberlaw. Cyberlaw attorneys are already handling medical malpractice cases related to accidents involving self-driving vehicles, robotic surgeries, and other types of technology-related cases.

In spring 2019, Boeing 737 max airplanes were grounded worldwide out of caution after two fatality crashes less than two years after first being put into service (Herkert et al., 2020). Lion Air Flight 610 took off from Jakarta Indonesia and crashed in the Java Sea on October 28, 2018 killing all 189 people on board. Ethiopian Airlines Flight 302 took off from Addis Ababa and crashed about 40 miles away from the airport on March 10, 2019 killing all 157 people on board. Investigators examined the black box units that recorded flight activity before the crashes. In hearings held by Congress, a Boeing executive suggested that the crashes were due to pilot error. The hearings also revealed that the airplane was equipped with software that automatically lowered the nose of the plane when it detected that the nose was too high. Not only did pilots not have knowledge of the existence of this software, they had not received training on when and how the software should be disabled. The majority of the lawsuits has been settled with the total financial compensation remaining secret; estimates place the total in the billions (Johnson, 2020).

**Summary**

 Cognitive science is the interdisciplinary study of the mind, including perspectives from philosophy, linguistics, anthropology, psychology, and computer science. With different histories, these disciplines became linked in the 20th century when the new field of cognitive science was born. Philosophers first posed the question of how the mind is related to the body as early as 400 BC in Greece as well as in Asia. Psychologist began studying mind and behavior in 1879, developing multiple approaches by the middle of the 20th century, including structuralism, and behaviorism. By the 1960s, the focus in psychology shifted to topics related to the mind, thinking, language, and cultural differences. This shift has been referred to as the cognitive revolution and is believed to have been influenced by the increasing use of computers. 20th century advancements in computer science and artificial intelligence by Alan Turing were inspired by the work of Charles Babbage in the early 1800s. There are critical social and ethical questions related to current and future innovations in cognitive science. These questions may lead to new social norms, laws, and occupations.

**Terms to Remember**

1. Analytical Engine
2. Applied Research
3. Automation
4. Basic Research
5. Behaviorism
6. Consciousness
7. Culture
8. Dualism
9. Functionalism
10. Localizationist View
11. Materialism
12. Monism
13. Neuroscience
14. Qualia
15. Reductionism
16. Research Hypothesis
17. Research Theory
18. Scientific Method
19. Strong AI
20. Structuralism
21. Voluntarism
22. Weak AI

**Questions for Review**

1. Discuss the distinctions of weak and strong AI. Provide an example for each type of AI from what exists now or from fictional portrayals of AI.
2. What is the mind-body problem? Describe at least two approaches to the mind-body problem.
3. Describe what is meant from the *easy* and *hard* problems of consciousness?
4. What are the three characteristics of phenomena that are most easily studied using the scientific method?
5. Describe how researchers in different disciplines use the scientific method.
6. How does the Buddhist view of the mind differ from that considered in ancient Greece and that proposed by Sigmund Freud?
7. Compare and contrast the three historical approaches to psychological research: structuralism, functionalism, and behaviorism.
8. What research topics came to be focused upon around the time of the Cognitive Revolution and later?
9. When and how did the field of cognitive science begin, how has it progressed in terms of basic and applied research?
10. Describe the ways in which society is beginning to be changed by AI and will continue to be changed, when AI becomes increasingly capable of simulating human mind and behavior.

**Recommended Resources**

Clark, A. (2013). *Mindware: An introduction to the philosophy of cognitive science* (2nd ed.). Oxford University Press.

Cognitive Science Society (2020). About. Retrieved September 16, 2020 from <https://cognitivesciencesociety.org/about/>

Mandler, G. (2011). *A history of modern experimental psychology: From James and Wundt to cognitive science*. MIT press.

Thagard, P. (2005). *Mind: Introduction to cognitive science* (2nd ed.). MIT press.

West, D. M., & Allen, J. R. (2020*). Turning point: Policymaking in the era of artificial intelligence.* Brookings Institution Press.

