Engineering Psychology

LISA JO ELLIOTT

Published by
PENN STATE UNIVERSITY LIBRARIES

The Pennsylvania State University is a member of the Open Education Network Publishing Cooperative.

Penn State University Libraries is committed to reducing barriers to learning and fostering equitable access to high quality educational materials for students everywhere. To learn more about open and affordable resources at Penn State, please visit open.psu.edu.

DOI: 10.26209/engin-psych

Recommended citation: Elliott, L. J. (2021). *Engineering psychology*. Penn State University Libraries. https://doi.org/10.26209/engin-psych

CONTENTS

SAFETY AND ERROR

UNIVERSAL DESIGN

PREFACE

Engineering psychology is an interdisciplinary field that relies on research in cognitive psychology, cognitive science, social psychology, sensation/perception psychology, industrial/organizational psychology, and clinical psychology as well as engineering psychology research. Engineering psychologists also rely on prolific research knowledge in electrical and industrial engineering as well as in information science and communications. As such, each of the chapters relies on an understanding of the previous chapters. Each chapter contributes to the knowledge before and after it.

This book has been written to a specific audience: undergraduate students who may not currently be psychology majors yet need an understanding of this topic to move forward in their careers. The topics are presented simply with many of the details about the research supporting these topics omitted in order to simplify the book and make it easier to understand. At the time of the writing of this textbook, all of the links work. If you find that a link doesn't work, please let me know. You can find me at Penn State or at LinkedIn. If the reader wishes to delve into the topics deeply, I recommend any of the textbooks on human factors / engineering psychology, the textbooks cited below, or the original literature cited. Finally, I urge readers to join the Human Factors and Ergonomics Society to learn more about this interesting field: www.hfes.org.

A NOTE OF THANKS

I am deeply indebted to Dr. Chris Wickens, his research and textbooks shaped our profession and many generations of engineering psychologists. I am also indebted to the following mentors in alphabetical order: Dr. Sylvia Acchione-Noel, Dr. Peter Foltz, Dr. Ken Forster, Dr. Doug Gillan, Dr. Melissa Guynn, Dr. Steve Hottman, Dr. Jim Kroger, Dr. Laura Madson, Dr. Jim McDonald, Dr. Dominic Simon, and Mr. Chris Wallen.

A deep thank-you to Floydd Michael Elliott, who spent countless hours giving feedback and support, and to Amanda Larson and Bryan McGeary, who worked tirelessly to shepherd this project through to its conclusion. Finally, a hearty thank-you to my many undergraduate and graduate students; it was through you that I learned that this book was needed and what topics should be covered. I wish you all successful and happy careers.

Thank you to the Open Textbook Network and Penn State for making this work possible.

Here are some other books on human factors and engineering psychology topics:

- Bridger, R. (2017). *Introduction to human factors and ergonomics*. CRC Press. ISBN: 9781498795944.
- Helander, M. (2005). *A guide to human factors and ergonomics*. CRC Press. ISBN: 9780415282482.
- Lee, J. D., Wickens, C. D., Liu, Y., & Boyle, L. N. (2017). *Designing for people: An introduction to human factors engineering*. CreateSpace. ISBN: 9781539808008.
- Niebel, B. W., & Freivalds, A. (2008). *Niebel's methods, standards, & work design*. McGraw Hill. ISBN: 9780073376363.
- Sanders, M. S., & McCormick, E. J. (1998). Human factors in engineering and design. *Industrial Robot: An International Journal*. ISBN: 9780070549012.
- Stanton, N. A., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C., & Jenkins, D. P. (2017). *Human factors methods: A practical guide for engineering and design*. CRC Press. ISBN: 9781409457541.
- Wickens, C. D., Gordon, S. E., & Liu, Y. (1998). *An introduction to human factors engineering*. Pearson. ISBN 9788120343719.
- Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology and human performance*. Psychology Press. ISBN: 9780205021987.

INTRODUCTION

CHAPTER 1

A HISTORY OF ENGINEERING PSYCHOLOGY

GOALS OF THE CHAPTER

- Learn what engineering psychologists research and how they are trained
- Learn about the different names for engineering psychologists and what they do in an industry or a government setting
- Learn how and why engineering psychology began

ASSESSMENT

- What are the differences between engineering psychologists and human factors engineers?
- How did engineering psychology begin according to psychologists?
- Does an engineering psychologist conduct research or provide therapy?
- Why are engineering psychologists often called engineers when they are not required to pass the ABET or take engineering courses?

WHAT IS ENGINEERING PSYCHOLOGY?

Engineering psychology is the study of how humans work in systems "from the neck up." "Systems" can be any coordinated goal-directed action between a machine or a series of machines/computers and a human. "Work" can be any activity that has a goal and measurable progress. The goal can be related to the person's profession, recreation, or personal life.

For example, a person may need a document printed for an upcoming meeting. In this case, that person may use a copy machine / printer that is connected to a network of computers in different offices or labs. The person accesses the copy machine / printer through the computer to print the document. The person must understand how to access the copy machine / printer, know the copy machine / printer's capability to print, and know the computer's ability to access the copy machine / printer. For the person to be part of this system, a complex mental model of the system must be available, including the physical location of the different system parts as well as their system states. In this example, the person is a part of the printing system as the operator of the machine. The machine cannot print the document without the person issuing the commands for it to print.

In another example, a person may want to find the closest public swimming pool for their children's birthday party event. In this case, the person may use their cell phone's mapping application and search for "swimming pools near me." This search triggers several systems that support the cell phone's application. The search engine determines what type of information the person typed into the search bar. The cell phone and the location of different public swimming pools are geolocated through cell towers, global positioning systems, and maps. These systems communicate with the cell phone application. The application displays the selected services for the person to choose from. The application may also have a function from which the person can dial the swimming pool's office phone number directly.

In a final example, a person may need to go to their physician for a checkup. In this case, the person does not initiate the action, but it is initiated on that person's behalf by other humans or the system. The person calls the doctor's office and makes the appointment. The administrator at the appointment desk sees the person's phone number, which triggers an application on the computer that accesses that person's medical records and other files. The administrator verifies that this is the correct information with the person on the phone. Once that has been verified, the appointment is made, and the system prepares the files for the physician. In this case, the automation has been designed to lessen the workload on both the person and the administrative assistant.

In each of these cases, knowledge of human cognitive capabilities and knowledge of human limitations play a role in the design process. Research has been conducted to determine the optimal settings for the systems to maximize human understanding and communication and lessen error reduction. This field of psychological research is called engineering psychology. People who conduct this type of research are called engineering psychologists, and those who put this research into practice may be called either engineering psychologists or human factors engineers.

In the United States, the persons who conduct research and make recommendations on how the human *physically* interacts with a system are called ergonomists. They study anthropometrics, or the measurements of humans and human movement in space. Outside of the United States, the categories of research are defined differently. In some places, the research in physical capabilities and limitations is combined with cognitive capabilities under the title of ergonomics.

THE HISTORY OF HUMAN FACTORS

The history of this subdiscipline of the profession depends on who is describing it. Psychologists trace the history back to **Wilhelm Wundt**'s experiments on sensation and perception in the 1800s and then Gestalt psychology in the 1940s, as described here: https://en.wikipedia.org/wiki/Wilhelm_Wundt. However, most engineering psychologists recognize that World War II played a pivotal role in developing the profession. In World War II, the use of technology was extensive. Walkie-talkies helped soldiers communicate on the battlefield. Airplanes delivered personnel, goods, and ammunition, as well as bombs. During the war, psychologists were employed to improve human performance. Psychologists investigated why some planes were more difficult to fly and crashed more often. This was the inception of the field of aviation psychology, which is still a strong subdiscipline within engineering psychology. Other psychologists investigated how to make walkie-talkies more effective. They implemented a system of codes that were easier to hear across the noisy battlefield. Many military and paramilitary organizations still use these codes today to convey important standard messages.

Those with a background in computer science may trace the origins back to management information systems and human-computer interaction. Human-computer interaction has a rich research background, as described in the journal article by Myers (1996). The origins of most of the devices that we use today can be traced back to several labs' groundbreaking work in the 1960s and 1970s, such as those at **Stanford**, **Bell** Labs (https://en.wikipedia.org/wiki/Bell_Labs), and **Xerox Parc** (https://en.wikipedia.org/wiki/PARC_(company)). These early studies laid the groundwork for the mouse, the Microsoft Windows operating system, graphical objects, and use of a graphic interface instead of a command line. Human-computer interaction is a part of the human factors discipline that focuses on how people work with computers.

Ergonomists may trace the origins of human factors to Alain Wisner's (1989) call to combine the efforts of psychology, sociology, anthropology, economics, finance, engineering, and politics. Wisner was one of the first to ask the members of the International Ergonomics Association (**IEA;** https://en.wikipedia.org/wiki/International_Ergonomics_Association) to address sociological disparities with technological solutions.

Other disciplines have contributed to engineering psychology such that it is one of the branches of psychology that has theories that cross many disciplines. In addition to theory, many professionals who are not trained as psychologists are considered human factors engineers. Versions of this class are taught from different disciplines' perspectives in departments of engineering, computer science, and communication. As such, the professional organizations for engineering psychologists reflect the broad perspective of the profession and the diversity of viewpoints.

THE PROFESSION AND PROFESSIONAL ORGANIZATIONS

The Human Factors and Ergonomics Society is the primary professional organization for engineering psychologists and human factors engineers. The San Diego Human Engineering Society and the Aeromedical Engineering Association of Los Angeles were the organizations that supported the discipline initially. The Human Factors and Ergonomics Society began in 1955, with the first meeting held in 1958 in Tulsa, Oklahoma. Their website states that human factors emerged during World War II to improve safety, signal detection and recognition, communication, and vehicle operations (Human Factors and Ergonomics Society, 2021). At the same time as World War II, human factors work in industry began with time-and-motion studies, task-analysis studies, and efficiency studies. Some of the first research questions asked, "Why do people keep crashing this particular plane?" "How do we make walkie-talkies easier to hear on the battlefield?" and "What is the most efficient arrangement of components on this assembly line?"

As the field of human factors grew, researchers continued to study aviation safety and operation, information-processing problems, safety and risk, control and display arrangement, and problems in the human perception of alarms. As automation began to be incorporated, additional issues of trust, dependence, attention, boredom, and awareness were also incorporated. Today, human factors professionals conduct research on nearly every product and complex system in nearly every sector of industry and government. At the time of this writing, the annual meeting of the Human Factors and Ergonomics Society attracted nearly five thousand students and professionals.

HUMAN FACTORS AND ENGINEERING PSYCHOLOGISTS: WHAT IS THE DIFFERENCE?

Typically, engineering psychologists are trained as experimental psychologists with an emphasis on research in cognition, sensation/perception, or social relationships. In addition to this applied research focus, they take additional courses in statistics, engineering psychology, methodology, and specific topic courses for their area of interest within engineering psychology. Many people assume that engineering psychologists either take courses in engineering or study engineers. Neither is the case. They "engineer" psychology to provide technological solutions through psychological methods. They earn the same salary as engineers with similar education requirements.

While engineering psychologists are a specific type of experimental psychologist, their work in the field of human factors means that they also must learn a great deal about industry, computer science, communication, media, information systems, anthropometry, and the other subdisciplines of psychology. It is not unusual for someone with a Ph.D. in industrial engineering to also have a specialty in human factors and work as a human factors engineer. Professionals with backgrounds in management information systems, computer science, communication, anthropology, kinesiology, media, mechanical engineering, and electrical engineering will take additional coursework that qualifies them to work as human factors engineers.

Human factors engineers work closely with engineers and product managers to develop documents of requirements for the system and the human side of a product. They may consider some of the ethical and safety ramifications of design decisions. They elicit feedback from potential operators and consumers of the product. They help the organization balance the ambitious goals of the product management and design team with the marketing and business goals of the organization. Finally, they ensure that the system or product complies with accessibility and safety regulations. Human factors professionals are often called to testify as to what types of research were conducted to ensure compliance and safety. Along with product managers, engineers, designers, and administrators, human factors engineers sign nondisclosure agreements, which require that they do not discuss the product or the research outside of the organization or courtroom without explicit approval from the organization where the research was conducted. This is one reason that very little is known about some of their work and few studies may be published on research related to product development.

Each subfield within engineering psychology has a rich history and requires a significant amount of domain experience before the engineering psychologist may be knowledgeable enough to make informed recommendations. A domain is an area where specific knowledge is needed. Examples of domains include medicine, the military, policing, firefighting, teaching, postal work, and aeronautics. These areas typically require people to complete a significant amount of training before they may begin work. An engineering psychologist working in these fields completes a portion of the training and demonstrates basic competencies so that they may understand and knowledgeably talk about the domain.

WHAT DOES A HUMAN FACTORS ENGINEER DO?

Let's look at an example in the field of medicine. Jenny is an engineering psychologist and works at a software company that specializes in electronic medical records. Her job title is a human factors engineer. One of Jenny's duties is attending the development meetings for her assigned product along with the software developers, the visual designers, the interaction designers, and the product manager. At the development meetings, Jenny observes the direction that the product is taking, new features that are being added or subtracted, and design questions that come up during the discussions. She asks questions that help her understand the history of the product and its users.

After the meeting, Jenny may contact expert users of the product to determine if these changes would help or hurt their ability to use the product. She may review similar products by other software manufacturers for their features and how these features work in their products. Finally, she may review the scientific literature in psychology, human factors, ergonomics, anthropology, medicine, or other fields to be sure that the changes considered by the team will indeed enhance the product's ease of use, its safety, its compliance with regulations, and its market share.

If she finds conflicts between what is proposed and what the users need/want and what the literature states, then she may recommend that the team consider gathering some data on how this feature will affect the product from all perspectives. For these data-gathering activities, she will need to know specifics about the product. She will need to know how electronic medical records are supposed to enhance the ability of the medical provider to do their job. She will need to understand a bit about medicine beyond what the average person knows. She may need to know how the program is written and how the system works.

This is so that she can design the experiment properly and communicate with the physicians and nurses who use the product on a professional level. She will need to know about the health privacy regulations, the other systems that her software interacts with such as drug formularies, and the environment where the software is used. She will need to know about noise levels, interruption frequency, how quickly the physicians/nurses learn new software, what other tasks they may be doing at the same time as entering the data, and many other things. All of this information will help her collect the right data with which the team will make the right design decision for the product and avoid costly litigation.

During her work with the medical teams, she may discover new ideas that she may bring back to the development team to create innovative services that they had not previously considered. Her understanding of how the medical community uses the product as well as how the software team develops the product puts her in a unique position to translate between the two communities and let each influence the other for the benefit of both. As she continues, she gains a deep understanding of the needs of both communities and a vast knowledge base about what does and does not work in the medical domain. Therefore, most human factors engineers do not switch domains because they would have to begin learning all over again.

REFERENCES

Human Factors and Ergonomics Society. (2021, February 21). About HFES: HFES history. https://www.hfes.org/About-HFES/HFES-History

Myers, B. A. (1998). A brief history of human computer interaction technology. *Interactions, 5*(2), 44–54.

CHAPTER 2

METHODS OF ENGINEERING PSYCHOLOGY

GOALS OF THE CHAPTER

- Learn about typical methods used to explore human factors as it relates to product design
- Understand that measurement is for the duration of a product
- Learn about the product development life cycle

ASSESSMENT

- What are the names of the types of tests that are done before and after a product is created?
- What are some examples in each category?
- What are three surveys typically used and what do they measure?

Let's go back to our example of Jenny, who works at an electronic medical records company. Jenny heard in the development meeting that her product is getting a new set of clients and family practice doctors in a large clinic. Previously, her product was used only by doctors in a hospital. She knew many of her customers well enough that she felt comfortable calling them on the phone. She called the hospital and asked for Dr. Hapt, whom she knew. Dr. Hapt worked in a geriatric medicine clinic. From her knowledge of medicine overall, she knew that a **geriatrician** would see a wide variety of clients and deal with a wide variety of health challenges in her patients. Dr. Hapt called Jenny back later that day, and they planned to have lunch together to discuss the challenges.

Jenny realized that this would be her opportunity to talk with a subject matter expert (SME) from the onset of this project. She drafted some questions that would help her understand the challenges that the developers, the new doctors, and the patients of these doctors would face when the new software was implemented. Here are some of Jenny's questions for Dr. Hapt:

1. What is the most challenging part of your geriatrician practice when keeping your patient records up to date?

2. Who are the stakeholders in an electronic medical records software on a practice level?

3. What part of a visit with a patient is the most critical—the part that you need to record the most information and get it absolutely spot-on, correct?

4. How do you like the hospital software? What is not working right? What could be better?

Jenny wanted to be sure that the questions were open ended so that there would be a lot of room for Dr. Hapt to give her more details. She wanted to keep to only a few questions and wanted the questions to be general enough that she would know more about the entire problem space, not just Dr. Hapt's specific challenges.

At lunch, Dr. Hapt was pleased to answer the questions. After lunch, Jenny wrote down what she found out in the notebook in her car before she went back to the office. When she got to the office, she typed the notes into a report along with additional information, such as how Dr. Hapt looked during each answer. Did she look angry? Did she seem happy? Did this part of the software frustrate her?

Jenny shared the report with her manager, Rick. She set up a meeting with him to discuss what she wanted to do next on the project. Then she reviewed the notes again and started to formulate the studies that she would need. Jenny determined that she would need the following:

1. A **comparative analysis** of the existing software on the market and the one that they had previously used

2. A **task analysis** of what a typical physician's day looked like when using the software and seeing patients

3. A **focus group** of family practice physicians describing what they liked and disliked about the software

4. A list of **user requirements** of the new software that included pain points and frustrations that she had observed in the previous studies

5. A **mock-up** of a design of the new software

6. Usability testing of the mock-up to discover additional **pain points** and frustrations

Jenny met with Rick. Rick shared that their budget had been cut and that they could afford only two of these studies that Jenny needed and that she would need to hire an outside consultant to meet the initial design deadline of 2 weeks from today. Jenny chose to do the task analysis and the usability test. She decided that she could do her own rough **mock-up** of the software design and that during the task analysis, she could ask questions of the physicians. Jenny asked Rick for a list of contacts at the family practice clinic.

She set up a task analysis for later in the week. Jenny would be working with Dr. Jean and Dr. Sherman on two separate days. Jenny started reviewing the materials that she would need for the task analysis and looking through her LinkedIn list for an outside consultant who could help her.

Let's talk about some of the methods that could be used to help design a product.

FORMATIVE TESTS

Competitive Analysis

A **competitive analysis** is a method where the analyst compares similar products or services to the one that they are helping develop. It is similar to a market analysis in that it examines all aspects of the competitive product or service, including the current and potential users. A competitive analysis is usually done before the product is developed or when the ideas for the product haven't been fully specified yet. This type of test helps the development team understand where their product will fit in the overall product landscape. They may get ideas for additional functionality or ideas of how to design a feature in a better way. The purpose of a competitive analysis is to discover what's currently available so that the team doesn't replicate what has already been offered.

To do a competitive analysis, you must obtain information on most of the products that are currently available to the public. This can be done through an internet search or by ordering the products. With large products such as car dashboards, it would not be feasible to obtain all the possible cars. This is where the analyst has to be inventive. The analyst might poll her friends and see if she could take pictures of their car dashboards. She might go to a rental car company and see if she could take pictures of their dashboards, or she may go to an automotive dealer.

The results of the competitive analysis list all the features of the different items being compared. These results are often used in requirement documents.

Task Analysis

A **task analysis** is done when it's unclear how the users usually do the task or when the design team wants to make the task as close to a real task as possible. For example, designers know that most people will not have used an automated or self-driving car before. They expect that new users will be wary and easily confused the first time they enter the car. The design team wants to make the starting of the car experience to be as close as possible to the way that people currently start their cars. Currently, there are many ways to start a car. Some cars require keys to be put in the steering column and turned.

Other cars have a push button that starts the car when the key fob is near, and you must press the brake pedal to start it. Which task analysis should you choose? Who are your potential users?

Let's say that you determine that the potential users of your automated car are people who have bought new cars in the past 5 years. You determine that all these cars have a push-button starter. In this case, you would do a task analysis using one of the new cars and its owner.

In the task analysis, you would ask the owner to show you her daily routine in the morning on the way to work. You want to know what she does before and after she starts the car. One typical morning, you and your assistant follow her around. You both take notes on what she does, how she does it, and how long it takes. You take the notes back to the design team and create a storyboard of what she is doing and what she does next. Throughout this process, you discover that she often keeps her keys deep within her briefcase and that there are other things in her briefcase that interrupt the radio frequency signal in the key fob. She complains that she often has to dig out the key fob in order to get the car to start. These are the types of findings that task analyses are designed to discover: small, seemingly insignificant habits that humans have that interfere with the system operation.

You want to know how frustrating and how common this habit is among new car owners, so you do another task analysis with another car owner. In this case, you decide to do a cognitive task analysis. In a cognitive task analysis, you ask the person what they are feeling and thinking at each step of their morning. When you do this, you find that men also lose the key fob in their briefcase. But the two men that you interview keep an extra key fob in the glove box of the car. Both seem to be unworried about security issues this might impose. You bring this finding back to the design team. One of the developers mentions that this is the way his car was recently stolen. The team decides to explore different ways of storing a key fob.

Focus Group

Whenever a product is in the initial design phase, the design team will call on previous and potential users to discuss the product. This is a **focus group**. The hope is that the group will recall incidents with the product or have suggestions for better designs. In reality, focus groups are very difficult to conduct, as the members will influence each other's opinions. Sometimes the discussion can evolve into a gripe session rather than a productive critique of the product. Other times, the members of the design team can get emotionally involved and subvert the conversations. It takes a very skilled moderator to lead a focus group to produce actionable results.

SUMMATIVE TESTS

Usability Testing

Once there is a design for the system, it can be tested on users. Usually, the test has 1 to 10 tasks that the user does with the system. Typically, about five users are tested on the system. The purpose of the test is to find usability problems. A usability test seeks to expand knowledge about the product being tested. The researcher simply wants to know how a person acts with the product. The results of the test are then taken back to the design group for review. They will determine how severe the problem is and what should be done about it.

Plain Old Experiment (POE)

Engineering psychologists are experimental psychologists with some extra coursework in human factors. Sometimes, the best way to discover what is happening is to conduct a carefully constructed experiment. Experiments are time consuming and costly. Often, industry professionals will choose to do a less expensive or quicker method first. To set up an experiment, the psychologist must identify an item that can vary (**independent variable**) and an item to be measured (**dependent variable**). Then the psychologist must determine which states the independent variable should be in to produce the clearest measurement at the different times. The results tell the psychologist which state produces the best human behavior and the best system behavior. A POE is used when there are no existing methods that will answer the design team's question. Usually, their question has to do with which interface is the best or which button yields the fastest response. The experiment's results are governed by the nondisclosure agreement protecting the product. Most results are not publishable and are only used to enhance the product's development.

Heuristic Analysis

Heuristic analysis simply means how well something adheres to a rule of thumb or a general design rule. A heuristic is a general rule. There are several types of heuristic analyses. The most popular heuristic analysis is Nielsen's 10 heuristics (Nielsen & Molich, 1990). They can be found here: https://www.nngroup.com/articles/ten-usability-heuristics/. Nielsen's 10 heuristics include many standard usability principles and have stood the test of time as universal design principles. When there is absolutely no budget for any type of user testing, experimentation, or other analysis, a heuristic analysis is a quick and budget-friendly way to be sure that some of the most damaging errors are avoided.

A heuristic analysis is done by at least one **analyst**; it is better if there are three analysts. The three analysts examine the product for each of the 10 heuristics separately. Then they discuss the product together. They make a list of the problems and errors along with the severity of each one. This list of problems is ordered according to severity and then delivered to the design team to discuss. Usually, a design team will address and resolve the problems in order of severity. Sometimes, a problem cannot be resolved, and the team creates training or another way to educate users on how the system works.

Measures of Workload, Usability, or Situation Awareness

There are several **surveys** that measure **different constructs** between the human and the system. Surveys can measure how usable the system is with the System Usability Scale (Lewis & Sauro, 2009) or how difficult something is to operate with measures of workload such as the NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988). Sometimes surveys measure how much of the environment the person is aware of at the time through a situation-awareness measure such as the SAGAT (Endsley, 1995). Each of these measures relies on a person's perception of their internal state at the time. Because of the **subjective nature** of these measurements, the surveys are typically given more than once to compare before and after. This assumes that the person's internal state remains constant and the only change is the system. However, surveys are usually paired with other types of observations before an expensive change is made to the system.

CONCLUSION

POE is the basis for formulating most theories in engineering psychology. Using POE, a lot of what we know about how humans have performed with other types of systems in the past is published in the journal comprising the body of knowledge for engineering psychology: *Human Factors: The Journal of the Human Factors and Ergonomics Society*. Increasingly, expertise plays a role in system design. One part of the system may be designed for main users (expert physicians), and another part of the system may be designed for novice users (patients). When expertise plays a role in the system design or evaluation, the *Journal of Cognitive Engineering and Decision Making* (*JCEDM*) may be used as the main body of knowledge. The remainder of this text will focus on what we know about how humans use systems, human performance, safety, ergonomics, and limitations. Within those sections, we will revisit some of the testing methods and add a few more.

The following is where you can go for more information:

http://www.hfes.org/

http://hcibib.org/

REFERENCES

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors, 37*(1), 65–84.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology, 52*, 139–183.

Lewis, J. R., & Sauro, J. (2009). The factor structure of the system usability scale. In M. Kurosu (Ed.), *Human centered design, HCD 2009, lecture notes in computer science* (Vol. 5619, pp. 94–103). Springer.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. In J. Chew & J. Whiteside (Eds.), *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 249–256). ACM.

THE DESIGN OF COGNITIVE WORK

CHAPTER 3

ATTENTION VIGILANCE AND FATIGUE

GOALS OF THE CHAPTER

- Learn about constructs in psychology and why they are necessary
- Understand the main theories of attention and processing
- Learn about vigilance and fatigue and how they impact system operation

ASSESSMENT

- How is attention studied experimentally, and can you give an example of an experiment?
- What is inattentional blindness?
- What is a self-terminating search, and what is the difference between a self-terminating and an exhaustive search?

ATTENTION

Arguably, **attention** is the most important area in engineering psychology. Without attention, accidents occur more frequently, and we are unable to process information and make decisions that impact our performance and the state of the system. A complete study of attention from both the psychology literature and the human factors literature could take more than a single chapter in a single course. In this chapter, we will briefly touch on some of the more pertinent topics in human factors.

WHAT IS A CONSTRUCT IN PSYCHOLOGY?

Unlike the other sciences, such as chemistry, biology, or physics, psychological science cannot directly observe many of the things that it studies. Because of this lack of direct observation, we have to think of unique ways to isolate and measure the thing in which we are interested. Isolating something to be studied and then describing it in a way that can be studied is often referred to as "making a **construct**." The thing itself is then called a construct because we are constructing a definition of what it is and what it is not and how it should be studied. In cognition, many of the things that we discuss are constructs. The first one that we'll discuss is attention.

What is attention? This construct is defined just as it seems. If someone is processing the information presented, then they are attending to it. If they are not processing it, they are not attending to it. However, we only know if someone is attending to something by their behavior. If they attended to the item, then they will be able to state that they noticed it, they will remember it, or they will behave as if they had noticed it.

WAYS THAT WE MEASURE ATTENTION

In behavioral experiments, the person is shown the item or the scenario, and then we measure if their behavior changes and how quickly it changes. Usually, attention experiments involve two tasks: a primary task and a secondary task. The person will be asked to do one item (primary task) such as ride a bicycle. Then we ask the person to watch a screen for a stop sign (secondary task). In this way, the two tasks require two different types of actions. The primary task requires that the person coordinate their legs and feet to ride a bike. The secondary task requires that they look for something in the environment. We say that the two tasks are in two different modalities: physical movement and visual search.

The experimenter would first ask the participant to do the easy task to assure us that this is something that our participants could attend to easily. Then we would assign the secondary task at varying levels of difficulty. For example, we might have a few levels that increase in difficulty, such as watching for a stop sign, watching for a pedestrian, and watching for a certain type of tree on the side of the road. If we were interested in additional processing capacity, we might add a third task of listening for the word *mango* while doing the first two tasks. This would make it harder as people concentrate on doing all three at once. This is called the **dual-task paradigm**.

Experiments such as this one led to the discovery that most of a person's attention is a step-by-step or serial process. It also led to the discovery that people will select a primary task and do that very well. The secondary and tertiary tasks will suffer neglect as the person focuses on the primary task.

The dual-task paradigm also tells us that people will switch which task is primary on their own in order to reallocate attentional resources. At the beginning of the study, the participant may ride the bike well and focus on maintaining a steady pace of 12 mph. Then as the study progresses and several stop signs appear in rapid succession, the participant may focus on identifying the stop signs, and the bike riding will slow down to 7 mph. If during the study, there are several stop signs in succession and then several incidents of the word *mango* on the headphones, the participant will struggle and may quit riding the bicycle. This is due to attention-switching costs.

The other way that we know if someone has attended to something is through eye-tracking. If their gaze has been directed at an object and they have behaved as if they noticed it, we count this as attending to it. Eye-tracking is a very popular way to measure visual attention in psychology, marketing, and other fields. More information on how eye-tracking is linked to attention can be found here: http://www.scholarpedia.org/article/Eye_movements.

WHAT IT MEANS TO ATTEND TO SOMETHING

Attending to something means that we have sensed the item that we were searching for or something in our environment. If we are searching, then once we have perceived that this is the correct item or the target, then we respond. Attention requires all of these steps: sensing, perceiving, and processing. We know from studies that mistakes in attention can occur at any of these steps.

In the bike-riding example above, the participant can fail to sense the stop sign because there was too much interference in the environment (noise). This might happen when there is a tree partially covering the sign or when another vehicle or pedestrian is passing and occluding the rider's vision. These would all be sensing errors that cause the item not to be initially seen. The rider would have the feeling that they never saw the sign at all because they did not.

In addition, the participant may see the sign but not perceive that it is a stop sign. This might happen when the sign is partially covered or when the light shines in a way that the identifying color, shape, or word is mistakable for something else. Or it might happen when the rider sees the sign but fails to recognize that it is the correct sign. Or sometimes people see a similar sign and mistake it for the correct sign. All of these would be perception errors: the participant saw the sign but was unable to perceive it properly.

Finally, the most frequent error in attention happens at the processing stage. We may sense and perceive that it is a stop sign, but the primary task for monitoring for traffic signals may have become the second or third task, and we fail to process the meaning of stop and fail to execute the behavior of stopping. This is what typically happens in texting-and-driving accidents. David Strayer at the University of Utah has done groundbreaking work in this area. An overview of his work can be found here: http://appliedcognition.psych.utah.edu/.

AUTOMATICITY AND ATTENTION

Automaticity happens when a person does a task so often that it takes up very little of the attentional bandwidth, as in riding a bike and watching for a stop sign. Riding a bike takes very little of the person's attentional effort and is automatic. As the person rides, it is very easy to do other tasks that don't involve manipulating the bicycle. Watching for a stop sign also does not take up many resources, as this is a task that is done frequently. So watching for a stop sign and bike riding leaves room for other tasks as well. Eventually, a person's attention will be consumed, and the person will be unable to attend to additional tasks. This is complicated by the fact that the person never knows which task will overwhelm the attentional system until it is too late and an error is made.

This is why texting while driving is so dangerous. In many cases, a person may do this successfully, as driving is an automatic task for many seasoned drivers. Monitoring the environment for traffic signals or for dangerous situations seems automatic. Drivers may choose to text when they feel that there is nothing that will demand their attention over what they feel is automatic. The accident occurs because driving is an unpredictable task. It is unclear when the outside environment will change because other humans are unpredictable. It is also unclear when the inside environment of the car will demand to be changed. Keeping the car stable in concert with other drivers/cars and the environment is the driver's primary task. When a person texts, the primary task changes to the texting. This all works out fine if the systems remain constant, but once an unanticipated change occurs, the driver does not have time to switch primary tasks back to driving, and an accident occurs.

WHAT IS TASK SWITCHING?

Task switching happens when a person has more than one task to perform, and they change the primary task to be the secondary task. The cost in thinking time is minuscule, so people think that there is no thinking time associated with the switch. Experiments in dual-task studies have demonstrated that this time is measurable and accumulates. For example, if you are studying and monitoring your phone for an important text message or alert, you are still studying. Each time you switch the primary task of studying to the task of picking up your phone, this attention switch costs a few seconds. If you compare the amount of information learned in this session to a session of only studying, there is significant information loss in the phone-monitoring and studying session. In this way, we know that task switching has a small but significant cost.

ADDITIONAL TERMS TO KNOW

There are some additional terms that you should know. The first is **selective attention**. Selective attention is when you change your attention to focus on one area. There is **focused attention**, which is when you select a particular item and attend only to it. Next is **divided attention**, which is when you try to attend to more than one item at once. **Sustained attention** is when you continue to attend to a particular item in your environment during a very long task.

There is also **multitasking**. This is when a person tries to attend to more than one item at a time. There is an invisible cost to multitasking, because each time a person changes their focus, there are a few milliseconds when the person needs to catch up, as in the task-switching example in the previous paragraph. Essentially, it takes effort to attend to one item and then to another. This can be visual, auditory, or physical attention. Usually, we think of attention as visual attention. Let's talk more about attention, but just in terms of visual attention. When you attend to an area, that area is called the **area of interest** (**AOI**).

Let's say you are watching the professor in class. The professor is in your AOI. There are certain things that you expect the professor to do, and there are certain things that you expect the professor not to do. When the professor does things that you expect, this helps your attention. If your attention is divided or you are multitasking and the professor does something that you don't expect, you're more likely to miss it if it is brief.

The professor will do things that are of high value or low value. She may tell you what will be on the test. This would be an example of a high-value item that will grab your attention. The professor will do things that are low value. Maybe she is previewing the reading for next week, and you already have read it.

There is a model that will predict when people will attend and when they will not attend to a certain item. This is called the **salience, effort, expectancy, and value (SEEV) model** (Steelman et al., 2011; Wickens, 2012). According to the SEEV model, when high-value items occur in a person's AOI and there is little effort that they must expend to attend to this item, then they will attend to it. If a high-value item requires effort to attend to it, then people are less likely to attend. Effort can be quantified by the amount of distance a person may have to move their eyes to see the item or the thinking effort that is involved (cognitive effort). For example, if the professor discusses what will be on the exam during the lecture, it is likely that you will attend to this information. If the professor writes the same information on the blackboard in the back of the classroom, according to the SEEV model, you are less likely to attend to it.

Change blindness happens when a person has simultaneous tasks and misses something that changes in their environment. Here is an example: A person is taking notes in class, and they are also listening to the professor, but they are really concentrating on the notes. The professor is wearing black pants and a purple shirt, and it is the first week of class. The professor steps out into the hall for a moment, and then a different professor enters with the same black pants and purple shirt and a similar voice and continues the lecture. The majority of students will not catch the change that this is a different professor. There are multiple YouTube videos on this effect along with Martens's (2011) research in this area. Harvard University Institute of Politics convention at https://youtu.be/uO8wpm9HSB0 is one of many YouTube videos on this effect.

Inattentional blindness is another term that is similar to change blindness. Inattentional blindness should not be confused with change blindness, as inattentional blindness happens when we see something but do not recognize it (Mack & Rock, 1998). Inattentional blindness differs from change blindness in that inattentional blindness is the failure to see something that is there, and change blindness is the failure to see a change that happens. For example, inattentional blindness occurs when we are texting and driving and see the traffic light change to red yet fail to respond properly. As we are attending to three different sets of things, texting may be the current primary task that consumes most of our attention. Hence the meaning of the red traffic light fails to prompt us to hit the brake pedal.

VISUAL SEARCH

Most of us are familiar with going through airport security and putting our suitcases through an X-ray machine. When we put our suitcases through this X-ray machine, a person searches for certain shapes and patterns in our bag. This would be an example of a **visual search**. In this case, it would be called a **serial self-terminating search**. When the person who is viewing the X-ray screen finds something that looks like it matches the outline of a knife or one of the other banned items, that person stops the conveyor belt, and that bag is set aside. Another officer takes that bag, finds the person that owns the bag, and then goes through the items one by one. In this task, the officer is searching for a target, such as a knife or a gun, among many distractors or nontargets. This methodical, item-by-item search that the officer performs is an example of a serial search. When the officer finds a target item, such as a small camping knife, we say that the search is self-terminated because the officer believes that they have found a banned item.

There are other types of searches as well. There is an **exhaustive search**, which is when the person searches all items in the field. Sometimes you are looking for individual features that are present together in one item, and sometimes you are looking for the absence of a certain feature in the area of the search. The literature surrounding visual search is based on the idea that we process things in a bottom-up or a top-down fashion. Here is a page that explains the difference between bottom-up processing and top-down processing: http://openpsyc.blogspot.com/2014/06/bottom-up-vs-top-down-processing.html.

AUDITORY ATTENTION

For both visual and auditory items, we have a very short-term storage area that collects and saves the sensations for us to select what we want to process and what we want to discard. For visual items, this is the iconic store. For auditory items, this is the echoic store. Both the iconic and the echoic memory stores last between 2–6 seconds for most people. During this time, our long-term and short-term memory storage areas will also interact with the iconic and echoic storage and sometimes help us select the most important sensations to process. When we store things for this short amount of time, we are usually unaware of them, although there is some processing. These storage areas are called preattentive (or before attention can engage), and we can consciously be aware of what we sensed. We know this from experiments with auditory attention.

Just as with visual attention, in auditory-attention experiments, the participants have a primary task and a secondary task. Usually, participants wear headphones, and different sounds are played through the right and left speakers. This is called the dichotic-listening task. A person may be listening to a conversation in one ear and then monitoring the other ear for a certain word. If that word is very meaningful to them (i.e., their name), then they are much quicker to respond. In fact, there is a phenomenon called the **cocktail party effect**, which describes how this happens in real life.

Here is an example of the cocktail party effect: Let's say that you are at a large party with friends. You are having a good time when you hear someone in the far corner say your name. Regardless of the amount of noise, you will be able to exclude it and hear parts of that far conversation where your name is used because of the importance of your name to you. Just as in a busy cocktail party with many people talking, important information is easiest to selectively attend to when it involves your name.

In addition to sensory paradigms such as auditory and visual tasks, there are cognitive attention tasks such as proofreading a paper or monitoring for a particular word or phrase on a screen. There are also proprioceptive and haptic tasks. Proprioceptive tasks are ones in which you monitor your body for a shift in its balance or the space that it occupies. A haptic task is when you monitor for a feeling in a control or touch pad that you are holding.

DETECTION RATES

Sometimes people will be more or less sensitive and detect a change more or less easily. We call this difference in detection rate bias. Bias can be modeled through signal-detection theory, where the number of correct targets found is the hit rate. The number of false targets identified is the false-alarm rate. The number of targets not found is the miss rate, and the amount of noise ignored is the correct rejection rate. A ratio of these numbers helps us estimate accuracy and bias in a search task or any performance task. For example, we might compare a novice to an expert.

We know that several things in addition to attention affect a person's hit rate or their ability to find a target when there is one. The intensity of a target, the value of the target, and the duration of the target all can determine whether a target will actually be found or not. For example, let's go back to the airport security visual search task. Let's say that each time an officer searches the X-ray screen for a knife and thinks that there is a knife in the bag, the officer gets a tiny raise in salary. If the officer sees the knife on the screen and then a subsequent bag search finds a knife, the officer gets a larger raise in salary. The officer is more likely to identify things that could be knives in the hope that one of them actually is a knife. In this example, there is no penalty for an incorrect identification or a false alarm.

Now to demonstrate bias, let's say that for the same task, the officer gets a pay decrease each time the officer states that there is a knife, and a bag search does not produce a knife. The officer gets a pay decrease for wasting the traveler's time and humiliating them in front of their fellow travelers. This officer's search bias will be more conservative, and the officer will be less likely to identify a knife unless there is enough evidence and the officer is fairly sure that one will be found. In the first instance, the officer is biased to produce as many hits and false alarms as possible because there is no penalty for a false alarm. In the second instance, the penalty is steep, so the officer avoids as many false alarms as possible.

Imagine your job is to be an airport security officer and monitor the X-ray screen and search bags. Let's imagine that one shift is 4 hours. For 4 hours, you will be watching the bags go across the screen, and you will be picking out knife- and gun-shaped objects. By the end of 4 hours, do you imagine that you will be producing more hits or more false alarms? You're right: it depends on how many targets you correctly identify and how much practice you have at this task. In general, your ability to produce hits will decrease the longer you are on the job, as this is a vigilance task.

A vigilance task is a task that requires a high level of attention for a long period of time, usually over an hour. As time passes in this task, you become tired or fatigued. Your sensitivity decreases. Your fatigue will increase as the task gets more difficult or the pace of the task picks up. For example, as a security officer, your sensitivity rate will be far higher when there are only three flights out of the airport compared to when there are 40 flights. Sustained demand theory (Matthews et al., 1990; Parasuraman, 1979) discusses this phenomenon. There are some things that can be done to help address fatigue during a vigilance task.

One way that has been implemented recently is to increase target noticeability. The next time you fly, if you peek behind the X-ray machine at the screen that the officer views, you will notice that there are green, yellow, and red outlines showing up on the screen. The newer X-ray machines incorporate machine learning to predict shapes of knives and guns at various angles within suitcases. If there is a pattern in the bag that matches one of these angles, a red outline of the shape will appear. This reduces fatigue in the officers by helping them in the visual search task through increased target noticeability and highlighting possible examples.

Another way to address fatigue in these officers is through a knowledge of their results. Was that odd shape really a knife, or was it an odd-shaped comb? The knowledge of how many hits and false alarms helps the officers improve their hit rate as they learn the shapes better. There are additional ways to increase sensitivity and appropriately set the criterion level as to what should be a hit and what is a false alarm. These include better instruction, an accurate estimation of the probability of co-occurrence, and shorter work periods. It is important to note that vigilance tasks also suffer when a human's internal state is not optimal. In other words, it is hard to pay attention when you are tired, hungry, bored, or unmotivated. Finally, the higher the workload or the more difficult the task, the shorter the period of work should be to address the decrease in vigilance over time.

UNDER AROUSAL

Just as important are tasks that are so simple that the operator is bored while they wait for the important parts of the job to happen. When there is much time between events, the miss rate goes up, and the hit rate decreases. When the value of the target is low, again the operator will become bored, and the miss rate will increase. Finally, when the target is difficult to detect, infrequent, and undervalued, the operator will find it difficult to monitor for the target and maintain vigilance.

CONCLUSION

Visual and auditory searches can begin in the preattentive phase, when the sensations are still in iconic and echoic memory stores. In a visual search, a person can use a top-down approach (e.g., "A gun has a barrel and a trigger, and so does this shape, so it must be a gun") or a bottom-up approach (e.g., "This shape looks like a trigger; is there a long cylindrical piece attached?"). A person can use a serial self-terminating search, a parallel search, a contingent search, or various other methods of searching depending on the goals of the search, what item is searched for, and where it is being searched. In an auditory search, a meaningful word can trigger a selective-attention mechanism. Detection rates can be quantified by the number of targets identified, the number of targets not identified, the number of items falsely identified as targets, and the number of nontargets ignored. This ratio can predict bias in the searcher and predict changes in bias and selection criteria based on external motivators or conditions during the search. It is possible to be overwhelmed and experience difficulty in maintaining attention or vigilance or underwhelmed and experience attention difficulties. Attention is the first part of information processing and therefore is key to understanding how humans sense and perceive a system and its tasks.

REFERENCES

Mack, A., & Rock, I. (1998). *Inattentional blindness*. MIT Press.

Martens, M. H. (2011). Change detection in traffic: Where do we look and what do we perceive? *Transportation Research Part F, 14*, 240–250.

Matthews, G., Davies, D. R., & Holley, P. J. (1990). Extraversion, arousal and visual sustained attention: The role of resource availability. *Personality and Individual Differences, 11*(11), 1159–1173.

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science, 205*(4409), 924–927.

Steelman, K. S., McCarley, J. S., & Wickens, C. D. (2011). Modeling the control of attention in visual workspaces. *Human Factors, 53*, 142–153.

Wickens, C. D. (2012). Noticing events in the visual workplace: The SEEV and NSEEV models. In R. Hoffman & R. Parasuraman (Eds.), *Handbook of applied perception*. Cambridge University Press.

CHAPTER 4

INFORMATION PROCESSING

GOALS OF THE CHAPTER

- Learn about the steps in processing information
- Understand the human limitations in information processing
- Learn about errors and how they happen

ASSESSMENT

- How does information go from the initial perception to final storage in long-term memory?
- What is an example of a retrieval problem and in what stage did it occur?
- What is cognitive fatigue, and can you give an example in your own words?
- What is cognitive engineering?

In Chapter 3, we discussed the psychological **construct of attention**. In this chapter, we will discuss another important psychological construct in engineering psychology: information processing. Information processing is a construct that describes attending to information in our environment and then doing something with it. That something may be an action (open up our homework assignment and turn it in), it may be a decision (I have to call the dentist and move my cleaning appointment), or it may be something we want to remember (I'll never forget where I was when I got that phone call). **Information processing** can be complicated, and a lot can go wrong as information travels from our senses to its destination.

First, let's talk about the types of information that are processed. There is information specific to our person, such as **internal sensory input** (my foot hurts) and **internal environmental information** (I am tired from not getting enough sleep). This is internal information. Internal information contains conditions that are specific to us and conditions experienced only by our person, no one else.

Then there is information specific to our external environment that is shared. This could be **external sensory input** (I heard my phone beep) and **external environmental information** (The heat in this room is making it difficult for everyone to concentrate). It can also be shared communication or relationships between humans or between humans and systems.

Let's think about how these types of information are processed in general. First, the senses take in information. Then this information is stored briefly in memory. There is memory for auditory information, called **echoic memory**. There is memory for visual information, called **iconic memory**. At this point in time, we believe that echoic and auditory memory is not selective and the memory store is very brief, only a few seconds.

Attention guides the selectivity of what is to be understood from **echoic** and **iconic** memory. Attention guides perception of the sensory store. The other things that the person is attending to or thinking about influence the processing of sensory information. Occasionally, this gives rise to controversy, as different people perceive the same sensory information in different ways. The black-and-blue dress controversy is one example (https://en.wikipedia.org/wiki/The_dress).

After the information is selected and processed through sensory memory, it goes into **working memory**. Working memory is a temporary storage place where information is processed further. Echoic and iconic memories that are not used are discarded/forgotten. In working memory, it is parsed and then stored in short-term memory, ready to be combined with other information. The Human Memory site (2020) has a reference page on memory (http://www.human-memory.net/types_short.html).

Once all the information is in short-term memory, a part of working memory called the **central executive** collates the information and decides what to do next with it. Baddeley called these components the visuospatial sketchpad (visual/spatial information) and the phonological loop (auditory information). Scholarpedia is a version of Wikipedia edited and maintained by scholars rather than the general public. Here is their page on Baddeley's work: http://www.scholarpedia.org/article/Working_memory.

The central executive in working memory is thought to be constrained in the same way that attention is constrained. In other words, every person has a limited ability to process items based on their internal environment and their specific capabilities. There are people who have vast capacities in working memory and can manipulate a wide variety of items quickly. Here is an example of working memory on the American Psychology Association Online Psychology Lab's demonstrations page under Memory and Memory Mapping in 3-D: https://opl.apa.org/src/index.html#/Demonstrations.

After the information has been organized in working memory, then it may be acted upon or stored in long-term memory for later retrieval. For example, let's say that you have a test in history class coming up in 2 weeks. You attend the lectures each week. During the lectures, your working memory compiles the information and stores it in your long-term memory. When you study for the test, you practice retrieving the right information and reorganizing it so that you can remember it later. When the test is given 2 weeks later, your working memory locates that information using the cues in the test questions. Hopefully, if everything goes correctly, the right information is retrieved, and you do well on the test (http://www.scholarpedia.org/article/Memory).

Working or short-term memory is thought to be very limited and lasts for about 20 minutes. In that short time, the information processing system is deciding which items need to be sent to long-term memory and making the right connections. The number of items that can be stored at any one time in short-term memory is thought to be about seven, according to Miller (1956). Other researchers since then have disputed his article and found that the true number that most people can juggle in working memory is closer to three or four

(Cowan et al., 2012). You can read more about this at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4486516/.

To expand this capacity, people may do things such as **chunking**. This refers to combining the seven items into a single item that is easier to remember. Let's say that you need to remember the code BMIP for a test. You remember that your father worked at the IBM Poughkeepsie office. You switch the order of BMIP to IBM-P. Then it is easier to remember based on your personal history. The test day happens, emotion overwhelms you, and you freeze. You are unable suddenly to remember the new code IBM-P. You remember that it had something to do with your father and something to do with where he worked, but the information is lost after that.

Two things may have happened. You might have retrieved something similar, and it interfered (Cisco SF is where your father worked after IBM-P). If you cannot remember the information at all and fail to retrieve it, it is called a retrieval error.

WHY DO RETRIEVAL ERRORS HAPPEN?

There could be a failure at encoding when the information wasn't practiced enough. There could be a failure at retrieval (the two places where your father worked could easily be confused). Or there could be a consolidation error in which the right information was remembered incorrectly (he actually worked at INTEL instead of IBM). All of these types of errors have implications for system designs in which the system relies on the human to rescue the system during a process that has resulted in an error or abnormal end.

WHY IS THIS IMPORTANT?

There are many times that systems are not fully automated. The system relies on the human to input information or give feedback. Sometimes the system fails and needs a human to restart it, enter information, or repair it. This is when the system design team needs to understand human memory and human information processing.

One of the most common instances in which a human needs to input information is when the system verifies that the human is an authorized user. Many systems currently use password authentication to verify authorized use. Knowing that a human can only remember about four to seven items in short-term memory helps the system design team accommodate this human limitation. However, the problem begins when the human needs to authenticate the system infrequently. Or the system keeps the password in memory for 30 days, and the human needs to remember the password after the 30 days have passed.

This is a difficult task for the human, as infrequent practice has failed to establish the links to the password in long-term memory. Therefore, password managers are necessary for many people who must remember 5–20 unique passwords per day. Typically, people will create a single password or suite of passwords and reuse them over multiple accounts, pushing the responsibility for the security of the account back to the organization providing the service. The human may have an incomplete understanding of how a system works, which contributes to an increase in risk taking at the service of convenience. When a breach happens because of insecure passwords, in the human's mind, the organization providing the account is liable, not the human. This can take a financial toll on both. To remedy this dilemma, many organizations are using biometric logins such as face, fingerprint, or iris recognition. These approaches do not rely on the human information-processing system.

WHAT ARE THE IMPLICATIONS OF PROBLEMS IN INFORMATION PROCESSING OR MEMORY?

If the person is unable to rescue or restart the system, the consequences vary. In most consumer applications, the person is left without the use of the application or device. In commercial applications, the consequences could become critical. In addition to memory errors, other things may impact a person's ability to rescue a system. These may be due to the person's internal environment overloading the person's cognitive system.

WHAT IS COGNITIVE WORKLOAD?

Imagine that you are a new pilot and it is the first time that you are flying solo. You have your instructor in the air traffic control tower to guide you. At the last minute, your mother or father wants to be on the plane with you. You are terrified. The addition of your parent joining you on the flight will compromise your working memory's ability to attend to the system's pertinent information by taking up attentional resources. If that parent talks to you during the flight, you will need to attend to their conversation in addition to processing the information to fly the plane. The addition of the internal emotional turmoil will overload your cognitive system, and you may find it difficult to concentrate. The impairment increases the likelihood of error and a crash.

In another example, imagine that you have a critical exam in an important class. The class is very difficult for you, with a lot of new information to memorize. You need to get a C in the class or you will have to change majors. You have already taken the class twice and failed it both times. During the exam, your ex-boyfriend/girlfriend comes in and sits next to you. The presence of this person will compromise your ability to concentrate and do well on the exam. You may find that you are unable to retrieve information that you knew before they arrived. The emotions that you are feeling interferes with your ability to retrieve the information for the test. Both examples are times when a person's **cognitive workload** had increased dramatically and impaired their information-processing ability.

WHAT IS COGNITIVE FATIGUE?

Cognitive fatigue is when your brain is tired. Maybe you have had several difficult exams in a single day, maybe there was an argument between friends, or maybe you watched too many episodes of *Letterkenny* on Hulu the day before. Or maybe you haven't had enough sleep for several days, and you feel foggy. Your brain doesn't want to process information, and it's difficult to attend to anything for more than a few minutes.

In both cognitive overload and fatigue, the human information-processing machine is overtaxed. If the human's cognitive fatigue happens during a critical system failure, there should be supports in place for the human to act on behalf of the system regardless of the human's temporary impairment. If the system failure is time sensitive, then immediate supports are needed for human memory and human information processing. These could be in the form of reminder cards or a second human backup. If the failure is not time sensitive, allow the human to resume the system support after rest.

INFORMATION-PROCESSING DECLINE

Sometimes rest does not fix the human information-processing problem. There are several instances where information processing from a human is unavailable to the system. In cases of dementia, brain injury, or aging, the human may not respond. The first two items cannot be remediated. The third item, aging, can be compensated for in many circumstances. As people age, their working-memory speed declines, but other memory systems will compensate. Their knowledge stores increase, as they have vast amounts of information from their years of experience. While they may not be able to process information quickly, they know from experience what needs to be done.

Another item, lack of practice, will also contribute to human information-processing problems. A person who hasn't driven a manual transmission car in many years may need extra time to get used to the pedals and controls again. In certain professions, decisions are made based on content knowledge that changes over time. For example, a computer scientist educated in 1970 would have to work to maintain her level of expertise in 2020 as the content of computers and computer programming has changed. This affects her ability to quickly process information, as she must select the most recent information from her vast experience and apply it while suppressing what she had learned initially in the 1970s. As the frequency of accessing the most recent information increases, so will the efficiency.

HICK-HYMAN LAW OR HICK'S LAW

The time it takes for a person to make a choice is related to the number of items they must choose from, the complexity of the decision, and the consequences of a poor choice. In other words, a choice between two items is rather simple and should be rapid. A choice between four items should be a bit longer. If the choice will impact later options, the person will take longer still. For example, a person deciding between two laptop computers for the next school year would typically take longer than the same person deciding what brand of coffee to buy at the grocery store. Aside from the price difference, the cost of a poor laptop will affect the person's ability to do well in their coursework. The brand of coffee is less likely to affect the person's ability to do well.

EXPERTISE

Expertise in a particular domain builds over time. Some researchers claim that the time needed to become an expert at something is 10 years or 10,000 hours (Ericsson et al., 2018). However, the exact number or length of time is controversial. Most researchers agree that deliberate practice is what builds expertise. Deliberate practice means that the novice has a coach or other expert critiquing their performance and giving tips on better performance. The novice incorporates the feedback, and the performance improves.

The concept of mentorship or expert critique continues to be what drives the accumulation of expertise in most domains. For research purposes, an expert is also called a **subject matter expert** (**SME**). Since the 1980s, efforts have been made to quantify and capture the knowledge of SMEs. Organizations wish to capture the institutional knowledge of their most senior executives. Military and educational institutions wish to model the knowledge of their most valuable personnel so that the knowledge continues to contribute beyond those experts' lives. Expertise is most valuable when it comes to decision-making within a domain. When Sully Sullenberger landed the plane on the Hudson River, that decision came from years of investigating plane crashes and years of piloting expertise (https://en.wikipedia.org/wiki/Sully_Sullenberger).

COGNITIVE ENGINEERING

Cognitive engineering is a new branch of human factors concerned with how people think and make decisions in real situations. It is popular in organizations where the **three Ds** (dirty, dangerous, and dull) describe the tasks and time constraints in human performance. Cognitive engineers combine what is known about cognitive psychology along with some of the investigative methods developed in engineering psychology to produce insights on how to improve problem-solving and decision-making. Military, first-responder, and health-care settings are ideal for cognitive engineering approaches.

Cognitive Task / Work Analysis

One of the methods used in cognitive engineering is **cognitive work analysis**. In this method, the analyst observes or transcribes interactions within a system. The analysis describes the people within the system, the nature of their interaction, their work domains and activities, strategies, social analysis, and competencies analysis (Stanton et al., 2013). The cognitive work analysis is done with the aim of developing a model of the system, system design improvements, system tools or training, team design, or interface design. It describes the constraints within a system along with dependencies and contingencies.

Critical Decision Method

The **critical decision method** (CDM) is an interview method that was an outgrowth of the critical incident technique (Flanagan, 1954). Since that time, it has been used to investigate decision-making in time-stressed situations where specific types of expertise are needed. Examples of these situations are firefighting, law enforcement, health care, and air traffic control. The CDM requires the analyst to learn about the domain so that they can ask the right questions at the right times. From these questions, the analyst can derive key decision points. These are times when the SME had choices on what to do next. The analyst will ask for additional information that informed the choice, the possible outcomes, and the possible weights. This helps the analyst understand the nature of the decision task and the nature of the domain. The output of a CDM can be constructed as a flow chart that describes networks of knowledge for expert systems, training, or evaluation.

CONCLUSION

Information processing is a complex cognitive process that uses sensory, short-term, and long-term memory. Attending to information is only the first step in this process. Along the way, information can experience interference or other errors. Attention and emotion can impact our ability to remember information by increasing cognitive load. Hick's law refers to the relationship between the number of items a person has to choose from and the time it takes to make a choice. Expertise is built over time with deliberate practice from a coach or mentor giving critical feedback on the person's performance. Cognitive engineering is a new area of human factors that investigates decision-making and the cognitive aspect of humans in a system. There are several methods based on observation and interviewing that reveal how humans process knowledge and make decisions in sociotechnical systems. These methods yield artifacts that can be used for training, creating expert systems, or optimizing the system.

REFERENCES

Baddeley, A. (2012). Working memory: Theories models, and controversies. *Annual Review of Psychology, 63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Cowan, N., Rouder, J. N., Blume, C. L., & Saults, J. S. (2012). Models of verbal working memory capacity: What does it take to make them work? *Psychological Review, 119*(3), 480–499. https://doi.org/10.1037/a0027791

Ericsson, K. A., Hoffman, R. R., Kozbelt, A., & Williams, A. M. (Eds.). (2018). *The Cambridge handbook of expertise and expert performance*. Cambridge University Press.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*, 327–358.

The Human Memory. (2020). *Short-term (working) memory*. Retrieved April 1, 2021, from https://human-memory.net/short-term-working-memory

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Stanton, N. A., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C., & Jenkins, D. P. (2013). *Human factors methods: A practical guide for engineering and design* (2nd ed.). Ashgate.

CHAPTER 5

TRAINING AND AUTOMATICITY

GOALS OF THE CHAPTER

- Learn about how procedural memory is different from other memories
- Understand the different types of learning
- Understand the relationship between practice and automaticity

ASSESSMENT

- How are operant and classical conditioning the same, and how are they different?
- What steps would you use to teach a dog to go to the refrigerator and get you a can of Pepsi?
- What are the two types of knowledge, and can you give an example of each?
- What are the two theories of automaticity, and can you explain them in your own words?

Information processing is affected by training. As a person learns what cues to react to and how to react, over time, this feedback loop becomes quicker. Training usually happens as a response to a specific job or a specific system that the human needs to operate.