Zalta, E. N. (ed.). (n.d.). The Stanford encyclopedia of philosophy. https://plato.stanford.edu/

[Begin Table 1.1 Examples of Strong AI in TV and Film]

TV or Film Strong AI character

Stanley Kubrick’s (1968) film 2001: A Space Odyssey Hal, the computer

Michael Crichton’s (1973) film *Westworld* multiple robots in Westworld amusement park

Ridley Scott’s (1982) film *Bladerunner* multiple replicants targeted for execution

James Cameron’s (1984) film *Terminator* the cyborg assassin from the future

Paul Verhoeven’s (1987) film *Robocop* Robocop

Chris Columbus’s (1999) *Bicentennial Man* Andrew

Lana and Lilly Wachowski’s (1999) film *Matrix* series earth dominating self-aware AIs

Steven Spielberg’s (2001) film *AI: Artificial Intelligence* multiple robots

Jake Schreier’s (2012) film *Robot and Frank* the eldercare robot

Andrew Stanton’s (2008) animated film *Wall-E* (Stanton, 2008) garbage collecting robot

Duncan Jones’s (2009) film *Moon*  GERTY, the intelligent computer

Charlie Brooker’s (2011) *Black Mirror* series multiple characters

Spike Jonze’s (2013) film *Her*  the cellular phone personal assistant

Alex Garland’s (2015) film *Ex Machina* Ava, the robot subjected to a Turing test

Neil Blomkamp’s (2015) film *Chappie* Chappie

Jonathan Nolan and Lisa Joy’s (2016-2021) *Westworld* series multiple robots in Westworld amusement park

Denis Villeneuve’s (2017) film *Bladerunner 2049* multiple replicants

Grant Sputore’s (2019) film *I am Mother.* Mother

[End Table 1.1 here]

[Begin Table 1.2 Interdisciplinary Contributions to Cognitive Science about here]

Discipline Description

Philosophy understanding reality, knowledge, and values

Linguistics understanding languages as formal systems

Neuroscience understanding the nervous system across species

Psychology understanding mind and behavior

Anthropology understanding human societies and cultures

Computer Science creating hardware and software that process information

Education teaching and understanding how learning optimally occurs

Sources: Cognitive Science Society (n.d.)

[End Table 1.2 here]

Figure 1.1 Santiago Ramón Cajal drawing of Purkinje cell



[Begin Box 1.2 Insights from Philosophy: The Black and White Room about here]

Insights from Philosophy

The Black and White Room

Mary is an accomplished scientist who has lived her entire life within a colorless room, seeing only black and white. She has been able to explore the world outside using a computer screen on which the images shown are also colorless, displaying only black and white. Her expertise is specifically in neuroscience and she knows about all the physical processes involved in the perception of color, including the specific wavelengths of light that are detected by photoreceptor cells on the retina and how these cells communicate with other cells in a chain reaction until the information about the wavelength of light is received by the visual cortex of the brain. She also knows the physiological processes that take place in the brain to generate the thought that leads to the articulation of utterances, such as *Apples are red* or *Lemons are yellow*. Jackson’s (1982) compelling question was whether Mary would learn some new form of knowledge if she were allowed to leave the black and white room and to explore the colorful world or if she were provided a color computer monitor on which she could see the world in color from inside the room. The implication is that Mary’s direct sensory experience of a colorful world is a type of knowledge distinct from all her other sources of knowledge about the physical perception of color. Jackson suggests that mental states, such as perceiving color, are the result of physical processes, but are independent of them, unable to influence them. He intended the example to be an argument against materialism. Others have criticized Jackson’s (1982) dualist view, suggesting that Mary’s experience of seeing color provides her with no new knowledge (Conee, 1994; Nemirow, 2007). The example has been influential, even if it has not resulted in a consensus among philosophers. Recent films have referenced the problem and/or the uniqueness of the human experience of color. Garland’s (2014) film *Ex* *Machina* refers to the problem in the context of testing the strong AI Ava. The character Jake in Kaufman’s (2020) film *I’m Thinking of Ending Things* points out that the experience of perceiving color only exists one place in the entire universe – within the brains of humans. As we will learn in Chapter 4, color vision exists in non-human species with different species able to perceive different ranges of color (Kelber & Jacobs, 2016). An intriguing question is whether there is a subjective (vs. physical) experience of color in non-human species similar to that purported in humans (who have the physical ability to perceive color)?

Discussion Questions:

1. Approximately five percent of the world’s population has some form of color-blindness. To what extent are the qualia related to experiencing the color in daily life likely to differ for those who can and cannot experience the full range of human color vision? Do the color blind have less knowledge about color than others?

2. Experiencing brain injury can sometimes lead to color-blindness in those who previously had normal color vision. Would it be accurate to say that someone who loses color vision after a brain injury has lost a form of knowledge?