Airplane pilots, air traffic controllers, nuclear reactor operators, all are examples of professionals who need to be trained to respond to the system's set of alerts, warnings, controls, displays, and operations.

Right now, you are engaged in classroom learning. In this type of learning, we are setting the stage for your future by explaining concepts, demonstrating skills, and asking you to integrate the knowledge and skills in projects, papers, and tests. Training is slightly different in that you are learning cues to look for and actions to do in response to specific cues.

People learn rapidly every conscience moment of their lives. They may learn what others intend for them to learn, they learn without guidance, and they learn what others did not intend for them to learn. Theories in psychology suggest that classical conditioning, operant conditioning, associate learning, and social modeling are the primary mechanisms by which people learn.

CLASSICAL CONDITIONING

Classical conditioning may happen intentionally or unintentionally when a stimulus creates a response, as with a pet dog or cat. When the animal hears the food bag or container, that animal automatically begins to salivate in response to the sound. The sound has been paired with the stimulus of food to signal to the animal that they are about to be fed. Usually, pet owners find that this unintentional conditioning is annoying but not troublesome. The same pet owners may try to intentionally condition their pet by telling their pet a command and then giving the pet a treat. The pet expects the treat after the command. The conditioning is in place. A similar thing happens when an alert sounds on a person's phone. The alert sounds, and the person picks up the phone, expecting the reward in the form of a message or social media posting. Addiction is a special form of classical conditioning where many stimuli signal that the reward of the drug is possible. In all forms of classical conditioning, breaking the pattern between the stimulus and response will, over time, erase the conditioning.

OPERANT CONDITIONING

Operant conditioning usually happens intentionally as one organism rewards or punishes another organism in order to get that organism to behave in a certain way. Again, this could be a pet owner who is trying to teach her dog to catch a Frisbee in midair. First, she needs to break down the action into steps that are easier for the dog to do. Think of the steps that it would take for a dog to learn to catch a Frisbee.

For each step, as the animal gets closer and closer to the desired goal, the owner rewards the dog. Soon the dog is doing each step: catching popcorn, then catching something larger, then catching a Frisbee that is tossed to it, then catching a Frisbee midair.

SOCIAL MODELING AND LEARNING

For cell phones, most people learn how to use different apps by watching their friends. This would be social learning, which has its origins in the work of Albert Bandura. Bandura suggests that when people watch other people do something (e.g., hit a doll), then they are more likely to do it the same way themselves. You might copy the way that your best friend dresses or something that you saw on Pinterest. This is also social learning. You infer that because someone else is doing it, the action is acceptable.

ASSOCIATIVE LEARNING

People also learn to pair things together with practice. This seems to be culturally based. Think of the three lines that appear at the top of many cell phone apps. This is interpreted as the hamburger menu. It is associated with additional functionality; if you cannot find the button that you are looking for on the screen, many people will click those three lines. Associative learning takes time, and some researchers believe that it is based on other types of learning that we have already discussed.

TYPES OF KNOWLEDGE

There are at least two types of knowledge: knowledge that you can explain to someone else and knowledge that you cannot. The first is called **declarative knowledge**—you can explain it, you can declare it. The second is called **procedural knowledge**—you cannot explain it clearly. You may explain the steps, but this isn't the same as declarative knowledge. This is important because declarative and procedural knowledge areas are thought to be separate information processing and retrieval areas. This is because of research with brain damage patients such as HM, whose actual name was Henry Molaison. His name was revealed with his permission after his death. Henry Molaison is a notable figure in psychology, as after a necessary medical procedure, he became unable to form memories, and so he volunteered to be a dedicated memory research subject. More information about Henry can be found here: https://en.wikipedia.org/wiki/Henry_Molaison.

Sometimes procedural knowledge is called **implicit memory**. This is because when you learn how to do something, you must relate the sensations that indicate the next steps with the motor movements. Implicit memory is also thought of as "muscle memory." If you are a longtime driver of a manual transmission car and you are training someone else how to operate a clutch, you will find that it would be easier for them to just do it themselves while you coach them as to how the clutch feels and sounds when you need to shift to another gear. You'll spend time describing how the vehicle sounds when it needs to shift. The lack of explicit direction may be frustrating. You will note that it would be easier to teach someone to bake a cake or tell them how to get to your favorite pizza place. This is because you associate sensory memory with muscle memory, but you must use explicit memory to convey the information.

There are several steps in learning a new motor skill using implicit memory. First, you must process the sensory information relevant to the action. You must break the skill into smaller parts so that you can direct your attention to the most relevant information. For example, according to Wolpert and colleagues (2011), where you direct your eyes, or gaze behavior, is task specific. As you need to process visual information, you will look for it. Experts and amateurs will have different gaze behavior, as they direct their visual attention differently in learning a task (Wolpert et al., 2011).

PROCEDURAL TRAINING AND AUTOMATICITY

Automaticity is when a declarative or procedural memory becomes so well known that you just know without thinking much about it. That can be knowing how to drive a stick shift, how to ride a bike, or how to tie your shoes. When you first learned these skills, it took quite a bit of concentration and practice. Then over time and with practice, these skills took less and less of your attention. Soon, you were able to do these things while doing other things that took attention—for example, holding a conversation with another person while driving or riding a bike or tying your shoes. Because these skills take almost no attention to perform, we say that they are automatic.

There are three different types of automatic skills. There are perceptual motor skills (e.g., riding a bike), combination perceptual/procedural skills (e.g., driving a manual transmission car), and procedural skills (e.g., making your favorite breakfast). There are two perspectives on how automaticity works. One is based on the use of very few cognitive resources: the resources perspective. The other is based on the use of a single memory that is retrieved very rapidly: the memory perspective (Logan, 1988).

As with many other things, we know much about how it exists but not a lot about how it works. Automatic processes are fast, effortless, autonomous, consistent, and unavailable to conscious awareness. Automaticity develops best when the skill is practiced in the same environment and in the same way until it generalizes. In other words, if you are learning to drive a manual transmission car, it would be faster to practice on the same car in the same environment over and over again until you become good at driving the car. Then switch to a different manual transmission car for the learning to generalize. This would be faster than trying to generalize the knowledge over several cars when learning to drive a manual transmission.

THE RESOURCES PERSPECTIVE

According to this perspective, automaticity develops with practice. Attention gradually withdraws as the person uses fewer and fewer cognitive processes to accomplish the task. As attention withdraws, the person becomes faster because they are not limited by the attention bottleneck as a serial or step-by-step processing task. For example, when a person learns to drive a stick shift, they first will press down on the clutch and then let it up slowly while depressing the gas as they put the car in the right gear. As the process is made automatic, this will be seamless rather than sequential, and the person will be able to operate the radio or hold a conversation while shifting the car because their attention isn't needed in shifting. In this viewpoint, people speed up through practice.

THE MEMORY PERSPECTIVE

According to this perspective, automaticity develops when methods are selected randomly, and if they are faster, they are practiced and retained. Then the resource demand is a function of practice rather than a function of the need to reduce resources. Once the skill has been well practiced, it changes from a series of steps (e.g., put the key in the ignition, turn the key, put the car into the right gear) to fewer combined steps (e.g., start the car and put it in gear) and then to a single step (e.g., start driving). According to the memory viewpoint, people retrieve the memory based on the different states of the environment. In this approach, requiring people to speed up requires learning additional things to perform the task faster.

Not every task can be made to be automatic. The task must involve motor movements in addition to cognitive cues. This means that any procedural task has the potential to be automatic with enough practice. Skills such as typing are good examples of things that are initially effortful and eventually automatic. Many people can type faster than they can handwrite because of the consistent daily practice in the task of typing. As the task is automatic, it becomes an implicit memory that is difficult to articulate. For example, your fingers know where the different letters are on the keyboard. If you were asked to label each letter on a blank keyboard, it is likely that you would have difficulty. You know which letters are where; you have internalized the task to such an extent that it has exited your declarative memory or explicit memory and entered your implicit or procedural memory.

CONCLUSION

Learning is complex and involves associations made with stimuli and feedback. In some cases, learning is explicit, and the students can be easily tested on the knowledge. In other cases, learning is implicit, and the students must be tested on the skills that they have developed as a result of learning. Both perspectives complicate training plans. Often, professional training involves a mixture of explicit knowledge and implicit procedures. When a pilot learns how to fly, their explicit knowledge includes the flight rules, national airspace rules, tolerances for the plane, air traffic control procedures, organizational procedures, and so forth; in addition, they need to have knowledge about the plane. Then there are the implicit procedures in learning about how to fly: how to communicate with air traffic control, how to feel where the plane is going to go next, and how to anticipate what the display panel will show as the plane engages in flight. All of these items make pilot training expensive and lengthy. The same is true for many other professions. Human factors professionals are often called upon to create training programs that will shorten the time or lessen the cost for the organization while creating the same learning effect in the pilot or other professionals. In this way, training and automaticity are important topics in engineering psychology.

REFERENCES

Logan, G. D. (1988). Automaticity, resources, and memory: Theoretical controversies and practical implications. *Human Factors, 30*(5), 583–598.

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience, 12*(12), 739.

CHAPTER 6

STRESS AND WORKLOAD

GOALS OF THE CHAPTER

- Learn about workload and how it interacts with stress
- Understand how to measure stress
- Learn about multiple resource theory

ASSESSMENT

- How does the information-processing pipeline break down when cognitive workload is high?
- What is an example of a high cognitive workload?
- How can stress be measured, and can you name three ways?

WORKLOAD

Workload refers to the amount of a human resource (e.g., attention) left to do additional work. When a task or series of tasks have light workload, there are many resources left for more work; when a task has heavy workload, there are few resources for additional work. These can be physical resources (strength, energy), cognitive resources (attention, information-processing capability), **temporal** resources (time), effort (whether the task requires a small or large amount of effort), performance (how well you are able to accomplish the task easily), and/or frustration (how emotionally easy or difficult the task is). If a task demands physical, cognitive, and temporal resources and is frustrating, a person may not be able to do any other activities, such as respond to an alert or have a conversation with a coworker. Kramer, Sirevaag, and Braune (1987, p. 146) define workload as "the cost of performing a task in terms of a reduction in the capacity to perform additional tasks that use the same processing resource."

A person's workload changes as their level of expertise changes; workload lessens as a person gains experience. For example, if a person is a new pilot, their workload is going to be much heavier when compared to an experienced pilot with many flight hours. An experienced pilot has developed a high level of automaticity as their experience has taught them exactly how to interpret different controls and different conditions. An experienced pilot's information processing is rapid and decisions are correct. A new pilot has low levels of automaticity, as each decision and action require deliberate thought. Their expertise is still being formed. A new pilot will misinterpret different controls and need to correct them as they learn the system. Hence their workload will be high.

Here is another example: Imagine that right now, you are reading this textbook on your phone in the middle of a construction zone on a hot summer day. You are on your break from your construction job, and you must read this chapter before tonight's class. Your cell phone signal is poor, and every time that you scroll to a different page, there is a significant lag. Your break is only 15 minutes long, and you are sweaty and thirsty.

In the construction zone example, it would be difficult to concentrate, as there would be a great deal of noise, the sun would be creating reflections on the phone screen, you may be sweaty, your cognitive system would be distracted constantly from the noise, and you would be watching the time so that you don't miss the end of your break. With this heavy workload, it is hard to retain information for information processing. You may experience encoding failures and do poorly on the reading quiz in class because of heavy workload. In this example of workload; as you read, you had few additional resources to process and remember the information.

According to Huey and Wickens (1993, p. 54), "The term workload was not widely used before the 1970s." Since then, there have been disagreements about workload measurement, the consequences of heavy or light workload, and what contributes to heavy workload. According to Cain (2007), workload can have three aspects: "the amount of work and number of things to do; time and the particular aspect of time one is concerned with; and, the subjective psychological experiences of the human operator" (p. 4-1). I would add expertise to this list, as experts experience different levels of workload when compared to novices. It is important to measure workload because then we can predict human performance with a system.

Workload characterizes the task in terms of the human's capabilities; as humans differ, so will their workload measurements for the same task. Typically, workload is measured in one of three ways: **subjective self-report rating**, **secondary-task measures**, or physiological measures. The most common of these is the first, subjective self-report rating scales such as the NASA Task Load Index (NASA-TLX) by Hart and Staveland (1988). This measure is freely available online and can be found here: https://humansystems.arc.nasa.gov/groups/TLX/.

The NASA-TLX was developed by the National Aeronautics and Space Administration to measure characteristics of aircraft and spacecraft. In the NASA-TLX, the human completes the self-report rating scale immediately after a particular stage of flight or immediately after a task. Workload measurements are common in aeronautical organizations, as they are interested in how difficult their vehicles are to operate. You can imagine that during particular stages of flight, pilots would have higher workload demands than in other stages of flight. This is true not only of manned planes but also of unmanned aerial systems. Often the NASA-TLX and other workload scales, such as the Cooper-Harper scale (https://en.wikipedia.org/wiki/Cooper%E2%80%93Harper_rating_scale) or the subjective workload assessment technique (SWAT; Reid & Nygren, 1988), are used to compare the performance of different system designs or different types of procedures or to determine how many operators are needed for a particular system. Sometimes, the results of these measures are used to certify a system for use by a federal agency such as the Federal Aviation Administration (Green et al., 1993).

The second way to measure workload is called the secondary-task measure or dual-task measure. This involves the person performing a primary task along with a secondary task such as riding a bicycle (primary) and listening for a word (secondary). The researcher may measure the primary or the secondary task for effectiveness (errors and mistakes) or efficiency (time). As Cain (2007) notes, the selection of the primary task is very important; it must have aspects of contextual relevance, be representative, and have a time element. The selection of what the human should do for the secondary task depends on the nature of the primary task. The secondary task must be demanding enough that it takes up the remaining attentional processing resources. The secondary task should not be something that is automatic or easy. Yet it cannot be so hard that the person cannot do the primary task.

For example, you may want to know which car's dashboard has the configuration that minimizes workload for a new driver. To measure workload, you might ask someone to recite the alphabet backward (secondary task) while driving a series of new cars at 48 mph each. It will be difficult for the person to do both tasks simultaneously, which will allow you to see which car has the easiest speedometer to read. As the person is doing the tasks, you would count how many times they are unable to keep the car at the target speed (effectiveness rating of the primary task), or you might count how long it takes before the participant states the next letter in the backward alphabet (efficiency rating of the secondary task). Often, researchers will use both methods, as some individuals will be very proficient at some tasks and not proficient at other tasks. Then both measures would be used to assess the displays.

The third way to measure workload is through physiological measures. In the same two tasks (the backward alphabet and watching the speed), you might ask the person to wear a series of sensors to measure their sweat (galvanic skin reactivity, GSR, https://en.wikipedia.org/wiki/Electrodermal_activity), heart rate, or brain patterns (electroencephalogram, or EEG, http://www.scholarpedia.org/article/Electroencephalogram). The display that produces the least reaction in performance across several participants would be the most desirable configuration. Some researchers separate workload from emotion, while others see workload and emotion as a system with an optimal amount of stress.

YERKES-DODSON LAW

There is a hypothetical amount of optimal workload for every operator and every task. This is modeled by the **Yerkes-Dodson law**, which states that if a workload is too high, performance will suffer. In addition, when a workload is too low, performance will also suffer, with more mistakes and errors. Finding the optimum range of arousal where the workload is just right is one of the tasks of a human factors engineer. More about Yerkes-Dodson law can be found here: https://en.wikipedia.org/wiki/Yerkes%E2%80%93Dodson_law.

STRESS

In short, stress is an internal subjective state based on emotions, the experience of a high workload, and other perceived pressures that cause the body's sympathetic nervous system to engage and produce a fight-or-flight response. The constant engagement of the sympathetic nervous system produces negative health consequences.

Stress and workload interact; therefore, as a person feels stressed, they perceive their workload to be higher than if they have the same task when they are not stressed. Stress magnifies workload, as it takes up additional resources and lessens the resources available to other processes. We can measure stress through the body's production of cortisol or through a survey. When measuring the production of cortisol, typically the researcher swabs the participant's mouth and measures the amount of cortisol in their saliva.

Survey measures can be subjective, as a person's perception of stress changes over situations and over time. One measure of stress, the positive and negative affect scale (PANAS) by Watson, Clark, and Tellegen (1988), measures emotions connected to stress. Another measure, the Dundee Stress Scale Questionnaire (DSSQ), and other stress measures can be found in the article by Matthews, Joyner, Gilliland, Campbell, Falconer, and Huggins (1999).

According to Hancock and Warm (2003), stress is a dynamic and changing construct. There is a trinity of stress that includes (1) environmental stress, (2) the adaptation to this stress, and (3) the output of the stress from the person back to the environment. You might think of it in this way: You come home from college, and your mom has had a difficult day at work. You want to do your laundry right away because your favorite jeans need a wash. You start to do your laundry, and your mom yells that you need to mow the lawn first. This is environmental stress. Your goals and your environment don't align. So you adapt and go outside to get the lawnmower. On your way to the shed, you step in dog poop and realize that before you mow the lawn, you must pick up all the dog poop. So you adapt again and begin to do this unpleasant task. You feel the chances of going out with your friends tonight slipping away. This creates stress for you because you are only home for two nights. You finish the lawn. Your sibling then makes a wisecrack about your new haircut, and you emit stress in their direction that starts the process over again for them.

Stress is an important component in workload because the emotion involved creates a cognitive load. The emotion hijacks some of the cognitive capacity and makes it harder to concentrate. Mowing the lawn is a simple task, but let's say that the lawnmower stopped working, and you needed to determine the problem and fix it. Because of your stressed state, it would be much more difficult to fix it than if you were not stressed at all. If you were not stressed, you would have additional cognitive resources to figure out that the air filter simply needed to be washed out.

Let's say that the same event happened to your best friend. Your friend's mother is always in a bad mood and always yells at your friend. Your friend has learned to accept this as the normal way of being for their mom. Your friend doesn't have an emotional response when yelled at and doesn't experience stress in the way that you did when your normally sweet and patient mom loses her cool.

Stress is an environmental event, a personal event, and an emittance event. The amount of negative emotion that any one person experiences as a result of stress differs among people and among events. People can be trained not to react to common stressful events and not to attach negative emotions to stressful events. If you have had military or first responder training, you have learned some of these skills.

While stress is important on its own for health and welfare reasons, in human factors, we are concerned when stress leads to changes in a human's performance with a machine or system that needs human input. If the human is so stressed that their workload is unable to accommodate any additional load, then the stress and the workload need to be examined.

MULTIPLE RESOURCE THEORY

Multiple resource theory (Wickens, 2008) identifies where performance deteriorates when workload exceeds capacity due to stress or other factors. Multiple resource theory is also an attentional theory that is discussed here: http://www.scholarpedia.org/article/Crossmodal_attention.

In the context of workload, multiple resource theory is based on the idea that we have three different stages of processing: perceiving, thinking, and responding. Then we have different codes and modalities for which the resources do not overlap. In other words, we can perceive, process, and respond to visual and auditory information simultaneously without experiencing a workload decrement because they are simultaneous. This model suggests that demand in each section differs, as do resource allocation and allocation policy (e.g., attention). Workload refers to only the first part: demand. As information is processed and the person is unable to compensate for several demands, the multiple resource theory can predict which challenges led to the overload.

REFERENCES

Cain, B. (2007). *A review of the mental workload literature*. Defence Research and Development Organisation. Retrieved April 2, 2021, from https://apps.dtic.mil/sti/pdfs/ADA474193.pdf

Green, D. L., Hart, J., & Hwoschinsky, P. (1993). *A perspective on the FAA approval process: Integrating rotorcraft displays, controls and workload*. Retrieved April 2, 2021, from https://ntrs.nasa.gov/api/citations/19940008825/downloads/19940008825.pdf?attachment=true

Hancock, P. A., & Warm, J. S. (2003). A dynamic model of stress and sustained attention. *Journal of Human Performance in Extreme Environments, 7*(1), 4.

Hanoch, Y., & Vitouch, O. (2004). When less is more: Information, emotional arousal and the ecological reframing of the Yerkes-Dodson law. *Theory & Psychology, 14*(4), 427–452.

Huey, B. M., & Wickens, C. D. (1993). *Workload transition: Implications for individual and team performance*. National Research Council.

Kramer, A. F., Sirevaag, E. J., & Braune, R. (1987). A psychophysiological assessment of operator workload during simulated flight missions. *Human Factors, 29*(2), 145–160.

Lysaght, R. J., Hill, S. G., Dick, A. O., Plamondon, B. D., & Linton, P. M. (1989). *Operator workload: Comprehensive review and evaluation of operator workload methodologies* (No. TR-2075-3). Analytics Inc.

Matthews, G., Joyner, L., Gilliland, K., Campbell, S., Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire: Toward a state big three. *Personality Psychology in Europe, 7*, 335–350.

Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. *Advances in Psychology, 52*, 185–218.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063.

Watters, P. A., Martin, F., & Schreter, Z. (1997). Caffeine and cognitive performance: The nonlinear Yerkes–Dodson law. *Human Psychopharmacology: Clinical and Experimental, 12*(3), 249–257.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50*(3), 449–455.

CHAPTER 7

DISPLAYS, MONITORS, AND SCREENS

GOALS OF THE CHAPTER

- Learn about the types of displays and their design options
- Understand the proximity compatibility principle
- Learn about the two versions of the relationship between distance and the size of a target

ASSESSMENT

- What is a card sort, and why would you do one?
- What is the proximity compatibility principle, and why do you need to use it?
- What is Fitts law, and how is it different from the Hick-Hyman law?

THE EVOLUTION OF SCREENS

For many years, we had a single monitor type: a cathode-ray tube (CRT) screen. Figure 7.1 is an example.

Modern displays and screens are much lighter and thinner. Now most screens are LCD monitors. LCD is an acronym for "liquid-crystal diode." These screens have two very thin pieces of material

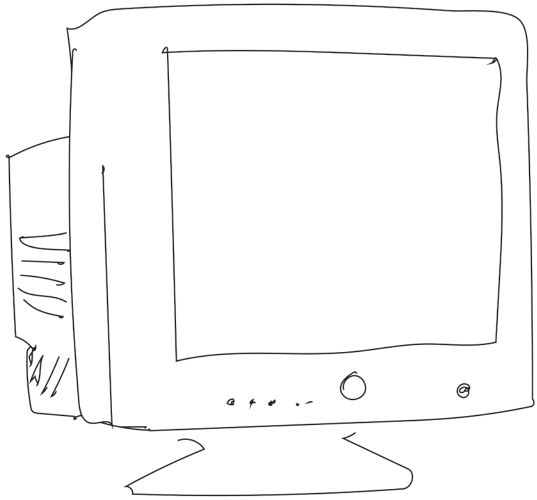

Figure 7.1. A CRT Monitor

encasing a solution. Electricity is pumped through the solution. Patterns are created with the crystals. Behind this crystal wafer is a light source. In older LCD screens, it was a type of fluorescent lighting. In newer LCD screens, it is light-emitting diodes, or LEDs (Kennemer & Waniato, 2019).

All screens have design constraints. In the case of CRT screens, these screens restricted the development of other technologies through their **refresh rates** and size. CRT screens flickered because of **refresh cycles**. A refresh cycle is when the screen redraws what it is displaying. In CRT screens, these refresh rates, or screen flickers, lead to eye fatigue. LCD screens also have flickers, but they are rarely detectable to the human eye because their refresh rate is so rapid. The second design constraint was size. The CRT screen designs must make room for a very large tube and other electronics. A modern, tiny screen, such as a watch screen, would have not been possible if LCD screens were not widely adopted.

As screen sizes became smaller, other issues were created. The first issue is real estate. In terms of monitors, this deals with the amount of stuff that a monitor can display at one time. For example, a laptop monitor is usually about 17 inches diagonally from upper left hand corner to lower right hand corner. A watch may be 3 inches as measured diagonally from upper left hand corner to lower right hand corner. The laptop monitor has much more screen real estate than a watch. To address this issue, most user interface teams will design for the smallest screen first, knowing that it is easier to scale up than down. Sometimes, user interface teams will develop unique ways to address the need for functionality in restricted screen real estate.

Interaction designers have introduced techniques to address the limited real estate problem. The first technique commonly used is the **hamburger menu**. Along with the regular menu in an application, the hamburger menu is the three to five lines in the corner of the screen. Google and Microsoft prefer to use a block of up to nine squares. Most users know that this is where the developers put things that they have no room for in the regular menu on the screen. It has taken time, but users have grown accustomed to this convention. The other concept that interaction designers have introduced is the idea of **above the fold** and **below the fold**. Above the fold is the idea that the most important items should be shown on the top of the screen first, with less important items shown as the user scrolls down the screen. Sometimes, users find it difficult to find what they are looking for in the different menus. One way to figure out what items go where is to do a **card sort**.

A card sort is an interaction design technique for menus. In this technique, people sort things into categories. They can be open card sorts (i.e., no structure is given: "Show me how you would organize these words into three groups") or closed card sorts (i.e., a structure is given: "We have three groups named home, draw, and format. How would you organize these words into these groups?"). Card sorts are based on the exemplar idea from language theory. The exemplar idea suggests that everyone who is a fluent speaker of a language has the same semantic categories. For example, sort the following words into two categories: Kenmore, General Electric, Haier, notebook, pen, stapler. Typically, people would put the first three words—Kenmore, General Electric, and Haier—together, as they are appliance manufacturers. They would put the last three words—notebook, pen, and stapler—together, as these are office supplies. It would be a rare person that would group Kenmore, notebook, and General Electric. This connection between the meanings of words and categories helps us create menus that most people can use in the common language of the interface.

Reading speed and comprehension have always been issues when it comes to screen size. The literature has been mixed, with the older literature stating that people should not read on screens, as the flicker rate causes headaches and reading issues (Gould et al., 1987). Recent literature has found that there is no loss of reading speed due to screen size (Elliott et al., 2019). In Elliott and colleagues (2019), they tested four different devices: a book, a laptop screen, a 3-inch cell phone screen, and an 11-inch tablet screen. Participants read for 10 minutes on each screen. This would be the typical length of time that someone may read a short news story or check their email. Within 10 minutes, the reading speeds across these four devices were not significantly different.

Often, we think of multiple screens as distracting; however, sometimes a person must monitor multiple screens as part of their work, such as an air traffic controller or a stock trader. Saleem and Weiler (2017) considered the usefulness of having multiple screens to monitor. They tested a single monitor, two monitors with a tablet, and then a single monitor with two tablets. Then they gave engineering students a difficult problem to solve. This problem required that they access multiple information sources. The researchers expected that the single monitor would produce the fewest errors and have the highest usability rating. However, it was the two monitors with a tablet that had the fewest errors. The single monitor with two tablets had the highest usability rating. No one chose the single monitor as their preferred mode, and it did not have the lowest error rate.

The researchers concluded that having several tasks viewable at a glance helped the engineering students solve the problem. The research question of having multiple monitors as distracting or not was unresolved. This study introduced additional research questions, such as which types of screens—tablet or desktop—influenced problem-solving. Another question that remains unresolved is the difference between using screens to gather information and monitoring screens for changes.

Sometimes a person operating a complex system, such as an unmanned aerial system (UAS), would need to monitor more than one screen at a time. Studies in human factors have determined that four screens are the maximum number to be monitored simultaneously in terms of information processing ability and workload. There are instances where more than four screens are required for a task. In those cases, the human factors engineer has a few additional tools to help maximize human information-processing ability and minimize workload and human error. The first is the proximity compatibility principle.

The **proximity compatibility principle**, derived by Barnett and Wickens (1988), states that tasks that require information integration from more than one screen are easier when the displays are nearby physically. Conversely, tasks that require information separation or focus on a single aspect while ignoring other aspects become easier when the displays are far from each other. For example, let's go back to the driverless car example. While driverless cars are technically driverless, many require a human to help monitor the system state in case of sensor malfunction or error. The human monitor will presumably be watching out the windshield for unanticipated changes in the environment. A heads-up display that is projected onto the windshield helps the human monitor the fuel level and other operating characteristics.

Wickens and Carswell (1995) discussed this principle. The proximity compatibility principle states that when items on the screen need to be accessed serially or simultaneously, they should be grouped together (i.e., in proximity to each other). For example, if a UAS ground operator is responsible for take-off, loiter (i.e., flying in place), and landing, these controls should be grouped together on the screen, as they will be needed in sequence. The ground operator may have other controls, such as battery level, gyroscope health, sensor health, headings, and speed. The controls that the ground operator uses together most often should be grouped together. For example, it would violate the proximity compatibility law to have the sensor health and the speed grouped together, as these are different tasks during flight. Grouping the battery level and speed would make sense and support the proximity compatibility law. The proximity compatibility law is one of the reasons that in large-scale UAS operations, there is more than one operator: a ground operator, a camera operator, and an outside observer. This division of labor in addition to the division of displays has helped reduce the number of crashes and increase safety.

Some of the reasons for the division of labor were found by Wickens and Andre (1990). One of these reasons was when operators view more than one display, their tasks change from a "viewing the information" task to an "integrating the multiple sources of information and determining the combined meaning" task. In other words, when a UAS ground operator would view the camera display as well as the vehicle display, they would fly with the camera display's **point of view** (POV) rather than the vehicle display's POV. In many cases, the camera display's point of view was very small and restricted to underneath the vehicle. As operators focused on the camera display, they would be concentrating on the small area visible through the camera and ignore the large building they were about to crash into with the vehicle. When there were multiple screens for the camera POV, the vehicle POV, and a ground observer POV, the operator would prefer a particular screen and use it as a primary place of monitoring while ignoring the other screens even when the other screens would be better for flight.

When more than one display is present, a person's monitoring task changes from monitoring a single screen to monitoring several screens with multiple meanings. Attentional processing as mentioned in the prior chapters becomes challenging. When a meaningful change takes place on any of the screens, extracting that meaning and that variable without distortion or bias from surrounding information is challenging. When screens are nearby in physical proximity or mentally grouped together, these three tasks are easier.

With the advent of touch screens, another law affects how quickly a person can react by pressing a button or making a selection. This is **Fitts law** (Fitts, 1954). Fitts replicated an 1899 study by Woodworth (1899) demonstrating that when a person took a pencil and connected three points into a triangle, the person would speed up the more often they did it. This led Fitts to explore the relationship between size, speed, and accuracy. According to Fitts law, the time that it takes to move from one point to another is determined by the distance between these divided by the width of the point that someone is attempting to select. It can be expressed as $T = C_1 + C_2\log_2({}_2D/W)$, where T equals time, D equals the distance, and W equals the width of the items that a person is intending to reach.

There are similar laws about the relationship between distance and target size, such as that from Schmidt, Zelaznick, Hawkins, Frank, and Quinn (1979), who described this relationship as linear rather than logarithmic. Their law can be expressed as $W_e = K_1 + K_2\ (D/T)$, in which T is a specified amount of time, W is the standard deviation of the movement distances several trials would produce, and D is the distance between the targets. Schmidt and colleagues suggest that their equation is accurate when the movements are very brief (under 200 milliseconds), and Fitts law compensates for longer movements when individuals make small corrections in the direction to reach the target, thereby creating a logarithmic function instead of a linear function. These types of investigations highlight the relationship between speed and accuracy called the **speed-accuracy trade-off**.

Other researchers found support for Schmidt and colleagues (1979) when the time was under 200 milliseconds and support for Fitts (1954) when times were over 200 milliseconds (Wright & Meyer, 1983). Still others found that the larger the diameter of the object, the faster a person could transcend a distance with the hand opened wide to compensate for over- or underreaching (Bootsma et al., 1994). All of these studies point toward the same finding: larger things are easier to grasp and easier to arrive at from a far distance. For example, think of a baseball diamond. If the bases on the diamond were larger, it would be easier to see and easier to arrive at the right target than if the bases were small. Fitts law is the most popular of these studies and is commonly referred to when practitioners or researchers want to explore the speed-accuracy trade-off or the size of a target button or location. In practice, professionals use Fitts law to determine the size of buttons on large interactive touchscreens or buttons on control panels for complex systems, such as nuclear reactor control panels. Often Fitts law is combined with the Hick-Hyman law of choice to determine the number, size, and configuration of buttons (Pennetti et al., 2019).

REFERENCES

Barnett, B. J., & Wickens, C. D. (1988). Display proximity in multi-cue information integration: The benefits of boxes. *Human Factors, 30*(1), 15–24.

Bootsma, R. J., Marteniuk, R. G., MacKenzie, C. L., & Zaal, F. T. (1994). The speed-accuracy trade-off in manual prehension: Effects of movement amplitude, object size and object width on kinematic characteristics. *Experimental Brain Research, 98*(3), 535–541.

Elliott, L. J., Ljubijanac, M., & Wieczorek, D. (2019). The effect of screen size on reading speed: A comparison of three screens to print. In W. Karwowski, T. Ahram, & S. Nazir (Eds.), *Advances in human factors in training, education, and learning sciences: Proceedings of the AHFE 2019 International Conference on Human Factors in Training, Education, and Learning Sciences* (Vol. 963, pp. 103–109). Springer. https://doi.org/10.1007/978-3-030-20135-7_10

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*(6), 381.

Gould, J. D., Alfaro, L., Finn, R., Haupt, B., Minuto, A. (1987). Reading from CRT displays can be as fast as reading from paper. *Human Factors, 29*, 497–517.

Kennemer, Q., & Waniato, R. (2019, October). *LED vs. LCD TVs explained: What's the difference?* Digitaltrends. Retrieved December 23, 2019, from https://www.digitaltrends.com/home-theater/led-vs-lcd-tvs/

Pennetti, C. A., Hollenback, K., Kim, I., & Lambert, J. H. (2019, November). Cognitive load variability from road characteristics should influence a safety requirement for vehicle stopping sight distance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 2129–2133). SAGE.

Saleem, J. J., & Weiler, D. T. (2017, September). Multiple screens and devices in an information-rich environment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (p. 361). SAGE.

Schmidt, R. A., Zelaznik, H., Hawkins, B., Frank, J. S., and Quinn, J. T., Jr. (1979). Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychological Review, 86*, 415–451.

Wickens, C. D., & Andre, A. D. (1990). Proximity compatibility and information display: Effects of color, space, and objectness on information integration. *Human Factors, 32*(1), 61–77.

Wickens, C. D., & Carswell, C. M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors, 37*(3), 473–494.

Woodworth, R. S. (1899). The accuracy of voluntary movement. *Psychological Review, 3*(3). i–114.

Wright, C. E., & Meyer, D. E. (1983). Conditions for a linear speed-accuracy trade-off in aimed movements. *Quarterly Journal of Experimental Psychology, 35*(2), 279–296.

CHAPTER 8

USABILITY

GOALS OF THE CHAPTER

- Learn about the product development life cycle
- Understand how usability is measured
- Learn about the types of measurement and design guidelines

ASSESSMENT

- What are the three constructs measured in usability?
- What is the ISO?
- What is an MVP?

As I write this chapter, I am at a conference with only my cell phone and keyboard. Although the type is very tiny, I am able to do incredible things that weren't possible many years ago. When computers were first introduced, only the government, businesses, serious hobbyists, and those with advanced degrees used them. This is because a person needed to write the code to make it work, and many of the applications that we use daily had not been invented yet. Computers analyzed large data sets, made complex calculations, and kept databases (https://en.wikipedia.org/wiki/History_of_computing_hardware).

Researchers realized the tremendous potential. Computer manufacturers, such as Apple, were pioneers in making computers that anyone could use. Making technology so that people without computer science training can use it is referred to as the construct of usability. We like to think of Don Norman (https://en.wikipedia.org/wiki/Don_Norman) and Jakob Nielsen (https://en.wikipedia.org/wiki/Jakob_Nielsen_(usability_consultant)) along with Alan Cooper as the pioneers of usability and interaction design (https://en.wikipedia.org/wiki/Alan_Cooper). Interaction design is the way that the interface reacts to the user (https://www.interaction-design.org/literature/article/what-is-the-difference-between-interaction-design-and-ux-design). Usability is a construct that encompasses how easy a computer is to use, how easy it is to remember, and how easy it is to learn, according to Hornbæk (2006). There are other components that have evolved since the first computers and first usability pioneers started their work. For example, now most people use computers on a small screen on their cell phones. Initially, nearly all computers were desktops with large monitors. As computer use has evolved, so has our understanding of usability, ease of use, and learnability. However, the way that we measure usability has remained the same.

MEASURING USABILITY

Depending on where the product is in the **development life cycle** (https://en.wikipedia.org/wiki/Product_lifecycle), the methodology changes. Usability investigations can begin at the product idea's inception and continue past the product release. Initially, usability was thought to correspond only to consumer products. However, industrial products, such as medical devices, are also subject to usability studies. There have been cases where consumer or industrial accidents have prompted lawsuits for an unusable device or a device that encourages injury. An entire organization, Exponent (https://www.exponent.com/), is based on the idea that usability consultants are needed in litigation.

There is an international organization that sets usability standards for engineered products: the **International Organization for Standards** (ISO). ISO sets forth definitions and standards for the functionalities of different classes of products. Usability is ISO section 9241, which is here: https://www.iso.org/standard/60476.html. These standards can be purchased by an organization or found at a library.

Typically, usability is thought of in terms of three constructs: **effectiveness**, **efficiency**, and **satisfaction**. Each product or system should allow the user to attain their intended goal and to do so in a reasonable amount of time, and the user should be happy with the interaction. Much of the literature in the field of usability also discusses learnability and error recovery. While these are also important, this chapter will focus on effectiveness, efficiency, and satisfaction.

DEVELOPMENT CYCLE

Let's discuss a development cycle for a moment. A product begins with someone's idea. The context in which the idea is generated matters a great deal. Here are four different contexts in which an idea happens: it could be (1) a person who has an idea for a better-designed pen during a classroom project, (2) an individual at home who likes to tinker who comes up with a better design for a pen, (3) someone who is participating in an entrepreneurial event such as **start-up week** who develops a superior pen idea, or (4) someone who works for a pen company who comes up with a novel design. Here we have four different scenarios as to the likely development path of the idea. Available resources will dictate how the idea is developed. Intellectual property (IP) concerns who owns an idea and how it can be shared. In the first three cases, the person with the pen idea is not working within a formal organization. The persons in the first three cases would want to be familiar with the laws in their state before sharing the idea. In the first three cases, it is most probable that the person who had the idea would own the idea once they created a recording of the idea (a drawing, a written document, or an audio description). Once the idea is fixed in some format, the idea is automatically copyrighted to the person who fixed it in the format (https://www.copyright.gov/help/faq/faq-general.html). In some cases, the individual would want to have a **nondisclosure agreement** (**NDA**) document made by a lawyer to protect the idea, as they shared it with others for developing a **proof of concept** (**POC**) or **minimally viable product** (**MVP**).

The fourth case is the most common, where an individual creates an idea within a formal organization. We'll focus on this scenario. Once the idea has been created and can be described verbally, the organization that this person works for owns the idea. For example, this textbook is copyrighted by me and Penn State, but they are made available to you with a certain type of usage license. Penn State owns the textbook, as I currently am employed by Penn State. If someone uploaded the textbook to their website and claimed that they wrote the textbook, they could be liable for intellectual property theft. In this case, Penn State would be the party prosecuting the theft, even though I wrote the textbook.

Commonly, these terms are agreed to when the person signs their initial employment paperwork. Any resources used by the person during the idea creation and execution along with the formal employment agreement allow the organization to own the idea. While this makes many people uncomfortable initially, it can be one of the best benefits when working in an organization. This arrangement allows the organization to provide resources to develop, promote, and sell the idea for the product, and it negates the need for an NDA. As the product idea develops, other professionals within the organization, such as usability professionals, can work on the idea. Once the product idea has been formalized and the organization has decided to put resources toward developing the product idea, a team is formed. This team typically has a project manager, a marketing person, a usability person, and a developer/designer/engineer. Additional personnel may be involved at different stages of the project.

Let's say that the product idea has been created, the development team has been assembled, and work is to begin on the new project. The development team will meet, and the usability person will immediately start to find out information about the product idea and will ask key questions of the person who had the product idea:

1. Who will use the product or system?

2. What will they use it for?

3. Are there any similar products or systems already in use that do something similar?

4. What frustrations do users have with these current products or systems?

5. What do users like about these current products or systems?

6. Why is this idea a good one?

For each question, the usability person will ask additional questions to find out some core facts:

1. Who are all the users of the product/system (their demographic information, salary range, education level, marketing likes and dislikes)? Who are the secondary and tertiary users of the system, and what are their demographics?

2. When does the primary user use the system/product and for what? When do the other users use it, or when does it affect them?

3. What is the competitor doing in this market, and who are they?

4. What are the frustrations that users have with any of the competitors' products? These are pain points or explain why we should make this new product/system.

5. What are the things that users love about current competitors' products? These are opportunities and what people value in the current way of doing things.

6. Why should we put company resources toward this idea, and why is this product idea one of the most difficult to address? Often the usability person will use a technique such as ***the five whys*** to find out the core value proposition. *The five whys* is a common technique that is explained further here: https://agileleanlife.com/5-whys.

Once these questions are answered, the usability person has the basis for developing an experience design strategy. The usability person will obtain the competitor's products/systems or gather information on them. Then, they will organize a group of users that represents the key stakeholders in the product. As the developers/designers begin working on the functionality of the project/system, the usability person will take the initial specifications and begin to determine additional specifications. Specifications determine what the product/system does and what it does not do. There are several different types of specifications, including user specifications, technical specifications, physical specifications, and so forth. Specifications are commonly referred to as the **functional product specifications** (**FPS**). This can be a picture, a single document, or several documents of specifications.

SPECIFICATIONS

The FPS can be as simple as a wall full of sticky notes. There are several ways to approach the development of an FPS. Some user-experience people prefer to create a task analysis first and then derive the specifications from that. Others prefer a meeting with all the main stakeholders in the room. Finally, some organizations prefer a room of users to devise specifications with the usability (UX) person as the lead.

After the FPS and other specification documents are finished, the team knows what they are building. How they choose to build it is up to the individuals and company culture. Different organizations will try different development approaches according to the types of products they create. Other organizations will use agile or lean exclusively. Some may rotate project managers and developers, while others arrange the development teams hierarchically with the most desirable products that require the most expertise available only to the most experienced development teams. Typically, a usability person or UX person will be assigned to a single product for an extended period of time so they get to know the usability problems endemic to that product and its users. The usability person may work with a variety of developers as they cycle through product teams.

After the specifications have been developed and agreed upon by all stakeholders, the UX person will work on the information architecture as they shape the experience that the person has when using the product/system. In large organizations, the UX person will work with a visual designer or an interaction designer to produce the first prototypes. They will do this by reviewing what others have done, devising an optimal flow, and then prototyping the flow of one or two main tasks that a user would do with the product using a paper prototype. They may go out to the field (local coffeehouse) and test the prototype on potential users of the product. The prototyping apps mentioned below help bring a consistent look and feel to all the applications that use the same operating system platform (e.g., iOS or Android) or are from the same organization. Different organizations will have their own **pattern library**. These templates have the standard buttons and menus that users already are accustomed to using.

USER TESTING

When the UX person tests the users, they will be asking what they liked and didn't like about the product. Once the UX person has started to hear the same comment from two or more participants, they can go back and fix this issue in the prototype. Then they may test again or add in a few more tasks and then test. This test and retest portion of the development cycle should continue until the UX person has developed the interface enough that it communicates well with the users and does what it is designed to do. That's what is called a minimally viable product (MVP). Once the UX person has accomplished a working MVP prototype in which the users do not find problems, they can turn it over to the development team. In reality, the process is more complex, with lots of discussions about adding and deleting functionality. The user testing never goes as planned. Developers find that what the UX person tested and what the users want is impossible to implement. Other complexities happen that confuse the process. Understanding how humans think and act can be critical to helping the process go well on both the development side and the user-testing side. Communicating between all parties and understanding all parties' unique points of view are critical in UX. Without empathy, a UX person may be creating arguments instead of solutions.

Work in UX is rarely complete. Once the product is developed and released to the public, there will be additional problems. Customers will complain about a certain function, or they will want additional functionality. Again, the UX person will prototype, test, and update the FPS according to their testing. Then they will recommend product improvements or reductions. Sometimes complaints appear that are unique to a certain type of user or a certain type of use. At that point, the development team may decide that a separate product is warranted, and the process begins again.

There are additional tools that the UX professional can use to determine specific types of problems, improve product/system organization, or improve functionality. Many of those techniques can be found here: https://www.usabilitybok.org/managing-ux. Stanton and colleagues (2013) have an excellent reference guide on different methods for human factors and UX as well.

Usability is intangible; hence an organization will occasionally need the UX designer to justify the expense of their salary and worth on the team. The organization will want to know the **return on investment** (**ROI**) of the work. Here is a good article on how to gauge that worth for a particular product (https://measuringu.com/ux-roi/). Consistently, organizations find that it is worth the trouble in reduced litigation, increased sales, innovative gains, and fewer customer complaints (https://www.experiencedynamics.com/blog/2014/07/making-strong-business-case-roi-ux-infographic).

Here are some additional links and prototyping applications:

https://www.usability.gov/

http://boxesandarrows.com/

https://uxmastery.com/resources/books/

https://design.google/

https://www.invision.com

https://www.sketch.com

https://www.axure.com

https://www.figma.com

REFERENCES

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies, 64*(2), 79–102.

Stanton, N. A., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C., & Jenkins, D. P. (2013). *Human factors methods: A practical guide for engineering and design* (2nd ed.). Ashgate.

CHAPTER 9

TEAMS AND PERFORMANCE

GOALS OF THE CHAPTER

- Learn about trust in teams and how trust shapes team communication
- Understand how trainers take individual differences into account when training
- Learn about the connection between outcome-based performance measures and errors

ASSESSMENT

- How do surgery teams address mental model differences?
- What is implicit communication in teams?
- What is an outcome of punishing the weakest link in a team?

Research on teams is usually focused in the research area of industrial and organizational psychology—as in, how teams of people work together. However, teams also work with and through technology. In this sense, team research through computer-supported cooperative work is a human factors research area. Technology can enhance team communication or hinder it. One of the first organizations to conduct research in **teaming** was the U.S. Army.

The U.S. Army was interested in knowing more about teams of soldiers on the battlefield. Here is an example of a scenario: Some soldiers are near the communication tower working on a tank as a team. Another team of soldiers is waiting on the battlefield for the tank to get refueled and pick them up. The soldiers working on the tank need to communicate with each other, but they are nearby one another and can just tap each other on the shoulder to get each other's attention to talk. The soldiers on the battlefield don't know how long it will take until the tank arrives, but they need to know, so they use their radios to connect to the tank soldiers. The waiting soldiers experience gunfire and are unable to communicate this to the tank soldiers. The tank soldiers are unaware of the sudden urgency. You can imagine the confusion on both sides, the anxiousness, and the need for good communication within and between both teams. The soldiers working on the tank presumably have no trouble with communication because they are nearby one another, but that may not be the case. They must have a good shared knowledge of the task that they are doing and the situation. Sometimes the soldiers who are near one another have more trouble communicating than the soldiers who are far from each other because of collaboration, trust, and coordination.

While military teams were some of the first to be examined, healthcare teams pose a great risk to human safety. It is estimated that up to 7,000 deaths per year happen because of medication errors (Caldwell, 2017). These errors can happen as a result of poor team communication and coordination.

Teamwork in medical settings can be challenging. Not only must the individual create their own mental ideas or mental model of the situation and the work that needs to be done, but in a team, this mental model is shared among others who have different perspectives, expertise, background knowledge, and communication styles. Individuals also have their own workload; working as a team can increase or decrease the different types of workloads as communication is either enabled or inhibited, which affects collaboration, cognition, and coordination. Individuals also have workload; working as a team can increase or decrease the different types of workload as communication is either enabled or inhibited, which affects collaboration, cognition, and coordination. Recently, surgery teams began to implement a checklist system to address these issues. In the checklist system, the lead nurse reads off who the patient is, what they are having done, and other items. Everyone on the team checks to be sure that they have the right patient and the right tools for the procedure and know who each member of the team is and who is performing the procedure and how it will be performed. This basic coordination has been responsible for a decrease in surgery errors.

Different tasks require different coordination demands. For example, Sellers and colleagues (2014) asked teams of rugby players to evaluate and choose a candidate for public office. In a subsequent study by Greenlee, Funke, and Rice (2019), the researchers asked participants who were not previously part of a team to engage in an evaluation task in pairs. This task required communication, discussion, and decision-making on the team level as well as the individual level. Sellers's study participants were used to working with each other; the Greenlee and Funke participants were a part of a new team. Both Sellers and Greenlee and Funke measured the perception of work for each member of each team. Perception of work is also called workload.

Greenlee, Funke, and Rice were interested in measuring how much work the team part of the task encompassed (2019). They found that the workload included mental demand, physical demand, temporal demand, effort, emotional demand, and performance monitoring. The team workload was composed of communication demand, team coordination demand, and team performance monitoring. There was a third area: team task balancing. This was composed of team support, team emotional demand, and time-sharing demand. Overall, Greenlee and Funke found that large, established teams such as the team of rugby players had a different distribution of workload compared to a team of two individuals and that the teams perceived workload differently.

This perception could be due to communication needs. In teams, members need to communicate verbally and nonverbally. Communication-related behavior is the primary way that teams interact. The effectiveness of any communication effort can be measured through the subsequent actions and the duration of the communication. The duration and nonverbal communication would be considered communication flow, while the content is, of course, content. Communication influences the performance of the team. A team that shares sparse information and cannot anticipate the information needs of others on the team will not perform as well as a team where the information is rapid between members.

Communication in teams differs by the proficiency at which the teams share information. As there is more interdependence on teammates as the task complexity increases, this becomes more important. There are two types of communication: implicit and explicit. Implicit communication is a shared understanding of what information will be needed by each party. Explicit communication is the sharing of that information with others when the other individuals on the team ask for it. Implicit understanding of what needs to be shared in the team depends on a good mental model of what the team needs to do and how everyone will contribute. It also depends on good **metacognition**.

Metacognition is the understanding of what is known, unknown, and yet to be known in a problem set. As the individual begins to understand the problem set, they will derive what they know and what they need to know in order to solve. In complex problems that require team coordination, the individual needs to consider not only their own information needs but also the information needs of the team. Highly effective teams have individuals with good metacognitive ability and individuals who can anticipate the information needs of others (Butchibabu et al., 2016). In order for information to flow rapidly, there must also be trust. Trust is based on a person's previous performance with other individuals and with the particular individuals on the team.

Trust is based in social cues. As individuals make up a team, the interaction between the individuals is key in making the team work well. As people on the team ask for information, they can do it in a supportive manner and encourage trust and collaboration. Or they can ask for information in a demanding, accusatory manner. There are nonverbal cues for trust as well, such as smiling or disgust.

An individual's level of social interaction anxiety based on a personality trait or based on previous experience influences their ability to function as a good team member. If many individuals on the team have a high level of social interaction anxiety, this can impair the team. Interaction anxiety is when a person perceives their behaviors and responses in connection with social interactions or social contingencies. As team roles are more visible, people who are prone to interaction anxiety are more likely to experience the anxiety.

Naber, McDonald, Asenuga, and Arthur (2015) examined this in conjunction with training to find the best way to train when individuals have social anxiety. Teams of people can be trained in two different ways. Individuals can be trained and be held accountable to individual criteria, or the entire team can be trained and be held to a team criterion. For example, when training a group of U.S. Navy submarine operators, each operator will need a certain level of proficiency individually and in collaboration with others. How you train that proficiency is a combination of individual training and team training. Some trainers may require that individuals train until they are proficient at their portion of the task before joining the team. Other trainers may require that individuals become minimally proficient and then train the entire team until it is proficient. One of the determining factors in this decision is the amount of interaction anxiety the trainer perceives in the group (Naber et al., 2015).

When mistakes happen in a team environment, often organizations will search for the "weak link," or the person who was responsible for the error. This could be a mistake, as it sends a message to the entire organization that errors are not tolerated. However, to be human is to make errors. While this should never be used as an excuse on a personal level, it needs to be reconciled on the organizational levels when supporting teams. Team communication, coordination, and trust need to be at their optimum for teams to use technology effectively. Teams of people are required to solve the complex problems of today. Interdisciplinary teams are becoming the normal way to approach work. An understanding of the importance of good collaboration skills enhances human performance.

REFERENCES

Butchibabu, A., Sparano-Huiban, C., Sonenberg, L., & Shah, J. (2016). Implicit coordination strategies for effective team communication. *Human Factors, 58*(4), 595–610.

Caldwell, R. B. (2017). *The influence of human factors in medication errors: A root cause analysis*. QSEN Institute. Retrieved October 18, 2019, from http://qsen.org/the-influence-of-human-factors-in-medication-errors-a-root-cause-analysis/

Greenlee, E. T., Funke, G. J., & Rice, L. (2019). Evaluation of the team workload questionnaire (TWLQ) in a team-choice task. *Human Factors, 61*(2), 348–359.

Naber, A. M., McDonald, J. N., Asenuga, O. A., & Arthur, W., Jr. (2015). Team members' interaction anxiety and team-training effectiveness: A catastrophic relationship? *Human Factors, 57*(1), 163–176.

Sellers, J., Helton, W. S., Näswall, K., Funke, G. J., & Knott, B. A. (2014). Development of the team workload questionnaire (TWLQ). *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58*, 989–993.

CHAPTER 10

SITUATION AWARENESS

GOALS OF THE CHAPTER

- Learn about situation awareness and why it is so important in operators of complex systems
- Learn about metacognition and how it is related to situation awareness
- Learn about the relationship between situation awareness and metacognition

ASSESSMENT

- Which cognitive processes compose situation awareness?
- What are the situation awareness processing levels, and can you give a real-life example of each?
- How is SA measured?
- Why would people say that you cannot have SA without metacognition?

Situation awareness (**SA**) is critical to human performance. SA is the knowledge of what is currently happening, what has happened, and what could happen in a situation or environment. Sometimes SA is confused with metacognition. However, SA is different from metacognition in that SA is flexible knowledge of the dynamic situation or environment. Metacognition is the knowledge of what you know as a person. For example, the metacognitive fact that Hondas take 89-octane gasoline is very different from the situation awareness that your Honda is currently in your garage.

Knowing the situation that your Honda is currently in your driveway helps you to decide if you are going to take the bus to school or drive the car to school. Other things may also influence that decision. In Erie, Pennsylvania, it often snows overnight, and when it snows, quite a bit of accumulation will make streets impassable for a small car such as your Honda Accord until the snowplow cleans the streets (metacognitive knowledge). You may want to find out the current situation: Did it snow last night? How much? In this way, SA is also domain specific. While it snows quite a bit in Erie, Pennsylvania, this is not common in Harrisburg, Pennsylvania. An Erie native who transfers to another campus would change their anticipation of the amount of snow over time.

It is possible to lose SA. Maybe you slept late and are in a rush. You put on your clothes in a hurry and grab a cup of coffee, only to find out that you cannot get out the front door because the snow drifted up to about 2 feet over the door. You forgot to check the weather last night, and this impacted your SA, which impacted your ability to get to school on time.

As attention is diverted to other tasks, the initial task that we were monitoring may change. This is called changing states during ongoing SA. The primary task changes state, and the change in state affects other tasks that were dependent on the first task. Driving is a good example. Driving requires a high level of SA, as the exterior and interior environment is constantly changing. As you are sitting at a stoplight, let's say the passenger spills their drink all over your seat. As you rush to find the towel that you usually keep in the driver's-side door, you fail to see the stoplight change from red to green (failure of SA as an ongoing assessment). The car behind you honks, and you look up and see the problem. You tell the passenger that you don't remember if you still have a towel in the car (SA as ongoing thinking process), and you ask them to look in their door pocket to see if it is there.

Sometimes you don't have enough **cognitive resources** to devote to the current matter at hand. SA is directly linked to working memory as well as attention. In the same car with the spilled drink, let's say that you also have three children in the back seat. The children are fighting over the new sweatshirt that one of them is wearing. As you focus on the spilled drink and the stoplight, you hear the sweatshirt tear but fail to act. As you are focusing on the stoplight and your wet seat, you do not have the cognitive resources to recognize that the back-seat battle has become intense.

Just as in attention, in SA, there are processing levels. There is noticing, understanding, and anticipating the next event. You noticed the spilled drink, as it was right next to you. You failed to notice the torn sweatshirt, as it was behind you, and other things that were more urgent were happening at the same time. Understanding is the processing of the meaning of the item. The spilled drink may damage the car seat. Anticipation is the projecting of what will happen next—the car seat may smell bad when the car gets hot in the summer. Your SA in this case has prioritized the spill over the commotion in the back seat and prioritized the driving over the commotion in the back seat. When a parent states, "If you don't stop fighting, I am going to stop this car," they are really stating a SA fact—that they cannot supervise and drive at the same time.

While we see many examples of SA in real life every day, SA becomes important when we are working with computers or other systems. In older video game systems, you may have been aware of a controller lag. If you pressed the button, it took a few milliseconds for the character on the screen to do something. This lag was predictable, and as you practiced with the system, you built this lag into your reaction time. You compensated for it. Still, because of the laws of physics, there is a lag between when we type in a letter on a keyboard and when it appears on the screen. We become accustomed to this lag when it is consistent, and we incorporate it into our expectations and performance.

Sometimes the lag is unpredictable, such as when the internet is down at school or when a computer needs more memory in the system. When this happens, you anticipate what is supposed to happen next but wait. You may not know what the system is doing and why it is taking so long. This is when SA is impacted. Often these are low stakes; you simply wait for the computer to catch up or for the internet to come back online. In other systems such as large mining equipment, a system lag can affect safety as well as performance. As a designer of systems, it is imperative that the system compensates for times when it lags and the operator does not know what is happening.

MEASURING SA

As in usability testing, SA is also measured using **real operators of the system**. In these measurements, the operator will be asked to do a real or virtual task. At certain intervals, the screen or the system will stop or go blank, and the operator will be asked questions that measure if they noticed, understood, and anticipated the next events or states of the system. One of the most popular of these measures is the situation awareness global assessment technique (SAGAT; Endsley, 1995). In the SAGAT, the operator recounts the current state of the system as they understand it.

When several operators fail to recognize a particular system state or a particular item in the environment, then the designer knows that something must be done. The system designer may choose to make a particular display larger or brighter or locate it somewhere different. For example, a person drives three different models of cars. During the tasks, they are stopped and given the SAGAT to measure their understanding of the controls in the cars. They and other drivers have all failed to notice that there is almost no gasoline in one particular car model. The system designers would want to redesign the gas gauge in that particular model by placing it in a different area on the console, making it a different color, or displaying it in a way that attracts more attention. The system designers might try three different interface designs and conduct usability testing and SA measurements before incorporating the top design into a new dashboard.

METACOGNITION

Metacognition goes with SA. Literally, you cannot have one without the other. As people monitor their environment, other people in their environment, and themselves for changes in the environment, they are also assessing what they know and don't know along with what others know and need to know. According to O'Neill and Abedi (1996), there are four subcategories of metacognition: awareness, cognitive strategy, planning, and self-checking. In awareness, a person must be aware of what is known and not known. In cognitive strategy, that person must have a strategy to find the information and a plan on where to find it. Then in self-checking, they see if there is additional information needed and how to fill in those areas.