[End Box 1.2 here]

[Begin Box 1.1 Cutting Edge Brain Research: Normal People with Unusual Brains about here]

Cutting Edge Brain Research

Normal People with Unusual Brains

Brain scanning is becoming a common tool in healthcare settings, in the form of X-rays, CT scans, or MRI. Sometimes, a routine scan reveals unexpected findings. In 2007, Feuillet et al. published a case report titled *Brain of a White Collar Worker* in a top medical journal. The middle-aged man was a French civil servant, married with two children, and experiencing two weeks of weakness in one of his legs. A diagnostic MRI revealed that he had almost no brain. There was a large empty space where brain tissue should have been. The brain tissue that he had was organized in a thin layer outlining the shape of the brain (See Figure 1.2). Subsequent IQ testing revealed that his scores were in the normal range, but slightly below average. In his life, he experienced normal social functioning. His medical history revealed that when he was six years old, he developed hydrocephalus, a condition in which fluid accumulates in the brain. He had experienced leg weakness when he was fourteen years old, which was successfully treated by a medical procedure. The leg weakness that brought him to the hospital also resolved completely following a brain procedure.

In recent years, researchers have reported other cases of individuals living relatively normal lives receiving a brain scan and finding out that they too have atypical brains. For example, there have been reports of a man in his late 80s whose brain scan revealed a missing corpus callosum (Brescian et al., 2014), a 24 year old woman whose brain scan revealed a missing cerebellum (Yu et al., 2015), and a retired man in his 60s who had served as a soldier in the Russian military whose brain scan revealed a missing left hemisphere (Daily Mail, 2019). Such cases provide evidence that more research is needed to document and to understand the full range of human biology and human functioning.

Discussion Question:

1. A team of researchers propose to investigate the full range of human neurodiversity using brain scanning. They are requesting a large sum of money to do brain scanning randomly selected people on every continent, even those living in hunter-gatherer tribes in remote regions. Is such a project worthwhile? Why or why not. Is there a limit on how much grant money they should receive for the project?
2. Imagine that a family member or friend received a brain scan and then told that their cerebellum were missing. They have not experienced any noticeable problems with movement, which is the type of behavior that the cerebellum is known to control. How would you advise your family member or friend about how to proceed with their shocking news?

[Insert Figure 1.1 about here]

 [End Box 1.1 here]

[Begin Box 1.3 Innovations in Computing: A Worthy Competitor about here]

Innovations in Computing

A Worthy Competitor

 One of the crowning achievements in 20th century artificial intelligence was the creation of a machine that could out-perform human competitors in gameplay. The first chess playing algorithm dates back to 1948, created by Alan Turing and David Champernowne (Kasparov & Friedel, 2018). They name it Turochamp. The algorithm has been called the *paper machine,* because at the time that it was created, there was not a computer powerful enough to implement it fully. By 1958, four years after Turing’s death, another researcher had successfully implemented a chess playing algorithm that could play an entire game against a human competitor (Bernstein & de V. Roberts, 1958). Decades later, IBM began the development of a chess playing algorithms in the 1980s, naming the first ChipTest and the second Deep Thought (Campbell et al., 2002). Deep Thought became the first computer algorithm to best a chess Grandmaster in an official chess tournament. The rank of grandmaster is the highest that members of the World Chess Federation can attain (Hooper & Whyld, 1996). The most celebrated chess game between a computer and a human competitor was held in 1997. IBM’s Deep Blue play played in a best of six match with Gary Kasparov, who was the reigning world chess champion. In a previous competition, Kasparov had beaten Deep Blue 4-2. In 1997, Deep Blue beat Kasparov 3 ½ to 2 ½ (draws are scored as ½ point, wins are scored as 1 point). IBM continued their work developing artificial intelligence applications, tackling next the world’s favorite television game show -- Jeopardy. Their creation was Watson (Baker, 2011). On January 13, 2011, Watson competed against two former Jeopardy all-time winners Ken Jennings and Brad Rutter, and beat them both. Since then, the power of Watson has been aimed at diagnosing cancer (Cha, 2015), making weather forecasts (IBM, 2019), and, most recently, providing support for business (IBM, n.d.).

Discussion Questions:

1. Do you ever play computer games using the option to compete against the computer? If so, do you find such games as enjoyable as those in which you compete against people? Do you enjoy games more when you play against people you know or strangers (as occurs in online games)? Why do such differences in enjoyment occur?

2. In the future, it may be possible for entire sporting events, such as football games or hockey games, to be created with such high levels visual displays that it would be impossible to know whether it was a game being played by humans or computer creation. Would you find a game created by a computer to be as enjoyable as a live game being played by humans? Why or why not?

 [End Box 1.3 here]