According to Vogel-Walcutt and Fiore (2010), when investigating training systems that teach high-level skills, delivering the training into sections and then providing deliberate practice with feedback are the best ways to develop overall metacognitive skills. One of the better ways to incorporate deliberate practice is to prompt trainees at different stages of the exercise with what they know, what they don't know, and what they think that they need to know. When prompted to answer questions in this way, the training instills the practice of self-monitoring and self-assessment. Then once the training is over, providing the trainee with a realistic assessment of their performance and ways that they could improve helps them reflect on the process and assess what they could have done differently. These build in a concept called strategizing, which is part of metacognition.

REFERENCES

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors, 37*, 65–84.

O'Neil, H. F., Jr., & Abedi, J. (1996). Reliability and validity of a state metacognitive inventory: Potential for alternative assessment. *Journal of Educational Research, 89*(4), 234–245.

Vogel-Walcutt, J. J., & Fiore, S. (2010, April). Insights from empirical metacognitive research. In *Proceedings of the 2010 Spring Simulation Multi-conference* (p. 33). Society for Computer Simulation International.

CHAPTER 11

EMOTION, MOTIVATION, AND BOREDOM

GOALS OF THE CHAPTER

- Understand models of emotion and motivation
- Learn about how these models predict and influence engagement with technology
- Understand the connection between boredom, motivation, emotion, and technology

ASSESSMENT

- Why did most human factors practitioners view the theories of emotion and motivation as simply theories unrelated to technology use?
- What is the technology acceptance model, and how does it influence technology use?
- What is the rational actor model, and how does it influence technology use?
- What are the two types of boredom, and how do they occur?

Emotion is our human system of attributing feelings to events, things, behaviors, and situations. Evolutionary psychologists believe that emotion is protective and saves our species from harm by allowing us to attribute the good and bad to otherwise meaningless events and behaviors (Plutchik, 1980). Emotion is complex, with meaning, drives, and "intelligences." Emotion drives behavior, attention, and motivation. It is from this perspective that emotion relates to human factors. As people derive emotion from interacting with a system, either they are motivated to use the system again, or they are not.

MOTIVATION THEORY

There are several theories of motivation, or what drives us to behave in ways that we do. One theory, **drive theory** (Hull, 1943), is based on the idea that when we have a deficit or a need, this need motivates us to solve it. For example, when we are hungry, we are motivated to eat. However, there are some problems with this theory: we also eat when we aren't hungry. In response, other researchers derived additional theories.

Another theory is suggested by **Maslow's hierarchy of needs**. Maslow (1970) suggested that we have needs, but a single behavior can satisfy more than one need. Basic needs such as hunger, thirst, and warmth are the most basic and frequent, so they must be met first in order to progress to the next level of needs. Next is the need for safety, then social needs, then the need for esteem, and then finally self-actualization.

People move through the needs as they progress through life and as they progress in their profession. However, drive theory and Maslow's theory don't explain why we use technology. Up until recently, most human factors practitioners viewed emotion and motivation as basic psychology theories that did not apply to technology. They viewed emotion as a design challenge, motivation as an organizational challenge, and boredom as an automation challenge (Szalma, 2014).

For example, if a particular technological design was engaging, it would tap into our emotions, and we would desire it. If we were unmotivated to use technology, it was due to stress at work or the wrong combination of incentives to excite us to action. If we were bored during the use of a technological system, it was because the automation was not adaptive. Now we have a better understanding of how technology engages us on an emotional level throughout our life and motivates us to use it.

According to Szalma (2014), motivation is direction, energy, and persistence when working toward a goal. In order for a human to desire to use a system, the system must be effortless, helpful (not frustrating), rewarding, and fulfilling. In some systems, human motivation is a goal superior to system performance.

For example, imagine using the current **learning management system** (**LMS**) that is used at your campus (e.g., Canvas). If students and professors are not motivated to use the system because it is frustrating, unrewarding, and stressful, it does not matter how well the system works. Moreover, the type of task that you are doing with the LMS affects your motivation to use it. This happens on both the instructor and the student side of the LMS.

For example, as an instructor, if I were to create an assignment on Canvas, it would take about nine steps. The same assignment distributed using paper in class would take two steps. I am more motivated to not use Canvas for assignments because they are time intensive. However, as a student, I would be very motivated to receive all my assignments through Canvas because it organizes my assignments, shows me the due dates, and shows me my grade. It takes a student two steps to submit an assignment through Canvas. The willingness to use Canvas or any LMS can be modeled with the **technology acceptance model** (**TAM**) and the **rational actor perspective** (**RAP**).

THE TECHNOLOGY ACCEPTANCE MODEL AND THE RATIONAL ACTOR PERSPECTIVE

The technology acceptance model (Venkatesh, 2000) is determined by the "perceived usefulness of the technology and the perceived ease of use of the technology" (p. 343). The perceived ease of use has several components that encompass how comfortable someone is with the technology and how effortful the use. The rational actor perspective (Bowman et al., 2012) refers to the complex interaction between technology and goals. According to these theories and a study by Elliott and Polyakova (2014), frequent use of an application or similar types of applications creates a learning effect. As use increases, the learning effect allows motivation to surpass design frustration and ease of use as the user overcomes poor designs in order to reach their goal.

In both TAM and RAP, there are extrinsic and intrinsic influences on the choice of technology and its use—extrinsic in the sense that there are outside pressures to use technology (e.g., Your friends are all on Instagram, so why aren't you?) and intrinsic pressures to use technology (e.g., I am very curious about what they are saying and posting on Instagram). These two types of motivation intertwine and are not easy to separate when using technology. In addition, there are also two other influences on motivation.

EGO INVOLVEMENT AND TASK INVOLVEMENT

The two other influences are task involvement and ego involvement (Ryan & Deci, 1989). When a person is motivated through task involvement, they are involved in the task for the sake of mastering it. In other words, you might sign up for Instagram because you like learning new applications and how they work. Or you might install a new operating system on your computer because you like to expand your knowledge of how operating systems work in general.

When a person is motivated through ego involvement, it is to prove a point. If you signed up for Instagram and were motivated by ego involvement, your motivation might be to post the perfect selfie, the pictures from last night's party, and a picture of your new car in order to show off to your friends. Some researchers claim that task involvement is intrinsic motivation and ego involvement is extrinsic motivation (Ryan & Deci, 1989). When these goals are in line with these other needs, the person experiences greater satisfaction, less stress, greater resiliency, and less fatigue (Szalma, 2014). The most useful portion of the discussion is the design recommendations that have evolved:

MOTIVATION-BASED SPECIFICATIONS

Szalma (2014) lists eight specifications of motivated design, or as it is also called, **eudemonic design**:

1. *Functional design*. The system demands should meet the user at their skill level.
2. *Eudemonic design autonomy*. The system provides the user with a choice in how they choose to perform the task.
3. *Competence*. The system allows the user to become proficient.
4. *Relatedness*. The systems avoid alienating the user or isolating them from others.
5. *Self-concordant goals*. The system is supportive of the user's values and goals.
6. *Need satisfaction*. As the user uses the system, their needs are met.
7. *Organizational context*. The system supports both short- and long-term goals and integrates with the sociotechnological system that the user already has in place.

DESIGN MEASUREMENT GUIDELINES

In order to use Szalma's specifications, they must be measured. Therefore, the following measurement guidelines are also proposed:

1. The designer should find out the user's goals and why they have these goals. For example, in a driverless car, the user may have several goals depending on the task or the riders, such as the following:
    a. Driverless car needs to be able to shuttle children to soccer practice and home.
    b. Driverless car needs to be able to take the elderly to medical appointments.
    c. Driverless car needs to be able to take people to work.

2. The user's goals will also address the need for **autonomy** (control and choice), **competence** (ease of use, skill development, challenge), and relatedness (Do they need to be connected to a system or an "other" to decrease feelings of loneliness and abandonment and to increase attachment with others and support high-quality relationships with others?). For example, in a driverless car, the user may have the following needs for autonomy:

    a. Children require autonomy for the convenience of the parents, do not require competence, and need to be in constant contact with their parents.

    b. Elderly require a great deal of autonomy and may require additional stops at a whim; they also require a great deal of competence and may request a great deal of relatedness.

    c. Working adults require a great deal of autonomy and competence but not much relatedness.

3. In addition, the measurements should address the following questions: What are the cognitive demands of the system?

4. What are the perceptual demands of the system?

5. Does the system promote **out-of-the-loop unfamiliarity** (**OOTLUF**) or learning?

6. How should the system reward and prompt users to maximize engagement?

These specifications and guidelines are still evolving. They are a great starting point when incorporating motivation into your design specifications. In addition to addressing motivation, design also should address the human emotion of boredom. Boredom is a bit more complex than motivation. Boredom is a negative emotion that can be brought on by a system state, by an individual's feelings, or by an individual's lack of attention. Boredom has been a topic in human factors as long as automation has been a topic, as an excess of automation in a system is thought to induce boredom.

BOREDOM

Boredom happens when a person feels under- or overstimulated, feels trapped, and is unable to maintain attention. Understimulation can happen when the information in the environment is either already known or not relevant to the individual's goals. The person has little motivation to direct attention to the information. Overstimulation can happen when there is too much information, and the individual doesn't know where to pay attention first or where to start. Feelings of being trapped happen when the individual feels social pressure to stay in the situation until a designated ending. This could be a class, a church service, a meeting, or a performance. The combination of being unable to change the situation and the disconnect between the stimulation level and the individual's goals creates a loss of attention. In automation, boredom occurs when the system fails to engage the human at the appropriate level. In automation, boredom can create a serious problem, as the human is unavailable to rescue a system failure.

Researchers suggest that there are two types of boredom: trait and state. The difference is that when boredom is a trait, that particular person is prone to being bored. Usually, trait boredom is associated with an attention disorder or depression (Hunter & Eastwood, 2018). State boredom is when a person is experiencing a temporary state of boredom at the same rate as most people in a similar situation. It is unclear if trait boredom can be designated as a true trait of an individual or is simply a failure to engage attention normally.

Boredom is a concern in automation. When the system is highly automated, the human is less involved and experiences a loss of SA and out-of-the-loop unfamiliarity (OOTLUF). Attention is diverted to other tasks such as watching a movie as the automated car drives you to your location. The overreliance on the automation with the OOTLUF conspires to create boredom in the human operator. Boredom creates the motivation to divert attention to more engaging tasks. The operator is unavailable for system rescue. Often, system designers will assume that the human operator will be available in case of failure and neglect to design for boredom. When there is an abnormal end of the system and its operator is bored or unavailable, disaster may result. This is the primary concern regarding boredom, emotion, and motivation: the human may be completely out of the system loop.

REFERENCES

Bowman, N. D., Westerman, D. K., & Claus, C. J. (2012). How demanding is social media: Understanding social media diets as a function of perceived costs and benefits—a rational actor perspective. *Computers in Human Behavior, 28*(6), 2298–2305. http://dx.doi.org/10.1016/j.chb.2012.06.037

Elliott, L. J., & Polyakova, V. (2014). Beyond Facebook: The generalization of social networking site measures. *Computers in Human Behavior, 33*, 163–170.

Hull, C. L. (1943). *Principles of behavior*. Appleton-Century-Crofts.

Hunter, A., & Eastwood, J. D. (2018). Does state boredom cause failures of attention? Examining the relations between trait boredom, state boredom, and sustained attention. *Experimental Brain Research, 236*, 2483–2492. https://doi.org/10.1007/s00221-016-4749-7

Maslow, A. H. (1970). *Motivation and personality* (2nd ed.). Harper and Row.

Plutchik, R. (1980). A general psycho-evolutionary theory of emotion. In *Theories of Emotion* (pp. 3–33). Academic Press.

Ryan, R. M., & Deci, E. L. (1989). Bridging the research traditions of task/ego involvement and intrinsic/extrinsic motivation: Comment on Butler. *Journal of Educational Psychology, 81*, 265–268.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67.

Szalma, J. L. (2014). On the application of motivation theory to human factors/ergonomics: Motivational design principles for human–technology interaction. *Human Factors, 56*(8), 1453–1471.

Venkatesh, V. (2000). Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information Systems Research, 11*(4), 342–365.

CHAPTER 12

DECISION-MAKING AND EXPERTISE

GOALS OF THE CHAPTER

- Learn about how expertise develops
- Understand how decision-making is tied to expertise
- Learn about expert systems

ASSESSMENT

- How is criticism vital to developing expertise?
- What are the three ways that expert status is determined?
- What is a recommender system, and how is that different from an expert system?

Expertise is the term that we use to distinguish someone who has significant practice and knowledge in a domain. A domain is any area that is defined by specific terminology and specific knowledge of how the structures, functions, and behaviors exist among entities. A domain contains knowledge that a person would need to get through either significant experience or significant training. For example, the practice of medicine is a domain area. Automobile repair is another domain area.

We have certain expectations of experts in any domain area. We expect that they will make few if any mistakes. We expect that there are standard ways of doing things within a domain that are associated with good outcomes. Finally, we expect that experts in any domain have amassed the knowledge and experience to make good decisions.

Expertise is difficult to quantify. One of the top authors in this research area is Ericsson (2017). Ericsson summarizes the field by stating that expertise is gained through deliberate practice regardless of the field or domain and regardless of the immediacy of feedback. **Deliberate practice** happens when a person works alongside an expert such as a coach and this expert critically assesses their performance. The expert suggests areas of improvement and how to improve. Then the person works on improving that area, only to be assessed again by the coach, who suggests further improvement. In countless research studies, the effect of deliberate practice has been quantified.

For example, I am certain that Aunt Norma makes the best custard pies and that she is an expert. She was trained by Grandma Ora, who criticized each and every pie that Aunt Norma baked. No pie was good enough for Grandma Ora. Through this criticism, Aunt Norma learned the qualities that created a good pie. She learned the cues to look for when making an excellent pie: what the oven smelled like when it was ready, how brown the crust should be, what the dough should feel like when you roll it out. Over 70 years of practice, with 20 of those years as deliberate practice with Grandma Ora's coaching, Aunt Norma learned to bake an excellent pie. I may feel that Aunt Norma's pies are the best, but you also have an Aunt Norma in your family, and maybe her custard pies are better.

HOW IS EXPERTISE DETERMINED?

Expertise can be determined in several ways. First, it can be quantified through professional certifications, or the nomination by peers, which are the most common. Medical doctors take medical board exams, engineers take an engineering exam called the ABET, and lawyers take the bar exam; all of these are examples of professionals getting certified as experts. Second, expertise can be determined through nomination by peers. This happens when a group of previously determined experts assesses the performance of a potential expert and then judges if the person has met the criteria for expertise. Maybe the known pie-baking experts from *The Great British Bake Off* watch Aunt

Norma bake a custard pie and then taste it. They will examine how she performs the task, what she does, her decisions, and then the result. These experts then determine if she has met the criteria for expertise that they set. A third example is obtaining your doctorate degree or terminal degree: other persons in your field supervise your work and examine your decisions, the knowledge you obtained through the process, your reasoning, and your ability to communicate. Then a committee decides if your preparation warrants inclusion into the doctorate or terminal degree.

Expertise is most evident when people are solving problems within their domain. Most researchers would agree that much of problem-solving is domain specific. This is because the way that you would solve a problem in chess is not the same way that you would solve a problem in auto repair. While the knowledge is completely different, so are the strategies that one would use.

Still, decision-making requires significant expertise in most domains. Not only do the decision-makers balance the factors that are impacting the decision, but they are considering the impact of the consequences of a poor decision. Several theories address decision-making. The most prominent ones in psychology are types of rational decision-making (prospect theory, Bayesian decision-making) and naturalistic decision-making.

Some of the first studies in decision-making were aimed at discovering why people make poor economic decisions. Economists teamed up with psychologists to develop expected utility theory, which describes how people will weigh the benefits and the risk of any decision before they decide. These benefits and risks are understood differently by different people, and the probability of the benefit or the loss (risk) actually happening is poorly understood by most people. **Utility theory** has several refinements in **prospect theory** and Bayesian decision-making (Hands, 2012; Scott, 2000).

Bayesian decision-making takes into account the benefits and risks but also the likelihood and prior information from similar events (Pan, 2016) along with uncertainty about the outcome. This incorporates the element of preference and past history into a utility-theory model. The common criticism of utility-theory models is that they fail to account for **cognitive biases** that people are prone to making and the outcome of many decisions.

To account for decision-making when experts consistently make good decisions, Klein (2008) developed **naturalistic decision-making**. Naturalistic decision-making requires that the person have some expertise in the problem area and that this expertise allows the person to pattern match. For example, in firefighting, when the smoke is a certain color and shape, the firefighters can ascertain what is burning and how long the fire has been consuming the building. This knowledge is from studying millions of fires over the course of the history of firefighting. Some of this fire knowledge is universal, but other types of fire knowledge are specific to the types of fires and buildings in that area (Elliott, 2014). For example, a fire in a building in Erie, Pennsylvania, in the winter may involve a heater, an electrical system, and wood structures. A fire in a building in a New Mexico winter may also involve a heater and electrical system and an adobe clay structure that does not deteriorate in the same way during a fire as a wood structure. An Erie firefighter working in New Mexico may decide to put a ladder on the building and make vents in the roof. This would be an error, as the adobe and the clay roof tiles become very hot and would make the ladder very hot as well. If the firefighter attempted to climb onto the roof, their feet may be burned from the heat emitted by the clay tiles.

Naturalistic decision-making involves the same risks as rational decision-making in terms of uncertainty, familiarity, expertise, and time. However, the time may be compressed, and the decision-maker may be presenting information or deciding for others. This is often the case in medicine. Physicians must decide how to present information to patients to allow them to make the best decision for their care. They must also allow the patient to determine the care that they prefer based on many factors, such as the probability of recovery and their preferred quality of life.

In general, studies have shown that when faced with critical health-care decisions, people are notoriously poor at understanding risk and the probability of loss associated with health-care decisions (Cokely et al., 2012). In persons with low numeracy skills, understanding the probability of loss is even worse. In an effort to address this challenge, Garcia-Retamero and Cokely (2017) have proposed that physicians use a system of visual graphics to explain the risk of nontreatment or different treatment options.

EXPERT SYSTEMS

Expert systems are computer systems that mimic the decision-making of an expert. These include systems that you can ask for a decision, such as a symptom checker on WebMD or systems that make a recommendation based on your past choices. These systems are becoming increasingly popular, as artificial intelligence is more accurate and reliable.

Many of you may subscribe to a streaming service such as Netflix to watch films and shows. Within Netflix, there are thousands of movies that can be viewed, and all of them have different qualities. Maybe you have a younger sibling that loves to watch *The Regular Show*, and you love to watch *Black Mirror*. Both shows have qualities that are in common and qualities that are very different. You use these qualities to decide when to watch what show.

You might also ask the system to make the selection for you. In this case, you might pick from Netflix's "Popular Right Now" list, or you might pick a show from what they recommend for you from your past selections. This is a **recommender system**. It records the frequency of the different qualities of the shows or movies you watch and finds movies and shows with similar qualities. As you choose different shows and films, the **algorithm** weighs the different qualities according to your past viewing preferences.

When you sign up for a streaming service, you report your age and gender, where you live, and a few other demographic details that don't seem to be important at the time. If you share your account with friends, your recommended shows are an amalgamation of your preferences and the preferences of your friends. Netflix's service is based on the idea that a person will consistently choose the same qualities in their entertainment. For example, a person who watches horror films may enjoy the jump-scare sequences. The jump scares are qualities across types of movies. If a person consistently chooses movies with jump-scare qualities, that quality will be assumed as a preference and entered into the **recommender system**'s list of preferences for that movie watcher.

Another type of system that helps with decision-making is an expert system. An expert system weighs the probabilities of good and bad outcomes along with the outcomes of previous decisions. Expert systems are good at solving **convergent problems**. These are problems with a single best solution. For example, in a multiple-choice test, there is one right answer to each question. Expert systems are not good at solving **divergent problems**. These are problems with many different solutions that are equally good.

The challenge with expert systems is that the areas in which expert systems would be most useful are the areas where there are many divergent problems and few convergent problems. For example, in medicine, it would be handy to know if a certain set of symptoms indicate a specific disease, except that is not the case, as each person's body has slightly different symptoms with the same disease and the progression of the disease differs by person. For example, think of the last cold that you caught from a friend. Your friend may not have had a cough, but you did.

There are some exceptions to medicine having only divergent problems. Some problems in medicine are clearly convergent and benefit from an expert system. For example, there are expert systems that aim to reduce prescribing errors in patients. One such system would alert a pharmacist if the patient was allergic to a drug or was known to be pregnant. These expert systems focus on the convergent problems, where there is a single best solution.

For example, if a patient was prescribed a drug that they were allergic to, then another drug would be suggested. Then if the patient was pregnant, the second drug may not be suitable for pregnancy, but the first drug was suitable. In this case, an experienced pharmacist and physician would be needed to determine which options would help the patient. An inexperienced pharmacist would be tempted to ignore alerts or become frustrated with the warnings.

When an expert system cannot accurately portray the risks and benefits of a particular decision, it produces false alarms. The false alarms teach the pharmacists to ignore the expert system through conditioning. When this happens routinely, the pharmacist overlooks the warnings. The errors that the system intends to fix are more frequent. However, this is an area of rapid change as researchers and engineers incorporate machine learning to expand the types of problems that an expert system may solve.

REFERENCES

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making, 7*(1), 25–47.

Elliott, L. J. (2014). *The role of analogical reasoning in cue evaluation for first responders*. GRIN Verlag.

Ericsson, K. A. (2017). Expertise and individual differences: The search for the structure and acquisition of experts' superior performance. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*, 1–6.

Garcia-Retamero, R., & Cokely, E. T. (2017). Designing visual aids that promote risk literacy: A systematic review of health research and evidence-based design heuristics. *Human Factors, 59*(4), 582–627.

Hands, W. (2012). Normative rational choice theory: Past, present, and future. *Voprosy Economiki, 10*, 1–30.

Klein, G. (2008). Naturalistic decision making. *Human Factors, 50*(3), 456–460.

Pan, Y. (2016). *Bayesian decision theory guiding educational decision-making: Theories, models and application*. Society for Research on Educational Effectiveness, SREE Spring Conference Abstract 2016. https://files.eric.ed.gov/fulltext/ED566993.pdf

Scott, J. (2000). Rational choice theory. *Understanding Contemporary Society: Theories of the Present, 129*, 671–685.

CHAPTER 13

LANGUAGE AND ARTIFICIAL INTELLIGENCE

GOALS OF THE CHAPTER

- Understand how the study of language gave rise to artificial intelligence
- Learn about the different approaches to language study and their influence
- Understand the connection between memory, knowledge, and language

ASSESSMENT

- What is the difference between syntax and semantics in language?
- Why should the people who write error messages be concerned about conversational implicature?
- How are both incorporated in modern AI?
- What is latent semantic analysis, and how does it work?
- What is the McGurk effect, and why is it a problem for AI?
- What is Searle's Chinese room, and how is it a critique of AI?

In previous chapters, we discussed expertise, decision-making, and training, all of which depend on memory and knowledge. Knowledge is an abstract term for a very important concept—the things that you know and things that help you make decisions. Here are some of the things that help you make decisions: facts, mental models, schemas, opinions, beliefs, patterns, observations, and a variety of other perceptions or behaviors that belong to you and your experience.

Knowledge is difficult to measure. We have only a few ways to measure knowledge: **reaction time**, accuracy, the number of things known or remembered, and language. Language contains our thoughts, and it is the relationship between those thoughts that contains meaning. Thanks to the pioneering work of Forster and Davis (1984) and others, we understand how second languages work depending on when you learn them. Thanks to Chomsky (1959), we understand that language is more than a stimulus-response learned behavior. And finally, thanks to Landauer, Foltz, Kintsch, and colleagues, we now have **computational ways** to represent language that allow us to type in common words into a search engine and get an accurate response.

All this research work is based on several ideas: some word combinations are more frequent than other word combinations (language is domain/situation specific), combinations of words have meaning (**semantic relatedness/neighborhoods**), and combinations of words have prescribed orders in which they should be used (**grammar/syntax**). In the example below, it is much more likely that the first sentence would occur than the second sentence:

1. Joan drank green tea and fretted about the book.

2. Joan ate the book green tea fretted.

In American English, sentences are constructed with noun-verb patterns, as illustrated by the first sentence. The first sentence is an example of that pattern. *Joan*, *green tea*, and *book* are all nouns, while *drank* and *fretted* are verbs. The second sentence does not follow that pattern; it follows a noun-verb-noun-noun-verb pattern. The order of the words can be referred to as **syntax**. The second sentence violates the syntactic rules, and it would not be spoken under normal circumstances by a normally functioning person who was a native speaker of American English.

Semantics are the meanings of words alone and in combinations. For example, in the first sentence, you understand that someone named Joan was drinking tea and worrying about some sort of book. In the second sentence, you typically would stop at *book*, as a person doesn't normally eat a book. Then you would search for the reason, but because the sentence does not make semantic sense, you would fail to process it further.

Programming the computer to understand that the meaning of the word combination must make sense is more complex. Both *book* and *green tea* are nouns, yet we normally treat each of these items differently because of the distinct difference in semantics. Coming up with a way to codify treatment, context, and meaning is what Landauer, Foltz, and Kintsch accomplished. They used the idea of neighborhoods because in a real neighborhood, you have people who are connected physically (their yards are connected or they live in similar houses next door to one another). It would be fair to state that most neighborhoods share other commonalities: there are neighborhoods in which most people are students, neighborhoods where most people work in factories, neighborhoods where most of the residents are retired, and neighborhoods where many young families live. In language, this idea is the same. Language neighborhoods are groups of words that are related in meaning in some way. There are some close neighbors such as *green* and *tea* that occur frequently together, and some far neighbors such as *heuristic* and *tea* that occur infrequently together. Neighborhoods can change with context, subject matter, the expertise of who is speaking, and their culture. As you grow into your profession, your language will change as a result of your use of specific words and the frequency in which you use them.

If you change regions of the United States, the language will also change. You will use some words together more frequently, and you will learn that some combinations have negative connotations and some new combinations have positive connotations. For example, in Indiana, stating that someone is a true Hoosier means that they are true to their Indiana roots and maybe fans of the Indiana sports teams. Yet the same phrase one state away, in Missouri, means that someone is from a lower-income class, is reckless, is slovenly, and lacks pedigree. These types of cultural artifacts are so embedded in our everyday speech that we have trouble detecting them ourselves.

Moreover, these artifacts can be problematic for computers to decode and the source of much humor when using Siri or similar computerized natural language processors.

The other item that you should understand when it comes to human factors and language is the idea of communication rules. The philosopher H. P. Grice (1975) wrote about meaning and communication conventions. In Grice's book, he states that in order for a conversation to proceed, there are common rules that we have for each person in the conversation, which are the rules of **quantity**, **quality**, **relation**, and **manner** (p. 45).

For quantity (Grice, 1975), it is expected that your contribution to the conversation should be informative but not too informative. For example, think of the last question you asked a friend and Google. Google, or a robot conversationalist, would provide all information on the given topic, hoping that the information that you needed would actually be there. Your friend would understand the meaning of your question and tell you the exact information that you asked for without any extra information. For example, let's say that you asked, "What time is the Thanksgiving dinner going to be at Joe's house?" Your friend would state, "Noon." Google would violate this Gricean rule and state among the 102 million results that there are two restaurants named Joe's, and the optimum time for Thanksgiving dinner according to the *Atlantic* article is 4 p.m. Yes, I Googled it.

For quality and relation (Grice, 1975), it is expected that your statements are not false and that you have ample evidence to support your assertions. For example, if you know that the moon is to be full in the next day or so, state that the moon will be full tomorrow or the next day. Do not state that the moon will be full soon. For relation, in all your communications, be relevant to the topics at hand. For example, if the conversation is about politics but you prefer to talk about lunch options, then ask to switch the conversation instead of answering incorrectly.

For **manner** (Grice, 1975), in your communications, be clear, orderly, and brief. In other words, refrain from using flowery language or adding unnecessary adjectives and adverbs. Do not try to impress others with your command of obscure words, and get to the point immediately. For example, when describing the lack of direction given for an assignment, it is expected that the person ask for more direction instead of stating, "The paucity of information has caused a decrement in my ability to understand the material."

Why am I mentioning these implicatures or rules of conversation? Because in human communication, when one or more parties do not follow the rules, then the other person in the conversation tries to find out why the rules are not being followed. They may ask questions of the speaker to try to figure out why the conversation is going so poorly. When humans communicate with robots or with systems, the rules do not change. Humans become confused as to why the system gives an instruction that does not seem relevant or asserts a statement that is untrue or provides too much information on the inner workings of the computer that are not relevant to the current computer problem.

Error messages are notorious for violating conversational implicature. Windows in particular will flash the "blue screen of death" or give a "runtime error." While this may be helpful for the cottage industry of decoding error screens for Windows, it violates conversational norms, and this is what creates frustration for most users, as few understand what to do next. Immediately, many extra steps are required for them to finish what they were doing on the computer.

While conversational implicature, as in the **Gricean maxims** (1975), are a pain point, our understanding of language processing has come a long way in the past few years. Most of this work is due to important research in language that happened in the 1980s. First, language processing was based on frequency (Forster & Davis, 1984) and processed in parallel (McClelland & Rumelhart, 1989; McClelland et al., 1987). Specifically, Rumelhart's idea that the predictable nature of language was due to the weighing of different aspects of words through hidden networks and intermediary networks of probability influenced Kintsch (1988, 1994) and Foltz, Laham, and Landauer (1999). They refined this idea of meaning being constructed by the relationships between words and assigned a numerical value to these relationships between words. For example, Foltz, Laham, and Landauer (1999) created an automated essay grader called latent semantic analysis. This essay grader would "read" an essay and compare it to previous good essays. The essay grader was not reading the essay; it was assigning numerical values to the relationships between words and calculating a similarity score. If the similarity score between the new essay that it read and the previous good essay was a match, the essay got that grade. Sometimes, people think that these algorithms provide a word-for-word match. This is not so; it creates a numeric value based on the neighborhoods that the chosen words live in and the relationships between words in different neighborhoods along with probability, grammar, and some other attributes. This approach to a machine learning to "read" was revolutionary and created the mechanisms for the semantic web, Google, natural language processing, Siri, and much of what we think of as artificial intelligence today.

Latent semantic analysis (**LSA**) is a way to analyze large bodies of language and compare them through meaning. In some of the original iterations of LSA, the system had to be trained on nearly 300 examples of the same topic in order to understand the relationships between the words as well as a human. For example, if you were to grade essays in an Introduction to Psychology course, and there were two essay questions on the final exam, you would need to find 300 examples of similar essays for the first question and 300 examples for the second question. While LSA is the basis for processing written words, it also has been a breakthrough for auditory natural language processing. While the current challenges in auditory language processing differ, the semantic processing is similar.

In order to see LSA in action, open up Google and type in a sentence with a clear and definite meaning, such as "What is the best weighted blanket to buy?" The LSA portion of Google will have no problem understanding "best" and pairing it with a "weighted blanket," which is a special type of blanket that is currently popular, and no problem understanding the word "buy." Next, try a sentence with a less obvious meaning using low-frequency words and putting them together in a less predictable order, such as "Where can I find witch hazel to bless my staircase railings?" You will receive a much different pattern of results. Google will focus on the high-frequency words of "staircase railings" and then "witch hazel" and completely ignore the verb "bless," as these words are in completely different neighborhoods and rarely appear together.

All this is fun until you begin to realize how hard people are working to make a world in which we can give voice commands to computers. Speech-to-text engines, chatbots, and telephone bots are all improving each day, yet few can hold a meaningful conversation with a human. But we know that the day will arrive when computers can understand and produce conversations as well as a human. The inhibitory issues of dialect, accents, cadence, and the **McGurk effect** interfere with the progress in communicating verbally with a computer system. The McGurk effect is when the motion of the lips is needed in order to translate the sound that the person is saying. The smallest meaningful part of a word is called a phoneme. The **phonemes** of *ma*, *ba*, and *va* are examples of the McGurk effect in action. For these particular phonemes, seeing the speaker's lips move helps us decode the word. There are many videos on the McGurk effect online. The McGurk effect, along with the other issues, has been quite a challenge for speech to text. Like LSA, speech understanding requires hundreds of examples in order to set the probabilities of what the likely next word or the next meaning will be in a conversation.

Historically, people believe that the true test of language understanding and therefore artificial intelligence is the **Turing test**. According to Goertzel (2015), who describes Turing's idea from 1955, the Turing test is passed if a machine can converse with human judges so that the humans believe that they are conversing with another human, not a machine. This has been the ultimate standard for many decades, and occasionally, a computer will claim to pass, and several judges will declare a winner. The bigger questions for most psychologists when it comes to issues of language and intelligence are, "Does the computer understand?" and "Can it attribute meaning?" Philosophers such as Searle (2009) claim that this question is more important than the Turing test. Searle claims that his thought experiment of the **Chinese room** describes our current state of machine intelligence better than the Turing test. In Searle's thought experiment, he suggests that a monolingual English speaker such as himself is in a room of Chinese symbols. He is given the instructions of what to do with the symbols in English when the symbols come through a door slot. When a person poses a question in Chinese to the room by sliding the question through the door slot, he follows the English instructions and delivers an answer in Chinese. Although he speaks no Chinese and doesn't understand any Chinese, over time, he becomes proficient at delivering the correct series of characters in the right order to any question. But he still fails to understand the meaning behind the characters and their combination because it is all in Chinese. Searle argues that this is essentially the Turing test (http://www.scholarpedia.org/article/Chinese_room_argument).

While the Turing test and Searle's Chinese room are striking examples of how language exists in computers but generative intelligence does not, many researchers and technologists would argue that generative intelligence is not needed in computers for them to replicate behavior and human decision-making. Computers make an important contribution to the world, and as they become more "human," ethical decisions on their rights, their consciousnesses, and their role with humans are to come. For example, who is liable when Twitter bots spread disinformation: Twitter, the bots, the humans that created the bots, the people who believed the false information spread by the bots and then shared it with friends, or the origin of the false information? Is the disinformation only criminal when harm is done? What type of harm qualifies as harm to a person, to an organization, to a profit-making enterprise, or to a country? These and many other questions are in our future.

REFERENCES

Chomsky, N. (1959). A review of BF Skinner's verbal behavior. *Language, 35*(1), 26–58. JSTOR.

Foltz, P. W., Laham, D., & Landauer, T. K. (1999). Automated essay scoring: Applications to educational technology. In *EdMedia+ Innovate Learning* (pp. 939–944). Association for the Advancement of Computing in Education (AACE).

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(4), 680.

Goertzel, B. (2015). Artificial general intelligence. *Scholarpedia, 10*(11), 31847. https://doi.org/10.4249/scholarpedia.31847

Grice, H. P., Cole, P., & Morgan, J. (1975). *Logic and conversation*. Harvard University Press.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*(4), 294.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*(2–3), 259–284.

McClelland, J. L., & Rumelhart, D. E. (1989). *Explorations in parallel distributed processing: A handbook of models, programs, and exercises*. MIT Press.

McClelland, J. L., Rumelhart, D. E., & Group, P. R. (1987). *Parallel distributed processing* (Vol. 2). MIT Press.

Searle, J. (2009). Chinese room argument. *Scholarpedia, 4*(8), 3100. https://doi.org/10.4249/scholarpedia.3100

Turing, A. M. (1955). 1954. *Biographical Memoirs of Fellows of the Royal Society, 1*, 253–263.

THE DESIGN OF PHYSICAL WORK

CHAPTER 14

ANTHROPOMETRY

GOALS OF THE CHAPTER

- Learn about how humans are measured
- Understand the importance of measurement variation and how it is addressed
- Understand the guidelines responsible for keeping workplaces safe

ASSESSMENT

- Why are humans measured, and who measures them?
- What organization ensures that workplaces follow best practices?
- What is the difference between an anthropometricist and an ergonomist?

THE SIZE OF CLASSROOM DESKS

I've noticed in my classrooms that some students seem comfortable in their desks and others do not. For those who do not, the desks and chairs seem either too big or too small. The students who seem to be most comfortable are of an average size and weight. For the tall students, the desks are too short, and for the short students, the desks are too tall. Have you ever wondered how desk size is determined?

Furniture manufacturers use **anthropometry** to determine the size of the desk. Anthropometry is the study of how people fit into physical spaces and the variation among humans' physical size. While we can have desks and chairs that work for everyone, it would be costly, as the university would have to guess how many students are of average size, how many students are tall, and how many students are short. It is most efficient to have a desk or chair that works for most people.

You will notice that for your taller classmates, the distance from their elbow to their fingertips is longer, and the distance from their knees to their ankles is longer. In smaller classmates, the opposite is true. When furniture designers create a new piece of furniture, they must decide on the height of the desk, the distance from the floor, and all the other measurements. To do this, they consult an anthropometric guide that gives the average human measurements for a range of years. Here is one guide: https://www.cdc.gov/nchs/data/nhsr/nhsr010.pdf.

Guides such as these contain lists of measurements for women and men as seen in Table 14.1. There are guides for measuring the anthropometric qualities of specific tradespersons such as bricklayers or athletes. There are also guides for the anthropometric measurements of the general public, which includes students.

In addition to the lists of measurements, there are guidelines that show how to take these measurements. One of these is called the International Standards for Anthropometric Assessment and is one of several books

Table 14.1. Measurements of Padel Players

	Percentiles						
Variable	**5**	**10**	**25**	**50**	**75**	**90**	**95**
Height (cm)	164	167	174	178	183	187	195
Weight (kg)	58	65	71	77	81	94	97
Body mass index (BMI)	19	22	23	24	25	28	30
Arm span (cm)	165	176	179	182	185	193	201

Source: Derived from the measurements of padel players from Sánchez-Muñoz et al. (2020, p. 7).

that outline standardized practices for taking these measurements (Stewart et al., 2011). The guidelines on how to take the measurements are just as important as the measurements themselves, as different practitioners will arrive at slightly different results. This illustrates the importance of taking many different measurements of slightly different populations over time by a variety of licensed anthropomorphists.

These measurements are used by several different types of professionals: furniture designers, clothing designers, architects, machinery designers, and ergonomists. In the last two categories, machinery designers and ergonomists, these professionals ensure that **Occupational Safety and Health Administration** (**OSHA**; https://www.osha.gov/) guidelines are met to reduce stress and physical injury in workers. Ill-fitting machinery, equipment, and furniture contribute to ill health and accidents in the workplace.

Work-related accidents contribute to two million deaths per year (**World Health Organization, 2010**). More international statistics can be found here: https://www.ilo.org/global/lang--en/index.htm. Many of us are familiar with white-collar jobs and cannot imagine that someone sitting at a desk would incur a work-related accident. However, there are millions of jobs that require the operation of large machinery or exposure to dangerous conditions. Pennsylvania is considered an agricultural state. Some of you may do farm work in the summer. According to the U.S. Department of Agriculture (USDA), in a year, over 300 youth die in farm-related accidents, and there are 23,500 nonfatal farm-related injuries; tractor-related injuries are the most frequent (https://nasdonline.org/1241/d001045/a-review-of-farm-accident-data-sources-and.html). This is one area where ergonomists use anthropometric data to increase the safety of tractors and other farming machinery.

Accidents in industrial settings often occur because of work posture and work rate. Work posture can be affected by the types of items a person is asked to lift or carry in an industrial setting. If a heavy item is not lifted properly or an item is dropped, a back injury can occur. Limits on the types of items and the method by which they are lifted or carried can be specified by a qualified ergonomist. Work rate is the pace at which employees are required to work. Again, a qualified ergonomist helps unions and organizations determine the number and length of breaks along with guidelines set by OSHA.

Workers should not incur injuries due to conditions, physical demands, or time demands. Anthropometric guidelines, ergonomists, and workplace regulation at the local, national, and international levels help ensure that this is minimized. As an employee, familiarizing yourself with these resources can help you ensure that these guidelines are followed and your workplace is safe.

REFERENCES

Sánchez-Muñoz, C., Muros, J. J., Cañas, J., Courel-Ibáñez, J., Sánchez-Alcaraz, B. J., & Zabala, M. (2020). Anthropometric and physical fitness profiles of world-class male padel players. *International Journal of Environmental Research and Public Health, 17*(2), 508.

Scientific American. (1998, June 29). Why are we getting taller as a species? Retrieved February 26, 2020, from https://www.scientificamerican.com/article/why-are-we-getting-taller/

Stewart, A., Marfell-Jones, M., Olds, T., & De Ridder, H. (2011). *International standards for anthropometric assessment* (*ISAK*). Lower Hutt.

Westat. (1998). *National health and nutrition examination survey II: Body measurements (anthropometry)*. Retrieved February 26, 2020, from https://wwwn.cdc.gov/nchs/Data/Nhanes3/Manuals/anthro.pdf

World Health Organization. (2010). *Healthy workplaces: A model for action: For employers, workers, policy-makers and practitioners*. Retrieved March 9, 2020, from https://www.who.int/occupational_health/publications/healthy_workplaces_model_action.pdf

CHAPTER 15

AUTOMATION

GOALS OF THE CHAPTER

- Learn about the levels of automation
- Understand how a human becomes less acquainted with a system as automation increases
- Understand how to address challenges when incorporating automation

ASSESSMENT

- What is OOTLUF?
- What happens as the level of automation increases to OOTLUF?
- Why is situation awareness important?

WHAT IS AUTOMATION?

Automation is when a single computerized or mechanical device or series of devices (machines) do something that may be normally done by a human. The use of machines has increased and will continue to increase as we augment our abilities with machines. Since the widespread use of machines beginning in the 1950s, the ability of machines to do things for humans has increased exponentially. Gordon Moore created **Moore's law** to describe this phenomenon in transistors (https://en.wikipedia.org/wiki/Moore%27s_law). Moore's law states that technology will increase exponentially. While simple machines such as toasters were common in the 1950s, now we have whole kitchens of machines that prepare food for us.

Machines make the whole experience of preparing food faster, more efficient, and sometimes safer. Moreover, they reduce error and the amount of labor required, and in cases of large commercial kitchens, the automated kitchen machines make their product consistent, quickly, and at a competitive cost. Humans like machines when they reduce the amount of effort a human must expend to accomplish a task. I would much rather type this manuscript on my computer than on a typewriter or write it out by hand. I would much rather take an Uber to school every morning than drive myself or take the bus (or even walk). Those three things incorporate different **levels of automation**.

Levels of Automation

Sheridan and Verplank (1978) along with others formed the idea that we should categorize automation by how much of the work is done by the machine and how much of the work is done by the human. This type of categorization is called a **taxonomy**. A machine that offers all the choices and the human decides—for example, a typewriter—would be a Level 2 automation on the Sheridan and Verplank scale. A system where the human makes the request and the computer executes the action upon approval would be a Level 5 automation, such as calling an Uber car service. See Figure 15.1 for an example of an automation taxonomy based on Sheridan and Verplank.

As the machine incorporates more and more automation, the human has less and less control overall. As the level increases, the machine can act without human knowledge and without human consent. In some cases, this is beneficial to both, such as an automatic braking system in a car or an automatic insulin pump for a diabetic patient. There are many times when human interference is detrimental.

In these cases, the machine makes decisions for the human. The process becomes complicated when the machine decision does not align with the potential human decision. Such is the current discussion in the justice system. In trials where the judge used a computer to

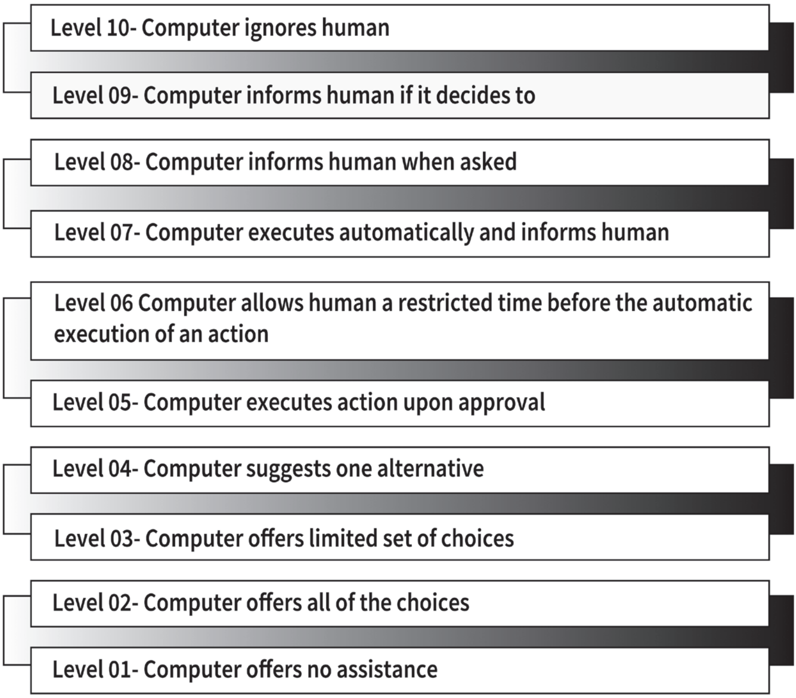

Figure 15.1. A Taxonomy of 10 Levels of Automation
Source: Based on Sheridan and Verplank (1978).

help them determine a sentence, the defendant may feel that they were unfairly sentenced because of the computer. The computer program has an algorithm that rates a defendant on various criteria. Then according to that rating, the computer predicts the probability that an individual would commit another crime. This probability influences the judge's decision on sentencing. A defendant with a high probability of committing another crime will likely receive a stricter sentence than a defendant with a low probability. An explanation is here: https://www.brookings.edu/blog/techtank/2019/03/21/algorithms-and-sentencing-what-does-due-process-require/. The algorithm that controls the ratings can be adjusted based on the criteria, the cases that it learned from, and the bias that is introduced into the program.

Out-of-the-Loop Unfamiliarity (OOTLUF)

While the justice system sentencing programs offer a striking example, in general, as the automation level increases, several things happen on the human side with trust, SA, and skills. As the automation level increases, the likelihood that humans will experience **out-of-the-loop unfamiliarity** (OOTLUF) increases as well. Let's take the example of automated cars. As humans drive less, they will become unfamiliar with driving, and their driving skills deteriorate. If they need to take the wheel, the OOTLUF will interfere, and they will be less proficient.

In addition to **OOTLUF**, **SA** and **trust** decrease with automation increase. For example, I am perfectly aware of what my toaster is doing, how it is doing it, and when the toast will be ready to eat. This is perfect SA: I know the past, present, and future situations and the contingencies and dependencies. The contingencies are the following: I know that if I smell bread burning, it is likely that small crumbs of bread became wedged into the heating coils, and I must turn off the toaster. If I smell bread burning, I must act to prevent a fire. And since I know that the system (the toaster) is dependent on electricity, its dependency is that the toaster must be plugged into the wall outlet to work. The system's performance is contingent on not having bread stuck in the heating coils. Remember that the toaster is on the lower levels of the taxonomy—about Level 2. Most of the time, I am 100% certain that when the bread is put into the toaster, the toaster will accomplish the goal and deliver a perfect piece of toast. I trust the toaster and have experienced no OOTLUF as a result of the toaster's low level of automation.

SITUATION AWARENESS AND TRUST

As the levels increase, my SA and trust change. When I call an Uber to pick me up from work, I have a vague idea of how Uber contacts the drivers and decides who will pick me up. But my SA is restricted to what the interface reveals to me. On the smartphone app, I can see the little cars driving around the map but have no control over which car comes to get me and who the driver may be. This lack of control along with the lack of SA contributes to my level of trust. If my previous experiences with Uber have been positive, I may trust the system to provide the best driver for me at that moment. If previous experiences had not been positive, I may feel frustrated over my lack of control and feel that the automated Uber algorithm will choose a driver who is not well suited to deliver me to my destination. Overall, in a perfectly operating system that never fails, humans trust the automation to perform correctly about 67% of the time.

Part of my loss in SA and loss of trust in the system has to do with some of the challenges inherent in automating a process. Whenever you decrease human error by automating a process, you increase the probability of system error. For example, if the system fails only 10% of the time, it is 90% reliable. If the system has multiple components that could fail, then the overall reliability is multiplied by each component's reliability. In a system with two components, each at 90% reliability, $0.9 \times 0.9 = 0.81$, or 81% reliability. This is discussed in Wickens, Hollands, Banbury, and Parasuraman (2015).

MEASURING SITUATION AWARENESS AND TRUST

In the situation awareness chapter, SA measures are discussed at length. All of these measures are sufficient in determining how well the operator can predict what the system will do next and what system state it is currently in. For trust, there are several new measures, including the Human Computer Trust Scale (HCTS; Madsen & Gregor, 2000) and the Trust in Automated Systems Scale (TAS; Jian et al., 2000). Often, human factors engineers will use measures of SA and measures of trust to ensure that the automation is well calibrated to human abilities and expectations.

OVERCONFIDENCE

Many systems rely on each other to deliver a product or service. For example, a very modern automated thermostat relies on the heating and cooling system to deliver heat to a large commercial building. If the heating and cooling system is controlled by an old computer language and the new automated thermostat is written in a new computer language, then something that knows both languages must provide a connection between the systems. People may trust one system more than the other. The person who runs the heating system may trust it even though it is a boiler system because it is predictable and has been reliable over the years. The person is overconfident in the system. When the new automated thermostat returns a system malfunction error and the old boiler doesn't heat the building, the person is more likely to look for an error in the new thermostat than they are to look for the malfunction in the old boiler. This is automation overconfidence. People will rely on perceptions instead of a thorough understanding of the system when assessing the automation for errors or missteps.

COMPLACENCY

In the same system, let's say that the heating and cooling system has been working without error for several years. The person managing the system may neglect maintenance because it runs so well. Or the person may ignore a minor error code that signals a possible future malfunction because of the positive history of the system. This happens with car owners and the tire pressure warning light. Because the tires may look fine and feel fine as the car drives, the car owner is more likely to attribute the tire pressure warning light to a sensor malfunction rather than a true problem with a tire. Then when the tire fails, the car owner may be surprised at the outcome despite having an automated warning.

Complacency issues are prevalent in aviation. There have been several incidents where airplane pilots turn off the autopilot because they don't believe that the instrument panel is correct. When this happens, the pilot will take manual control of the plane only to discover that the instrument panel has not failed and they should have left the autopilot in control of the plane. At the time of publication of this chapter, the Boeing 737 MAX airplanes are still grounded because of automation issues. A system named MCAS (Maneuvering Characteristics Augmentation System) overrode the pilot's manual control automatically and caused several planes to crash. In order to stop the MCAS system from overriding the pilot's control at the wrong time, the pilot must know how MCAS works and how to turn MCAS off and allow it to reboot. The MCAS sensor system could qualify as a Level 8 or 9 automation on the taxonomy.

This phenomenon is not new. In 1997, a master's student devoted a thesis to the issues of automation in planes. While this document is not widely read, it is very informative to those without a background in aviation. It can be found here: https://apps.dtic.mil/dtic/tr/fulltext/

u2/a327119.pdf. While it is easy to disparage automation because of these events, an increase in automation does save lives and does prevent human error. For example, pilots can die or become incapacitated midflight; automated pilots reduce the workload until another person is able to take control of the plane. Because of the automation, unskilled citizens may potentially land a distressed plane with the right conditions, supervision, and instruction. This would not have been possible without automation.

The current view on automation supports increased automation along with better cooperation with humans. It also recognizes the view that humans fail to value automation appropriately. Research is actively searching for ways to increase human SA and trust with automation and reduce OOTLUF. Some of the most recent research has suggested that automated systems should behave more like humans to increase acceptance. This means making occasional errors as humans do, providing varied and consistent feedback, asking for permission, apologizing for errors, asking questions, and occasionally failing on purpose to help maintain human operator skills and SA. Do you think that these recommendations are good or bad?

ADAPTIVE AUTOMATION

When we think of automated systems, we often think of static automation. This is when the level of the automation is constant. For example, my toaster will always toast the bread at the designated setting. The toaster will not sense the environment and change its settings on its own. **Adaptive automation** does exactly that; it senses the environment and changes or adapts to it. The environment could be the physical environment or the human. I like to think of an adaptive automobile braking system in cars as an example. First, the braking system waits for the human operator to react. If the sensor determines that there is no longer enough time for the human operator to press the brake to avoid a crash, the automated braking system will engage and prevent the accident. Future adaptive automation in vehicles could sense when the driver is not paying attention through an eye-tracking system near the windshield. In either case, the adaptive automation senses something in the environment, and this triggers a change in the automation.

AUTONOMY

Autonomy is the highest level of automation on any taxonomy of automation categories. Autonomy occurs when the machine operates independently of any human control. Typically, when we think of robots, this is what we assume. The robot acts on predetermined or learned objectives and performs the tasks independently. There are many concerns regarding this technology. These include system feedback and system state, the goals and directives of the system, and etiquette. Sometimes it's hard to think of a robot having etiquette, but if you think of the way that people interact, then it makes more sense. People know how and when to interrupt each other without being disruptive. There are stated rules between people to maintain civility. There are additional rules that govern communication between humans. These communication rules are the **Gricean maxims**: https://psychology.wikia.org/wiki/Gricean_maxims. The Gricean maxims are discussed in the chapter on language.

Ultimately, as with automation, autonomous machines' greatest hurdle is human acceptance and trust. In order to increase trust, there should be excellent system feedback, humans understanding what the system is doing at any given time, and automation etiquette. This helps with human SA and allows the human to step in at any point when the system falters. Treating the automation and the human as equal team members ensures that both give their full attention to the goal and help each other. The study of how humans and robots interact is called human-robot interaction (HRI). This is separate from the study of how humans and computers interact, or human-computer interaction (HCI).

REFERENCES

Jian, J. Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics, 4*(1), 53–71.

Jian, J. Y., Bisantz, A. M., Drury, C. G., & Llinas, J. (1998). *Foundations for an empirically determined scale of trust in automated systems (No. CMIF198)*. Center for Multisource Information Fusion.

Madsen, M., & Gregor, S. (2000, December). Measuring human-computer trust. *11th Australasian Conference on Information Systems, 53*, 6–8.

Sheridan, T. B., & Verplank, W. L. (1978). *Human and computer control of undersea teleoperators*. Massachusetts Institute of Technology, Cambridge Man-Machine Systems Lab.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology and human performance*. Prentice Hall.

CHAPTER 16

HUMAN-ROBOT INTERACTION

GOALS OF THE CHAPTER

- Learn about the different uses of robots in human-robot teams
- Understand some of the challenges in human-robot teams
- Understand anthropomorphism as it relates to trust in human-robot teams

ASSESSMENT

- What are the three Ds?
- How might a human factors professional measure trust in human-robot teams?
- How are telerobots and teleoperators different?

HUMAN-ROBOT INTERACTION

Robots are incorporated successfully into many areas of life, such as space exploration, aviation, undersea research, surgery, rehabilitation, agriculture, education, package delivery, mail processing, policing, and military operations (Sheridan, 2016). Robots are the perfect solution for tasks that are within the **three Ds**: **dirty**, **dull**, or **dangerous**. Robots can withstand conditions that humans cannot, such as space and undersea exploration. Robots can maintain attention for extended periods of time without lapse, which is useful in aviation, surgery, agriculture, logistics, and military operations. Robots can be constructed to possess enhanced strength, which allows them to carry heavy military packs, lift large elderly patients, or move pallets of packaged goods. Robots can be constructed to have enhanced vision for surgery or enhanced precision for aviation based on advanced sensor technology. Robots deal with dangerous situations better than emotional humans. In addition, they are not seen as culpable or capable of intentional harm. Many people believe that because of these attributes, they are therefore ideal for policing, aviation, and military applications.

As in automation, robots have varying levels of automation, with some robots being simply assistive, such as helping a nurse move a patient from a bed to a wheelchair. Other robots are fully automated and produce humanlike capabilities on an automotive manufacturing assembly line. These two different types of robots engender different levels of trust, reliability, social interaction, error, and acceptance based on their automation characteristics and their level on the taxonomy (Sheridan, 2016).

You might remember from the automation chapter that Sheridan created a taxonomy of automation; Sheridan has also created a taxonomy of robots (2016). In the first level, **telerobots** are robots that have sensors in that they are aware of their environments with limited mobility and a limited number of actions that they can carry out. Generally, these robots do things such as sort packages, provide industrial manufacturing support, or control substances in an office, lab, or pharmacy. In the second level, **teleoperators** remotely control objects in different types of spaces, such as oceans, interstellar space, airspace, and inside human bodies. Telerobots that are reprogrammable are also considered second-level robots. The third level consists of automated cars and other vehicles, automated highways, automated farm equipment, and automated aircraft. The fourth level is what most people think of when they think of robotics: social robots for companionship, entertainment, teaching, and assistance for children, the elderly, and others.

According to Sheridan (2016), the primary physical problems when working with robots are the ability to see the entire environment from the robots' point of view and the problem of communication delays or lack of communication.

TRUST, SOCIAL BONDS, AND ANTHROPOMORPHISM

Communication delays and the perception of the entire environment are important factors when establishing trust. When a robot has inconsistent communication or the inability to perceive obstacles, human trust in the robot erodes. This is particularly true when a person's physical health relies on the robot. In patients with limited mobility, robots are used to move a patient from the bed to a gurney or wheelchair. This solution assists the nurses and orderlies in this ergonomically challenging task, but often the patients do not feel comfortable being raised and lowered by a robot.

There are several measures of trust in human-robot interaction (HRI), according to Volante, Sosna, Kessler, Sanders, and Hancock (2019), such as the following:

- Interpersonal Trust Questionnaire (ITQ; Robinson et al., 1991)
- Negative Attitudes Toward Robots Scale (NARS; Nomura et al., 2006)
- Propensity to Trust Machines Scale (Merritt et al., 2013)

Some researchers have observed bias and conformity issues in human-robot teams. This suggests that there is a **social component** in trust and a social component when humans interact with robots. Volante and colleagues (2019) found that when social messages about the robot team member were presented to the humans through an interface not connected to the robot rather than through the robot, humans would rely on the norms of the other humans in the team regarding trust. When the robot could communicate with the humans directly, trust increased. However, when both were present, the trust decreased slightly.

Other researchers had observed the effects of oxytocin on trust in human-human interactions and explored the effect in human-robot interactions. In the study by de Visser, Monfort, Goodyear, Lu, O'Hara, Lee, and colleagues (2017), they asked participants to take an oxytocin spray orally and then participate in several experiments exploring trust, compliance, and decision-making. Trust and compliance with the team's norms increased in participants who had taken the oral oxytocin. This increase was found in human-robot teams as well as the previously studied human-human teams.

Trust and compliance also seem to interact with the level of **anthropomorphism** in the robot (de Visser et al., 2017). In this sense, anthropomorphism is the level of human features or the likeness that humans perceive in the robot. The question of how humanlike the robot should appear has a measurable balance. Overly humanlike robots have an increased creepiness rating. A nonhuman-like robot decreases trust. However, this balance can be manipulated. When humans took oxytocin orally, they were more receptive to social cues, compliance, and team decision-making, especially when the robots had a middle level of anthropomorphism. Both findings with oral oxytocin suggest that trust, compliance, and team interaction have a neurological factor not driven by human volition.

HUMAN-ROBOT TEAMS

Another study looked at conformity in human-robot teams. Hertz and Wiese wanted to know if humans would conform to robot behavior and if the conformity was task dependent or team dependent. If they were in a task-dependent situation, humans incorporated a competence judgment in their consideration and assigned more analytical or math-related tasks to the robot than to the human. In the teaming situation, humans were more likely to incorporate social strategies with other humans than with robots (Hertz & Wiese, 2018).

There is a great deal of work that still needs to be done to clarify the role of human-robot teams. Many of the aspects of human-human teams, such as communication, collaboration, trust, compliance, conformity, and decision-making, are also present in the human-robot teams. Human-robot teams have different levels of these factors based on the anthropomorphism of the robot and the **mental model** of the humans regarding the robot's capabilities and normative social rules imposed by human society (i.e., robots are good at math, humans are good at socializing). Defining the characteristics of human-robot teams will define how humanlike different classes of robots need and the attributes required of each class. For example, do telerobots who work in hospitals need to look like humans for patients to trust them? Do they need to laugh and cry with the patients in order to enhance trust? Do robots within industrial settings need to be the least humanlike for their assembly-line team members to trust their competency? If they tell off-color jokes in the assembly line, are their human teams more likely to feel a sense of kinship when they work alongside them?

LEARNING AND ROBOTS

Machine learning is an artificial intelligence technique in which computers/robots apply pattern matching and inference by statistical modeling to predict an outcome. When the outcome is validated, it becomes part of the learned pattern. This advance has allowed robots to display rudimentary learning. It is anticipated that with experience in pattern recognition, advanced sensors, and better language translation and production, robots will be able to match humans in many tasks in the future. For humans to take full advantage of these advancements in technology, addressing issues of human-robot teaming is needed, as is more research and study.

REFERENCES

de Visser, E. J., Monfort, S. S., Goodyear, K., Lu, L., O'Hara, M., Lee, M. R., Parasuraman, R., & Krueger, F. (2017). A little anthropomorphism goes a long way: Effects of oxytocin on trust, compliance, and team performance with automated agents. *Human Factors, 59*(1), 116–133.

Hertz, N., & Wiese, E. (2018). Under pressure: Examining social conformity with computer and robot groups. *Human Factors, 60*(8), 1207–1218.

Merritt, S. M., Heimbaugh, H., LaChapell, J., & Lee, D. (2013). I trust it, but I don't know why: Effects of implicit attitudes toward automation on trust in an automated system. *Human Factors, 55*(3), 520–534.

Nomura, T., Kanda, T., & Suzuki, T. (2006). Experimental investigation into influence of negative attitudes toward robots on human–robot interaction. *AI & Society, 20*(2), 138–150.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). Criteria for scale selection and evaluation. *Measures of Personality and Social Psychological Attitudes, 1*(3), 1–16.

Sheridan, T. B. (2016). Human–robot interaction: Status and challenges. *Human Factors, 58*(4), 525–532.

Volante, W. G., Sosna, J., Kessler, T., Sanders, T., & Hancock, P. A. (2019). Social conformity effects on trust in simulation-based human-robot interaction. *Human Factors, 61*(5), 805–815.

CHAPTER 17

VIRTUAL ENVIRONMENTS

GOALS OF THE CHAPTER

- Learn about what creates an immersive experience
- Understand how to minimize virtual reality sickness
- Learn about the different uses of virtual reality

ASSESSMENT

- What three items make up the triad of a virtual environment?
- What is virtual reality sickness?
- What are two things that you can do if a person has virtual reality sickness?

Virtual reality (**VR**) is a way to transport a person to a different reality that they physically do not exist in (Rebelo et al., 2012). The first VR machine is attributed to Heilig in 1962: https://en.wikipedia.org/wiki/Sensorama. This device had a person viewing a scene with three-dimensional glasses, which would allow them to view a scene as they "rode" through New York City. A fan created the effect of wind, and the scents of pizza, the images of buses, and various noises were pushed toward the viewer in an effort to create a multisensory environment.

As video games became popular in the 1980s and 1990s, developers incorporated surround sound and motion-producing controllers to create a sense of presence within the video game. However, to give the user a full experience of being inside the environment, the game needed to sense where the player/user of the system was in the three-dimensional space. This proved to be very difficult, as humans vary dramatically in size and shape and live in complex environments.

When the Nintendo Wii was introduced in 2006, it was the first system that could detect three-dimensional input and give feedback to where the user was in physical space. When the Kinect was introduced in 2010, it was the first system that did not require a special controller to be held by the user. The Kinect was discontinued for gaming but remains a reliable component of many virtual reality systems in education, training, and medicine. With the Wii and Kinect, the groundwork was laid for truly immersive virtual reality environments.

In 2012, several developers started the Oculus Rift, which requires the user/player to wear a cumbersome headset wired to the computer (https://en.wikipedia.org/wiki/Oculus_VR). As the product has evolved, there has been a move to eliminate the wires so the player/user may move freely in space. Managing the sensors for each headset when there is more than one player has been the biggest challenge. Many users/players have been quite pleased and feel that for the first time, they have a true virtual reality experience. Oculus's success has inspired others such as Microsoft's HoloLens, Amazon's Sumerian, and Google's Cardboard.

The three principles of interaction, immersion, and imagination guide the success and intensity of a VR application (Rebelo et al., 2012). The player/user must feel that they are immersed in the environment and have a high level of interaction with the virtual system. This can be accomplished by the 360-degree view of the cameras, the lack of camera lag, the surround sound, and the integration of these elements in real time. The virtual environment must respond to where the user is in three dimensions and to the user's input through motion trackers and sensing gloves. Some systems also allow users to wear a Lycra suit embedded with sensors. Interaction is when the systems of sound/motion/video/pictures respond immediately to the player/user. If the player/user moves forward, the "sound" of their feet and the movement of the video must sync perfectly with the sensors. Lag in the system will contribute to motion sickness in the player/user.

Immersion is the sense that the user has that they are actually in the virtual environment (Rebelo et al., 2012). There are three types of immersion: full immersion, semi-immersion, nonimmersion. In full immersion, the user wears a headset/goggles and maybe a suit. In full-immersion systems, the user has a 360-degree view of the virtual environment. In semi-immersive systems, these are similar to special rooms with projectors that create a virtual environment. The user interacts in the room, and the projected environment changes as a result of gestures and other actions. In nonimmersive environments, users move through a simulation that is presented on a desktop or laptop machine. User involvement happens when the player/user is motivated by the system to continue through the simulation or game. This is through the narrative, the characters, or the entertainment value. A break in any of these can cause a break in all and a feeling of unrealness in the user, which then interrupts the imagination.

Imagination is the creation of the virtual environment in the user's felt experience. The user knows that they are not actually in another world. The synchronous sights/sounds and feedback help the user create the imagined world and interact with it. The amount of realism can be provided in the virtual environment by high-quality equipment, and the tasks that the user is asked to perform must be easily doable and provide a great deal of interactivity. Both of these provide the user with motivation to continue through the virtual environment and provide the user with the feeling that they are experiencing the environment in its truest sense.

There are several advantages to using a virtual reality system to simulate an environment rather than to use the actual environment. As with other types of automation, when considering an appropriate use, think of the three Ds: dirty, dangerous, and dull. For example, VR is often used in medical training where it would be dangerous to expose a patient to unskilled practitioners. The virtual environment can provide realistic sounds and sights to inform new practitioners when they are doing a procedure correctly or not based on the virtual patient feedback. In military operations, the risk to personnel is great during encounters with the adversary. A virtual environment based on previous examples can help train the proper behavioral and tactical responses without endangering personnel. Finally, in aerospace flight training, there are long lulls of time where the pilots may be tempted to disengage with the aircraft and then have to reestablish situation awareness to address an emergency. A virtual environment can train pilots on important tasks during these lulls as well as train them on strategies to reengage quickly.

VR-related nausea and motion sickness are still challenges in some individuals. Rebenitsch and Owen (2016) provide a nice review of the technological issues that cause VR-related motion sickness. Rebelo and colleagues (2012) suggest that repeated exposure, breaks, and optimization of the cues in the virtual environment will help. They also note that small children and people over the age of 25 years do not seem to have the VR-related motion-sickness problems. There are two questionnaires used to assess simulator sickness and motion sickness in virtual environments. The first is the simulator sickness questionnaire (Kennedy et al., 1993). The second is the motion sickness assessment questionnaire (Gianaros et al., 2001). During the virtual simulation, other questionnaires can quickly assess the user's state: the short symptom checklist (Cobb, Nichols, & Wilson, 1995), the misery scale (Wertheim et al., 1992), and the fast motion sickness scale (FMS; Keshavarz & Hecht, 2011).

REFERENCES

Cobb, S., Nichols, S., & Wilson, J. R. (1995, December). Health and safety implications of virtual reality: In search of an experimental methodology. In *Proceedings of FIVE '95 Conference, Framework for Immersive Virtual Environments* (pp. 227–242). Queen Mary & Westfield College.

Gianaros, P. J., Muth, E. R., Mordkoff, J. T., Levine, M. E., & Stern, R. M. (2001). A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, Space, and Environmental Medicine, 72*(2), 115.

Heilig, M. L. (1962). *U.S. Patent No. 3,050,870*. U.S. Patent and Trademark Office.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3*(3), 203–220.

Keshavarz, B., & Hecht, H. (2011). Validating an efficient method to quantify motion sickness. *Human Factors, 53*(4), 415–426.

Rebelo, F., Noriega, P., Duarte, E., & Soares, M. (2012). Using virtual reality to assess user experience. *Human Factors, 54*(6), 964–982.

Rebenitsch, L., & Owen, C. (2016). Review on cybersickness in applications and visual displays. *Virtual Reality, 20*(2), 101–125.

Wertheim, A. H., Ooms, J., De Regt, G. P., & Wientjes, C. J. E. (1992). *Incidence and severeness of seasickness: Validation of a rating scale* (TUDelft, Faculty of Marine Technology, Ship Hydromechanics Laboratory Report 957-O). TNO-IZF A-41.

CHAPTER 18

VISION AND VISUAL SEARCH

GOALS OF THE CHAPTER

- Learn the different types of visual searches
- Understand how people search for things in their environment
- Understand what causes people to end a search

ASSESSMENT

- What are three things that influence a person's visual search?
- What is the difference between a serial search and a parallel search?
- What do the gestalt principles do for us in terms of search?

Have you ever searched for your friend in a crowded room? Maybe you have looked for your car in a parking lot full of similar cars? If so, you have engaged in a visual search. **Visual search** is when we direct our gaze to process information relevant to a target (your friend, your car). Since only a small portion of your eye processes detail, you move your eyes to potential targets as you take in the relevant information and decide if it is indeed your friend or your car. As you take in a small portion of the scenery in search of the friend or car, we say that you are sampling the environment in search of your target (friend/car). You might think that you see your friend walking across the room; your eyes will follow the object in order to catch up and see if it is really your friend. This is called saccade—your eyes will move to catch up. During this time, when you move your eyes to catch up, you are temporarily ignorant of what you are seeing in the intervening time. Let's say that it really is your friend. You follow her across the room trying to decide if she sees you. This is pursuit. Once you see her, you may look for a bit or dwell on her image as you discern if it is really her or her sister.

Visual search (http://scholarpedia.org/article/Visual_search) is the term that human factors psychologists use when someone is using only their vision to process items while looking for a target. Think of looking for your phone or your keys—you search for them. Parallel search is when you are searching for more than one item at a time, such as in an airport security scanner. The person looking through the bags is looking for a range of banned items, not just guns. Serial search is when you are looking for items one by one. This is the most common type of search, such as searching for a book in the library. You look for the approximate location, and then within that location, you find the shelf, and on the shelf, you look through the books and find the one that you are looking for (http://scholarpedia.org/article/Visual_search#The_mechanics_of_search).

All of these are visual-search terms. A person moves their useful field of view toward the item that may match their search. They try to maximize the usefulness of the search while minimizing the costs. In other words, they try to be efficient. Several conditions specific to the person can influence a visual search: **mental models**, **event rate**, **arrangement**, **memory**, and **processing/search strategies**. First, people will have an idea of what to expect. This can also be called a mental model. They might expect the friend to be talking to a certain person or be at the bar. They might expect the car to be in the middle of the lot: "That is where I usually park." In other words, the expectations will guide the search.

People will also have an expected event rate and arrangement. They may expect the friend to arrive before 8 p.m. or not arrive at all. They may expect the event to happen in some circumstances and not others based on previous experience, their mental model, and other factors. They may expect to see the friend as they scan the room horizontally but not vertically. Finally, a searcher's memory is imperfect. They may misremember how they found the friend last or how they found the car the last time and might repeat the same mistake. Finally, searchers may use a particular strategy that was not successful in previous attempts, and they may have a very narrow focus or cognitive tunnel as they search.

They may become so fixated on a particular area that they fail to look in other places. In addition to impairment on the person who is doing the search, there are also potential impairments in our perception system.

Gestalt psychology is a subdiscipline in psychology that focuses on perception and sensation. It proposes several principles on how we organize what we see. These were first proposed by Wertheimer in the early century (http://scholarpedia.org/article/Gestalt_principles). These principles suggest that we group visual items together in certain ways. The most basic gestalt principles are as follows: figure/ground articulation, the proximity principle, and the similarity, continuity, and closure principles. These principles govern how we separate a figure from the ground, how we know which items are in front of other items, and how we see some separate items as all in one group due to their similarities or differences.

There are a few different theories of how we search for things. These models help human factors engineers determine how long it will take a person to find a target given certain constraints. One model by Drury (1975, 1990, 2006) states that there are two stages. In **Drury's model**, the first stage is related to how likely the person is to find the target. The probability of finding the target increases as the search time increases (Wickens & Hollands, 2000). However, it will diminish when the search does not include the whole search field or useful field of view. In the second stage, the person makes a decision and uses the expectancy of a miss to set a criterion of whether something is the target or not. This would be an example of a serial search. The person inspects each item in turn until the target is found. When the target is found, the search is terminated. This is called a serial self-terminating search.

In Drury's model, the time that a person takes to conduct a serial self-terminating search can be modeled by the equation below. If each inspection takes about the same amount of time (I) and the expected location of the target is unknown, then it is possible to predict the average time that it will take to find the target. In the equation below, T stands for time, N stands for the total number of items to be inspected, and I stands for the average inspection time for each item. The division is necessary because on average, the person finds the target after about half of them are searched (Wickens et al., 2013, p. 79).

$$T = (N * I)/2$$

People in the Western world tend to search in the same way that they read: top to bottom and then left to right. A parallel search can occur when the target has a defining characteristic, such as color, size, brightness, and motion. This makes the target "pop out" from all the other items. Some researchers believe that this parallel search takes few attentional resources and can be done all at once, hence the term *parallel search*. As the number of distractors or similar nontargets increases and as the number of features a person must search for increases, the search will become more parallel.

REFERENCES

Drury, C. G. (1975). Inspection of sheet metal: Model and data. *Human Factors, 17*, 257–265.

Drury, C. G. (1990). Visual search in industrial inspection. In D. Brogan (Ed.), *Visual search* (pp. 263–276). Taylor & Francis.

Drury, C. G. (2006). Inspection. In W. Karwowski (Ed.), *International encyclopedia of ergonomics and human factors* (Vol. 2). Taylor & Francis.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance*. Prentice Hall.

Wickens, C. D., Lee, J., Liu, Y. D., & Gordon-Becker, S. (2013). *Introduction to human factors engineering*. Pearson Higher Ed.

CHAPTER 19

AUDITION AND NOISE

GOALS OF THE CHAPTER

- Learn how auditory signals work and when to use them
- Understand the benefits and drawbacks to using auditory signals
- Understand some of the challenges in interpreting audition computationally

ASSESSMENT

- What are 5 of the 11 instances where an auditory signal is preferred over a visual signal?
- What is JND, and how is it used?
- What are the four tasks that sound conveys?

HOW THE EAR WORKS

Sound is created when a physical object changes the air waves around it. This can be through electronic transmission, such as when you listen to music through headphones, or it can be through a clash of two objects, such as when a dish hits the floor. In both cases, there was an action that happened, and sound was produced. Sometimes the produced sounds are pleasing, and sometimes they are not. This depends on the frequency of the sound waves and the amplitude of the sound waves.

Frequency is produced by the creation of a waveform or a disturbance in the air. Frequency is perceived as pitch by the tiny bones in our ears and then the cilia in the eardrums, which convey the information to the brain. The brain interprets this information as sound and attributes categorization, meaning, emotion, and further processing into memory. Frequencies are measured in hertz (Hz).

The **amplitude** of sound waves is perceived as loudness. The closer we are to the dish that has hit the floor, the loudness, or the amplitude, is greater. Amplitudes are measured in decibels (dB).

This is where the simple part ends. Sound waves are complex combinations of different frequencies and amplitudes. When sound is examined, sections are created that separate the different frequency bands so they may be examined individually.

Sometimes, sounds will mask other sounds, such as a hairdryer drowning out or masking the sound of someone talking to you. Headphones and earbuds use the masking technique along with other techniques to make background noise disappear so that you may hear the music better. One of the problems with using audition to create warnings and signals and present other system information is that we are rarely in a sound vacuum. The auditory information must always compete with other information in the environment, which makes it dynamic and unpredictable. However, according to Sanders and McCormick (1993, p. 169), there are a few instances where auditory alerts are appropriate. These are the following:

1. When the origin of the signal is itself a sound (i.e., breaking)

2. When the message is simple or short

3. When the message will not be referred to later

4. When the message deals with events in time

5. When warnings are sent or when the message calls for immediate action

6. When continuously changing information of some type is presented, such as aircraft, radio ranges, or flight-path information

7. When the visual system is overburdened

8. When speed channels are fully employed

9. When illumination limits the use of vision

10. When the receiver of the sound is moving from one place to another

11. When a verbal response is required

Imagine that you are monitoring playing children for problems. There are a lot of children, too many to watch individually. You would monitor the children for sounds of fighting (e.g., grunting, tearing, thud), raised voices, or crying. In these cases, audition works better than a video display because the message is simple and short, you are dealing with an event in real time, the event would call for immediate action, and the information is changing rapidly. Parents become very good at this task, so good that children will be amazed that their parents are alerted when they feel as if they have only intended to do something but have not actually done it. Parents will exclaim, "Oh, I have eyes in the back of my head," when, in reality, they have become very adept at processing auditory signals from children.

Another example would be when you are looking for your car in a dimly lit parking lot, all of the cars may look as if they are the same color, as the dim lighting has made color perception difficult. Sunset and sunrise are the most difficult times to see for humans, as things appear to be the same intensity and saturation in color in our environment. If something is not moving at dusk or daybreak, it is difficult for us to see. If you have a keychain fob with an unlock or alert button, you can press it and hear your car chirp. In a similar example, when you forget where you placed your cell phone last, you can call the phone, and the sound helps you locate it.

In these examples, the visual system is not the appropriate information sense because it is overburdened, not available, or unable to process the amount of information quickly. Sound is quicker than visual information. However, sound conveys only a few pieces of information. These are **detection**, **relative discrimination**, **absolute identification**, and **localization** (Sanders & McCormick, 1993, p. 169). In **detection**, either the sound is present, or it is not, such as in a fire alarm in a fire station. The firefighters identify the sound and start their process of getting ready to fight the fire. When the sound is absent, the firefighters are monitoring for the sound to appear while doing other things. In **relative discrimination**, the person must identify when the sound is at a level where they must act. The children will make small crying sounds or have loud voices at times, or there will be thuds and grunts when playing. The caregiver knows the level at which these things demand their presence. In **absolute presence**, as you drive your car, the squealing sound of your brakes indicates that the brake pads need to be replaced or that other parts of your car need maintenance. The presence of a specific sound signals an alert. **Localization** occurs when you are trying to find your phone or your car or another item. For example, when large trucks back up, the driver cannot see what is behind the truck and relies on anyone behind them to move when they hear the signal approaching them.

When implementing auditory signals when designing an alarm as part of a system, there is a specific process to determine the proper frequency and amplitude of the alarm. This specific process is called a **just noticeable difference** (**JND**) process. The signal is tested on the target users of the future alarm system. When 50% or more of the users tested can detect the signal in a normal noise environment for which that signal is present, the signal passes the test. In addition, people will report if the signal sound is pleasant or unpleasant.

Sometimes it is appropriate to use visual signals, and other times auditory signals are best. Auditory signals are best used when an additional signal is necessary and the other signal is visual or when a visual signal cannot be used. When pairing an auditory signal with a visual signal, if you are trying to convey an increase in the level of the alert, the auditory pitch and intensity should also increase in sync with the visual signal. The opposite is true for a decrease in the level of the alert. When an emergency needs to be conveyed, human response is quickest to sound. When an emergency needs to be conveyed, shrill and intense tones are best. Often in emergency signals, there will be a combination of auditory and visual alerts.

Other times, a two-stage signal is needed. The first part of the signal is to gain human attention. The second part is to convey a message. In these types of signals, the two auditory signals should be distinct and separate from any background noise. The same signals should be used consistently so that each time the signal is heard, it is distinct and conveys the same information each time. The signals should be no longer than necessary in order to be heard and recognized according to our JND standard. More information about auditory signals can be found in Sanders and McCormick (1993).

Signals, alarms, and alerts are complex tools, as they attempt to garner human attention and promote action. The duration and tone of the signal in terms of feedback should be brief and pleasant for positive feedback (i.e., the system confirms that this was correct). The duration and tone should be slightly longer and unpleasant for negative feedback (i.e., there is something wrong in the system). Specificity matters in terms of signals; the more specific the signal is to the problem, the better.

Ideally, signals should function as the "**cocktail party**" effect functions. This effect refers to the phenomenon of standing and talking at a party and then hearing your name being mentioned across the room. You are subconsciously monitoring for your name. You are able to hear your name through a great deal of noise, perhaps more noise than you would normally hear something. Important alerts and warnings should function in a similar way. Humans should be able to hear and recognize them over a great deal of environmental noise.

REFERENCE

Sanders, M. S., & McCormick, E. J. (1993). *Human factors in engineering and design* (7th ed.). McGraw Hill.

CHAPTER 20

HAPTIC CONTROLS AND VIBRATION

GOALS OF THE CHAPTER

- Learn how different types of interfaces are used to control a system
- Understand unique ways to negotiate control in automation

ASSESSMENT

- How are a GUI and a TUI the same, and how are they different?
- What is one use of a haptic interface?
- What is one use of a proprioceptive interface?

Sometimes visual feedback is hampered by lighting or a complex environment. Sometimes auditory feedback must compete with loud noise in the environment or there is a need for near-complete silence. These are opportunities for a different type of feedback based on our skin and sense of our body in space. Feedback that incorporates touch is called tactile feedback. Many laptops have a touchpad instead of a mouse. Feedback that incorporates pressure against the skin is called haptic feedback. Feedback that incorporates limb or body position is proprioceptive feedback. Finally, feedback that incorporates motion is called kinesthetic feedback.

First, let's focus on **haptic feedback**. Underneath your layers of skin, you have neurons that respond to pressure. These neurons tell you when you have grasped a cup in your hand so you can carry your coffee across the room. The neurons tell you that there is a wall when you stub your toe. Haptic feedback is very important in virtual environments, as it tells the user when they have grasped an object in virtual space. Haptic feedback is conveyed by haptic interfaces such as gloves or controllers. **Haptic interfaces** are also called **tangible user interfaces** (**TUIs**). In contrast to visual **graphical user interfaces** (**GUIs**), TUIs offer features for the visually impaired to create music or interact with a system without the need for vision (Vetter, 2019). This application of TUIs is very useful; however, haptic-interface development is still in its beginning stages.

Outside of virtual environments, TUIs are used most when natural information is needed about the environment or the user needs to maintain situation awareness in a complex environment or while doing something else. For example, a haptic gas pedal might alert a driver to unsafe road conditions that cannot be seen, such as black ice, veering off the road, or other hazards (van Paassen et al., 2017). The second case already exists in the sense of grooves cut in the side of highways that make the vehicle's tires bounce. When you drive over these grooves, your tires bounce in a way that makes it very clear that your vehicle is outside the lines.

In automated vehicles, TUIs can work cooperatively with users to cede/relinquish or assume control unobtrusively when the person may be doing other things. When haptic interfaces are used in this way, they are called shared control interfaces and are most often used in autonomous situations where a robot can take full control over a vehicle or device. According to Terken, Levy, Wang, Karjanto, Yusof, Ros, and Zwaan (2017), several interfaces are being developed to increase trust in automation by having the automation ask for control from the user with either a haptic or a gesture-based interface. One of the interfaces, named Stewart, uses **servos** and an **Arduino computer** to communicate the automated system's intention to swerve or speed up in an automated car. A second gesture-based interface called Comrade also communicates the automated system's intention through a light across the dashboard that either moves toward the steering wheel or moves away from the steering wheel. The movement of the light to the steering wheel tells the driver that the automation wishes to give control to the human driver. Both the driver and the car can control the light to negotiate who has control of the vehicle: "The [human] can push the light aside to take over control or pull the light to the steering wheel to hand over control" (Terken et al., 2017, pp. 4, 5).

In the shared-control application of a TUI, the development team hopes to increase trust in automation between the users and the automation through a shared situation awareness. This application also inspires the human user to stay in the loop and be aware of what the automation is doing. The automation does not take the ability to control the environment away from the human users, as the humans may choose to disengage the automation or refuse to relinquish control of the vehicle to the automation.

Similar TUIs are being developed in medical robots; however, the strength of the human user is a consideration. Some humans have a strong touch, while other humans have a light touch. Some of this discrepancy is due to the expectancy of gain by the human and is dependent on what the TUI is controlling. A human who is using a TUI with a large industrial robot may expect a large response to a slight touch. A human who is using a TUI with a virtual environment may expect a small response to a heavy touch. As the user continues to use the haptic device, the user will calibrate their touch to the response it generates in the device (van Paassen et al., 2017). In this way, calibrating the haptic control is important, but recalibration is only necessary from the system side, as the human will adapt.

In both haptic (TUI) and **proprioceptive feedback**, the user may be wearing a device such as a glove or a vest instead of a discrete control button or pad. When this is the case, the user's perceptions of the glove or vest are very important. At least one survey has been constructed to measure the amount of anxiety, movement, perception of change, harm, attachment, and positive or negative emotion associated with one of these devices (Knight & Baber, 2005). A user who is not comfortable with or in the device is likely to be less effective. Small changes in the construction of a glove or vest can make a significant difference, and a survey should be used to measure human perception.

Proprioceptive feedback conveys where a body is in space. A mouse or a joystick conveys where the cursor is on the screen in space and can be considered proprioceptive feedback. An isometric control is one that responds only to pressure but does not move. A hard press will move the item farther than a soft press on the control (Wickens et al., 2013). Proprioceptive feedback is important in flight simulators, medical applications, and all virtual environments.

Proprioceptive interfaces are usually embedded in the design of the system itself, as in boots that help stroke patients learn to walk (Ramakrishnan et al., 2019). In the boots application, the system within each boot measures the gait of the wearer and calculates the action of the boot. At first, interfaces such as this can be difficult for the human until the system and the human have acclimated to each other. Once the partnership has been made, the boot makes small adjustments in an effort to retrain the human through operant conditioning. As the human makes correct movements, the system will work properly. As the human makes incorrect movements, the system doesn't respond. The human adjusts to the automation. In this instance, the use of proprioception allows the system to retrain the human to walk.

Beyond these applications, proprioceptive interfaces are yet to be developed. As with TUIs, this is a rich area for development. It allows humans to engage in other tasks or use the device in a sound and visually rich environment. Moreover, it allows humans who have limited capabilities in vision or audition to control a system.

REFERENCES

Knight, J. F., & Baber, C. (2005). A tool to assess the comfort of wearable computers. *Human Factors, 47*(1), 77–91.

Ramakrishnan, T., Kim, S. H., & Reed, K. B. (2019). Human gait analysis metric for gait retraining. *Applied Bionics and Biomechanics, 2019*, 1–9. https://doi.org/10.1155/2019/1286864

Terken, J., Levy, P., Wang, C., Karjanto, J., Yusof, N. M., Ros, F., & Zwaan, S. (2017). Gesture-based and haptic interfaces for connected and autonomous driving. In *Advances in human factors and system interactions* (pp. 107–115). Springer.

van Paassen, M. R., Boink, R. P., Abbink, D. A., Mulder, M., & Mulder, M. (2017). Four design choices for haptic shared control. *Advances in aviation psychology*: Vol. 2. *Using scientific methods to address practical human factors needs* (pp. 237–254). Routledge.

Vetter, J. (2019, March). Tangible signals-physical representation of sound and haptic control feedback. In *TEI 2019: Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 741–744). ACM.

Wickens, C. D., Lee, J., Liu, Y. D., & Gordon-Becker, S. (2013). *Introduction to human factors engineering: Pearson new international edition*. Pearson Higher Ed.

CHAPTER 21

SPATIAL SENSE AND MAPS

GOALS OF THE CHAPTER

- Learn how people understand two-dimensional space
- Understand the theory of FORT
- Translate ideas about display design to map design

ASSESSMENT

- What is the theory of FORT?
- What are the differences between egocentric and exocentric views?
- What is spatial cognition, and why is it important in understanding maps and displays?

DISPLAYS

Displays convey the system state and what the system is doing, as in the fuel level in an automobile. The system could be moving physically, as in an automobile, or it could be static. In moving systems, the movement in the display should match the direction or the intensity of movement in the part of the system that the display monitors. This is called the **principle of the moving part** (Roscoe et al., 1981). For example, in a speedometer, the arrow moving to the right or to the positive side of the number line is associated with greater speed. There are instances when the principle of the moving part should be violated and instances where this principle can be combined with either pictorial realism or the **proximity compatibility principle** to describe to the human what the system is doing. In pictorial realism, the display imitates what the system is actually doing. For example, if a plane were banking left, a pictorial realistic display would have a small icon that looked like an airplane with the right side higher than the other. The proximity compatibility principle (Barnett & Wickens, 1988) states that displays for tasks that require intense attention or concentration should be grouped together or have a high proximal location. Displays for tasks that require moderate or low attention or concentration should be grouped apart from each other to allow room for the displays of tasks that require intense attention.

Display designers will use artistic cues to help describe what the system is doing and how it is doing it. These could be shadows, linear perspective, binocular disparity (e.g., when the scene should appear slightly different to each eye), convergence (e.g., when parts of the scene seem to align), and accommodation (e.g., when things in the foreground are detail rich while things in the background have less detail). Shadows can be used to indicate where one item is in relation to another. **Linear perspective**, **convergence**, and **binocular disparity** are all used to communicate distance or travel toward an item. Linear perspective is when something that would normally be large is small, such as a semitruck in the distance. Convergence is when two lines that would normally be parallel, such as the traffic lines in a road, get closer and closer together until they become a single point in the far-off distance. Binocular disparity is when the two eyes would perceive different parts of an image, which conveys distance. With this in mind, an item that is close appears with more detail, while an item that is far away has less detail or could be slightly blurred.

MAPS

While these principles will assist most designers with screen design, much of the work in how people understand a two-dimensional space and relate it to the three-dimensional world around them is from work on map reading and navigation. Map reading and orienteering use landmark knowledge, which is the mental images of the places. If you are looking for a friend's house and they say, "I live two doors west of the Flintstones' house," that person is attempting to conjure an image from your knowledge of cartoons (*The Flintstones*) and an image of their large stone house. Your friend is trying to convey landmark knowledge. Landmark knowledge can only be used when a person has direct interaction with the environment. In other words, your friend would have to live there to know that there was a house to the west that looked like it was made of stone. There would be another way that your friend could describe how to get to their house: through route knowledge. Route knowledge often uses landmark knowledge. Your friend might say, "Take a left out of campus, drive to the first church on the right, and take a right there. Then my house is one block away on the left, after the stop sign." Finally, your friend could draw you a map. This is called survey knowledge. The map would have places other than the direct route. Landmark knowledge and route knowledge are called egocentric knowledge, as one must have personal knowledge, and that knowledge is described from that person's point of view. Typically, each person would describe it slightly differently. Survey knowledge is said to be **exocentric**, as it is more generalized and from anyone's point of view. Typically, everyone would draw the same map with similar routes to the same place (Wickens & Hollands, 2000).

The **frame of reference** or viewpoint on a map can help or hinder a person when using one to orient themselves spatially. Imagine the inside of your home and the different ways in which you could portray the route to the bathroom from the living room. You could have a virtual reality camera and allow the person to see landmarks/anchors as you travel the route. This would be the **egocentric view**. You could also have a camera in the corner of a key room that would allow the person to see the route being traveled. This would be a **tethered view**. Finally, you could have a camera in the ceiling that showed the route downward. This would be a **plan view**. This would be a plan view. In the plan view, you can imagine that the farther away the camera is from the actual route, more of the house or apartment will be seen. In this case, the map becomes more exocentric, as the navigator must imagine themselves in a smaller and smaller portion of the map. Closer maps, or more egocentric maps, are easier to navigate, but they are made for fewer individuals. Far-away maps, or more exocentric maps, assist more people in understanding spatial distance.

Different applications such as MapQuest or Google Maps offer an ideal combination of both **exocentric** and **egocentric** viewpoints, as the navigator can zoom into the precise location and turn off or on different levels of detail, such as landmarks/anchors, "street view," and traffic patterns. However, one of the challenges that these online maps still face is the field of view. Typically, in "street view," the map takes on the forward field of view (FFOV). While this is the preferred way of navigation, many people need peripheral judgments to help them understand distance and space. The limited capacity of online maps cannot offer these accommodations and instead have offered turn-by-turn verbal directions. The timing of the directions can be tricky and imprecise, causing the driver to turn too early or too late, which leads the driver to miss a turn in heavy traffic. In addition, the online maps do not always understand bodies of water. When I was living in Florida, I learned to use paper maps, as an online map application would often instruct me to travel through a lake. However, the online maps applications are getting better as developers learn of the challenges.

YOU ARE HERE (YAH) MAPS

You are here (YAH) maps are orientation maps that you see on campuses or at shopping malls. Typically, they have the major landmarks portrayed on the map with a star or arrow indicating where the viewer is located, or "you are here." These maps are easier to read when the map is oriented with north as pointed upward and when forward travel is also oriented up. Sometimes it is not possible to accomplish both of these goals; an effort to accomplish either will help the person navigate quickly and develop a better understanding, or **cognitive map**, of the environment. Developing a cognitive map of a specific space or understanding space and distance seems to be a trait in which humans differ. Research is ongoing to identify the factors that contribute to someone being able to create a cognitive map quickly and help those who have trouble developing cognitive maps.

SPATIAL COGNITION

Spatial cognition and cognitive maps are the two terms that refer simply to the human's understanding of where things are in three-dimensional space. This can be projecting where something is going to go given a trajectory, translating a two-dimensional map into an understanding of where to turn next to arrive at a destination, imagining where something is so that you can give directions or create a map, and understanding where your body is in relation to other landmarks/anchors in your environment. Everyone has some level of spatial cognition so that we don't trip over furniture. However, some people have a particularly acute sense of spatial cognition and are very good at using it in sports, driving, dancing, and academic endeavors such as engineering. In fact, most engineers are tested at one point for their spatial cognitive ability, as they must be able to imagine the items that they are creating.

Typically, spatial cognition has been tested using a task called **mental rotation** (Shepard & Metzler, 1971). Mental rotation is the construct that describes how a person translates what they are seeing into a useful spatial model. However, with "brain performance" apps, many of these measures are available to the public. This creates a **practice effect** in the general population. It is still unknown if cognitive tests improve IQ, as the resulting evidence is mixed. However, the worry that a practice effect taints experimental psychology results is real. A practice effect is when someone gets naturally better at something over time simply by repeated exposure. They know what to expect, and the repeated exposure has allowed the person to create better answers and react faster. An example is if you repeatedly took a final exam several times and were given the answers to each question after you answered it. Naturally, you would get better over time and would eventually earn a perfect score in record time.

In pilots, spatial cognition is key to being able to navigate the plane. Pilots must translate between two frames of reference: **egocentric** and **world-centered views** (Aretz, 1991). The egocentric view can be thought of as the pilot's view out of the cockpit of the plane. The world-centered view can be thought of as the map with north facing up. Pilots use mental rotation to align their egocentric view with the world-centered view and **triangulation** to establish the relationships between the two views. Sometimes, maintaining navigational ability competes with the pilot's other responsibility of flying the plane. In large aircraft, there is a separate pilot responsible for navigation. In small aircraft, one pilot takes on both tasks. It is for these pilots that spatial cognition and map reading are key. Map design could make a difference in the plane's safe arrival. Research in map reading has found that if the map shows an egocentric view, with the direction that the pilot is headed in the up position, this enables the pilot to localize where they are fastest. When the map is in the north-up position, this enables the pilot to figure out where they are in relation to landmarks/anchors fastest (i.e., "I am north of the airport, and I need to turn left").

These ideas are based on a theory on map reading called **frame of reference transformations** (**FORT**; Wickens et al., 2010). FORT describes the mental rotation that people must make to translate the map from a two-dimensional piece of paper to a three-dimensional representation of the environment. Some of the solutions that the FORT theory suggests are the following: (1) training on how to translate the map and (2) multiple maps for the same location, such as Google Maps (Wickens et al., 2013). As a person develops an understanding of a new city or locale, they are engaging in translating the map training. First, they learn the landmarks that people use when describing routes. Next, they learn the main and alternate routes. Finally, they learn **spatial proximity**—as in, a **mental representation**—of the area. By the time that people get to the final stage, they have a good idea of how far something should be from a specific location (Wickens et al., 2013). Google Maps has an excellent example of multiple maps for the same purpose. In this application, you can view the directions from a landmark perspective with satellite view, a turn-by-turn perspective, or a street-view perspective. Each perspective allows the reader to add more information to their spatial model and relate it to what they already know about the area.

REFERENCES

Aretz, A. J. (1991). The design of electronic map displays. *Human Factors*, *33*(1), 85–101.

Barnett, B. J., & Wickens, C. D. (1988). Display proximity in multi-cue information integration: The benefits of boxes. *Human Factors, 30*(1), 15–24.

Roscoe, S. N., Corl, L., & Jensen, R. S. (1981). Flight display dynamics revisited. *Human Factors, 23*(3), 341–353.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701–703.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance*. Prentice Hall.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology and human performance*. Prentice Hall.

Wickens, C. D., Keller, J. W., & Small, R. L. (2010, September). Left. No, right! Development of the frame of reference transformation tool (FORT). In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 54, No. 13, pp. 1022–1026). SAGE.

Wickens, C. D., Lee, J., Liu, Y. D., & Gordon-Becker, S. (2013). *Introduction to human factors engineering*. Pearson Higher Ed.

CHAPTER 22

CONTROLS AND CONTROL PANELS

GOALS OF THE CHAPTER

- Learn about the different types of controls
- Understand what makes one type of control preferable to another in a particular application
- Understand the challenges of different control types

ASSESSMENT

- What is gain?
- What is lag?
- What is Fitts law?
- What is the speed-accuracy trade-off, and why is it important in design?
- Why is an eye-controlled device impossible for a human to use?

Controls are a means of transmitting information, such as a response or input, between the human and the system or computer. There are several types of controls, but they fall into two classes: (1) **discrete** or (2) **continuous** (Sanders & McCormick, 1993). Discrete controls turn something on or off or make a single choice among many options.

For example, when you turn on your stove's burner or press a button on your microwave, you are making a discrete choice. A continuous choice control allows the user to set something along a continuum. When you turn on your oven, you choose a temperature for the oven. This is a continuous choice; the oven can be set at 350 degrees or at 349 degrees. When you press the accelerator on a car, again, it is a continuous choice. When you move your mouse, this is a continuous choice. Table 22.1 is adapted from Sanders and McCormick (1993, p. 336) and shows the four types of controls with examples.

Table 22.1. The Four Categories of Controls

	Discrete control (on/off)	Continuous control (set a level)
Force required: small	Toggle switch	Joystick
Force required: large	Foot push button (automobile high beams)	Accelerator pedal in an automobile

Source: Sanders and McCormick (1993, p. 336).

IDENTIFICATION OF CONTROLS

There are instances where controls are mistakenly pressed. Sanders and McCormick (1993) state that in airplanes, the landing gear and flap controls are often mistaken, and in locomotives, the signal lights and fuel pump can be easily confused. When elderly or impaired drivers confuse the brake and the gas pedal, this is an identification of controls error. As these pedals are usually operated with the same foot and require the same movement in different areas of the floorboard, a simple misalignment of the foot can cause this mistake. An engineer in Japan has suggested changing the accelerator pedal to be a lateral push pedal next to the downward brake push pedal, which was mentioned in this article (https://www.nippon.com/en/news/yjj2019081601044/hopes-on-onepedal-growing-for-preventing-car-accidents.html). The location accuracy can be enhanced when switches are arranged vertically, with about 2.5 inches or more between them. When switches must be arranged horizontally, at least 4 inches are needed for humans to recognize them by location (Sanders & McCormick, 1993).

Indeed, as Sanders and McCormick (1993) state, the misidentification of controls is typically an issue that can be addressed by differentiating the similar controls by switching the direction of travel, the type of control, the location of the control, the texture of the control, or the size of the control. Several studies have identified different variations of texture and size in knob controls that humans can easily distinguish. These studies take into account the conditions of use and if the human is wearing gloves, which may dull the ability to distinguish control size and shape (Bradley, 1967, 1969a, 1969b, 1969c).

In optimal conditions, humans can view the controls and the control labels as they orient themselves to the system and operate the system. However, as anyone who has driven a rental car knows, this is not usually the case. In a new rental car, when it rains, the driver must find the windshield wiper controls while paying attention to the road. When the driver must operate the controls purely on touch and location, then the shape, size, texture, and type of control become the only identifying information. For rental car drivers, usually the windshield wipers are located on the steering wheel and have a similar look and feel across cars. In addition, they usually have the same three settings of intermittent, slow, and fast. The resistance within the control is usually about the same among cars. The interval between turning the switch to the on position and the wipers actually working is about the same: a 1–2-second delay. These two factors of resistance and operational interval are important when humans determine if they have chosen the right control or not. The resistance within the control consists of the amount of effort the human has to exert on the control and the ease at which the control moves (Sanders & McCormick, 1993). The human operator will quickly adapt to similar controls that have similar resistance qualities. There are also controls that have little to no resistance. These controls are best used in continuous applications that require small movements, such as an oven dial.

According to Fitts law (Chapter 7), the size of a control knob can be determined by how urgent it would be to access it in a timely manner. Controls that are rarely accessed or rarely needed can be small and distant from the optimal operator location. Controls that are rarely accessed but are important to be accessed quickly can be large and distant from the the operator's location. This principle is the basis of Fitts law (1954; https://en.wikipedia.org/wiki/Fitts%27s_law).

TELEOPERATORS

A **teleoperator** is a robot under direct human control with few automated components. The automated components that do exist would be to extend a person's reach, strength, or capacity to endure a harsh environment, such as outer space. Some professionals consider teleoperators a special type of robot (e.g., see Chapter 15), while other human factors professionals consider teleoperators a human-machine system designed to augment a human's physical skills (Sanders & McCormick, 1993). Let's consider a teleoperator as an augmentation device that allows humans to have extended control.

These types of teleoperators are used when people encounter the three Ds (dirty, dangerous, or dull) portion of a larger task. In the case of teleoperators, the dirty and dangerous situations are the most common. In dangerous situations, the human operator is at risk of serious illness or death if they encounter the environment, such as with nuclear waste management or deep-sea exploration. Creating the teleoperator is challenging, but it must match the human operators' kinesthetic and proprioceptive abilities. **Kinesthetic** refers to the way people move in space, and **proprioceptive** refers to how people know where their body is in space. The goal of most teleoperator developers is to create a device that feels as if it is a natural extension of the user's own arms and hands or legs and feet.

One of the primary challenges with teleoperator systems is communication between the robotic arms and the controls that the human operates. The **feedback loop** between the human and the robotic tool is critical. A tiny increase or decrease in the control could direct the robot to use too much or too little force. This is where haptic feedback and plenty of practice are useful. You can imagine that in a teleoperator system—for example, the da Vinci Surgical System—the delicate movements must be practiced, or the system could create devastating damage. More on the da Vinci Surgical System can be found here: https://en.wikipedia.org/wiki/Da_Vinci_Surgical_System. Within such a system, the sensitive ratio between the movement of the control that the human who operates the movement of the system's arm or tool is called **gain**.

Often humans using a teleoperator system need to see what they are doing. Large systems can use a camera that shows the entire field of view, such as in the mining telerobot here: https://youtu.be/ewRUlh48Aog. However, in surgical and outer space applications, an entire field of view may not be possible. To address this, developers will add cameras to the end of specific teleoperator appendages. However, cameras are imperfect, as they distort images, have time delays, must be well positioned, and must zoom in or out to give the correct viewpoint for the teleoperator to do their job. An example is here: https://youtu.be/K82fCFqnLBk.

Teleoperators can become visually disoriented and lose their spatial map when their point of view is isolated to several single cameras. The following is a video that describes the use of teleoperators controlling heavy cranes: https://youtu.be/35j9RrG1ncA. At the end of the video, you can view the cameras they used to orient themselves and create the connections. As you watch the video, consider how the crane operators understand the movement of the crane through the toggle controls on their waist packs and why their excellent spatial cognition and kinesthetic sense are necessary for them to complete the job without becoming disoriented and mistaking the movement of the crane in relation to their viewpoint and self-movement.

GAIN AND LAG

Gain must be set to a sweet spot for most operators to not experience instability: if the gain is too high, the system overshoots the target; if it is too low, the system undershoots the target, and there may be a perceptible delay or lag (Wickens et al., 2013). **Lag** refers to the time between when the control has been pressed and the system reacts. In this case, we will only refer to transmission lag, but there are other types of lag as well. In transmission lag, there is a delay between when the control is pressed and the system executes the command. In old computer systems, the delay was significant and perceptible between when a key was pressed on a keyboard and the letter would appear on the screen. In modern systems, that time exists, but it is nearly imperceptible. There is always transmission lag in a system, as the system must process the input. Calibrating user expectations to the system lag is part of determining the usability of a system. In our example before, when the driver turns on the windshield wipers, they expect that there will be a second or two before the wipers clear the windshield. In a cell phone, when the user presses a key on the virtual keypad, they expect that the image of the letter or number will appear nearly instantly. These differences in lag expectations affect how long the user will wait before they prompt the system to respond in a different way. These expectations are built over time with similar systems and similar experiences.

STIMULUS-RESPONSE COMPATIBILITY

In addition to lag and gain, humans will also expect that the control movement is in sync with the system movement. This is called **stimulus-response compatibility** or **display-control compatibility** (Wickens et al., 2013). First, the control should be near what it is controlling (Fitts & Seeger, 1953). For example, the dials on the stove should map to the stovetop element that they control. If a dial is on the far right, it should control the burner element farthest to the right. This is called location compatibility. When you turn the dial up, the burner element should heat up. This is called movement compatibility. The direction of the knob or dial should resemble the action that you want the system to take (Wickens et al., 2013).

SPEED-ACCURACY TRADE-OFF

In teleoperator systems, the quicker the operator makes a decision or executes a movement, the more likely it is that there will be an error. The more time that it takes, the more accurate the operator will be. This is called the **speed-accuracy trade-off**. As an operator increases their speed, the operator is more likely to incur more errors and be less accurate.

VERBAL CONTROLS

Verbal controls based on speech-recognition require specialized training. Increasingly, developers are embracing a universal design that requires little to no user training. Speech recognition can be classified into several dimensions: **isolated-word**, **connected-word**, and **continuous-speech systems** (Sanders & McCormick, 1993). Of these systems, the isolated-word control is the simplest to implement. In this type of control, the human utters a single word, and the system responds. Alexa, Cortana, and other current speech-recognition systems work well with a single-word utterance. In the single-word-utterance control, a predefined list of words controls the system. Here is one example: https://developer.amazon.com/en-US/docs/alexa/custom-skills/choose-the-invocation-name-for-a-custom-skill.html.

Connected words and continuous speech are more challenging for speech-recognition controls to recognize, as different humans' pronunciation of the same word differs and combinations of words meld together to seem as if they are one word (Sanders & McCormick, 1993). This often happens when one word ends in a consonant and the next word begins with a vowel. For example, if a human says "You gave the cat your dinner" to a speech-recognition system, it is interpreted as "You gave the catcher dinner" (Sanders & McCormick, 1993, p. 375). One of the ways to address this problem is to have a larger collection of words that the system can learn from. This is called "enlarging the corpus," with *corpus* referring to a body of words. However, as the corpus grows, computational time and power must grow as well. This is a significant problem that many organizations are working on. One of the main purposes of many household modern voice-recognition systems is to gather a corpus of words across many different humans in order to refine connected-word and continuous-speech systems. As you can imagine, this corpus or collection of human utterances must be vast in order to capture the diversity for a machine-learning application. Eventually, we hope speech-recognition applications improve as the corpus increases.

EYE-ACTIVATED CONTROLS AND BRAIN-ACTIVATED CONTROLS

These types of controls have been in development for several decades. Eye-activated controls rely on a device called an eye tracker to measure the location of the eye in relation to the target area of the control. An eye tracker is a complex set of infrared beams that measure the location of the eye from the reflections on the cornea (https://en.wikipedia.org/wiki/Eye_tracking). Most modern eye-tracking is noninvasive and fairly accurate. However, as a control mechanism, the use of eye-tracking can be problematic. Humans are highly distractible and find it difficult to control their eye movements; often, they shift their gaze momentarily or fail to maintain a gaze at the control for the specified amount of time.

Brain-activated controls or **brain-computer interfaces** (**BCI**) have had more success. These controls take a brain wave elicited at a certain frequency as a signal to control a computer or robotic device. The challenge with BCIs has been training the user to elicit the correct wave at the correct time. Also, humans have different skull densities and slightly different geolocations of specific brain wave activity. This has delayed the implementation of a broadly used BCI. Current uses include implantation of the electrodes in specific patients or systems developed with a certain segment of the population such as paraplegic patients. Preliminary work has spawned BCI games where developers can compete (https://youtu.be/5jGcNbQhbg8). In certain populations, BCIs have been very successful in providing control of full-body robots such as these: https://youtu.be/At3PHNkTWqg. More on BCIs can be found here: https://en.wikipedia.org/wiki/Brain%E2%80%93computer_interface.

REFERENCES

Bradley, J. V. (1967). Tactual coding of cylindrical knobs. *Human Factors, 9*(5), 483–496.

Bradley, J. V. (1969a). Effect of gloves on control operation time. *Human Factors, 11*(1), 13–20.

Bradley, J. V. (1969b). Glove characteristics influencing control manipulability. *Human Factors, 11*(1), 21–35.

Bradley, J. V. (1969c). Optimum knob diameter. *Human Factors, 11*(4), 353–360.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*(6), 381.

Fitts, P. M., & Seeger, C. M. (1953). SR compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology, 46*(3), 199.

Sanders, M. S., & McCormick, E. J. (1993). *Human factors in engineering and design* (7th ed.). McGraw Hill.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology and human performance*. Prentice Hall.

Wickens, C. D., Lee, J., Liu, Y. D., & Gordon-Becker, S. (2013). *Introduction to human factors engineering: Pearson new international edition*. Pearson Higher Ed.

SAFETY AND ERROR

CHAPTER 23

HUMAN ERROR

GOALS OF THE CHAPTER

- Learn about the types of human error
- Understand how humans interact with systems to produce an error

ASSESSMENT

- What is the SCM model?
- What are latent variables?
- How are stress and performance related?
- What does a reliability analysis demonstrate?

Key to many aviation, medical, and industrial accidents is human error. In fact, most errors in these systems are due to normal humans misperceiving, misunderstanding, and lacking an awareness of the system within which they are working. For example, a 6-year-old boy named Michael had a benign brain tumor removed. The physicians were checking that all was in order before releasing him. The last step was an **MRI** (**magnetic resonance imaging**). They sedated Michael and took him to the MRI room. An MRI machine has a powerful magnet used for imaging. Once the MRI machine is active, anything metal will stick to it. Medical personnel must be careful to leave all metal objects outside of the machine's room.

As Michael was having the MRI, the anesthesiologist noticed that his oxygen levels were low. Usually, the MRI room has its own oxygen system for the patients, with tanks kept in a safe compartment away from the machine. That day, the oxygen system was not functioning, and the MRI technicians went to investigate why it wasn't working properly. They had failed to close the door to the MRI room. The anesthesiologist called out for oxygen. A nurse walking past the room heard the anesthesiologist call for oxygen and saw the tanks next to the room in the hall. Unwittingly, the nurse wheeled a tank into the room. The MRI machine was active, and the metal oxygen tank flew into the MRI, striking Michael and killing him (https://www.patient-safety-blog.com/2011/11/02/young-boy-killed-by-projectile-during-mri/).

Who is at fault? Yes, many things were in error. The subsequent analysis is on the patient safety blog page and lists that the door was not shut and locked. The MRI technicians were busy doing something else and were not available to stop the nurse; at least one technician should be in the room monitoring it at all times. The anesthesiologist called out that oxygen was needed rather than wait for the technicians to determine the problem and devise a solution. The nurse was unaware of the special procedures when administering oxygen in an MRI room. Finally, the oxygen tanks were stored next to the room and easily accessible for anyone to make this mistake.

Human error and accidents are not caused by solely one human but by a variety of factors inherent to the system. These can be **sociocultural factors**, design factors, environmental factors, and system weaknesses; they all contribute. There are several views of human error. One of the first is to examine the preconditions that may cause a human to behave in a less than optimal way. In engineering psychology, much of what we know about human error began with investigations of airplane pilot error. Approximately 80% of all airline accidents are due to human error, specifically pilot error, according to Li, Baker, Grabowski, and Rebok (2001). One of the first steps in correcting an error is in reporting; the Aviation Safety Reporting System keeps track of aviation incidents here: https://asrs.arc.nasa.gov/. Since reporting began in the 1970s, the system receives over 30,000 reports a year (Spencer, 2000).

We know that error is linked to stress. Stress can be created by errors existing in a system, or stress can be created when a system fails. Stress impacts human **information processing** in that it narrows **attention**, reduces **working memory** capacity, and increases the probability that people will **perseverate** (Wickens & Hollands, 2000). Perseverate is a person's tendency to follow a course of action that they have used in the past regardless of their previous success or failure. The tendency to perseverate when under stress can cause problems in system recovery, lengthen the system recovery process, and create further errors. In addition, perseverating negates the seeking of novel behaviors that may eliminate errors in the future, as the previous course of action may not have been successful (Wickens & Hollands, 2000).

THE YERKES-DODSON LAW

Yerkes and Dodson explained the link between stress and performance in their well-known Yerkes-Dodson law (1908). Briefly, the **Yerkes-Dodson law** states that for every individual and every task combination, there is a relationship that can be modeled with a stress/performance curve, as shown in Figure 23.1. In Figure 23.1, on the initial training side, during the ascending part, stress (black line) and the quality of performance (gray line) increase until there is a perfect balance between stress and performance. At the top, the performance is optimized. During the descending part, increased stress will decrease the quality of performance until the stress (black line) becomes overwhelming and a quality performance (blue line) is no longer possible. The Yerkes-Dodson law may be moderated by expertise and training as shown on the right side of Figure 23.1. Training enables the person to withstand additional stress with smaller performance decrements and flattens the Yerkes-Dodson curve. This principle is one of the reasons for "overtraining" humans in situations where human error could have disastrous effects. For example, first responders who are not actively attending an emergency are often engaged in training drills to refine their skills. These drills are similar to elementary school multiplication table drills; the goal is to overlearn the material. The training is so well learned, it overrides the tendency to perseverate and

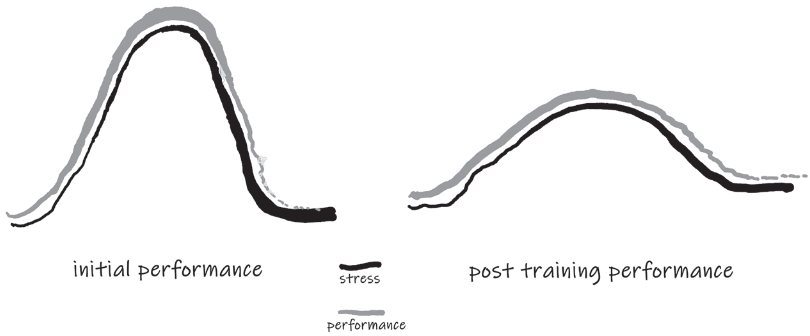

Figure 23.1. Yerkes-Dodson Performance Curve

becomes "second nature." This type of training can cause the Yerkes-Dodson curve to flatten so that stress has no effect on performance. This approach is used to address a disaster response, as seen in the Red Cross's "Make a Plan" on their website (https://www.redcross.org/get-help/how-to-prepare-for-emergencies/make-a-plan.html). This strategy is aimed at reducing feelings of helplessness during a disaster. When a natural disaster occurs, uninjured people will endure stress that is so overwhelming, they will fail to act and put themselves in danger. First responders recognize that the uninjured but "stressed-out" people are severely injured emotionally rather than physically due to stress. The Red Cross's strategy is to train the public in order to decrease stress and flatten the Yerkes-Dodson curve so that when a disaster does strike, there will be fewer severe emotionally injured people. This will increase the amount of time that first responders have to help the severely physically injured.

The study of human error is an entire research area within engineering psychology and human factors. In other words, researchers have devoted their professional lives to this single topic. There are several taxonomies or structures of why and how human error happens along with the types of human error. This chapter will cover a subset and will provide suggestions for additional reading. Some researchers classify human error as **errors of commission** and **errors of omission**. In commission-based errors, someone has made the wrong choice. In omission-based errors, someone has failed to do something. Within both omission- and commission-based errors, there are mistakes.

MISTAKES

Mistakes happen when the person misinterprets or misdiagnoses the state of the system or misunderstands what they are supposed to do in response; they may or may not follow an "if-then" rule. For example, you may have a new car and mistakenly turn on the blinkers instead of the windshield wipers. You did not understand where the windshield wipers were on the steering wheel column, as the car is new, and you turned the knob where your old windshield wipers were located, but you did follow the "if-then" rule: "if" the windshield is wet, "then" turn on the wipers.

Other mistakes can happen when you don't follow the "if-then" rule and you simultaneously misunderstand the state of the system. For example, you are driving an unfamiliar car and see that the fuel tank is nearly empty. You see a gas station but pass it by, as you think that you can make it to the next one. You run out of gas. You have misunderstood how far the car can travel when the gas tank reads empty and failed to follow the "if-then" rule of filling up the tank when you see that the car is on empty and a gas station is nearby.

SLIPS

Slips happen when someone understands but does not execute the right action. For example, in your car, you may see the stop sign and understand what it means but fail to stop. Slips can also happen when your attention is distracted during a familiar execution of an action. For example, you put a package of lunch meat on top of the refrigerator so that your dog doesn't grab it off the table. The phone rings, and you learn some upsetting news, so you forget to put the lunch meat away, and it spoils. In this case, you have the knowledge and know the rule—if the lunch meat is out of the refrigerator, you must put it back inside the refrigerator—but your attention has distracted you from executing the right action.

LAPSES

Lapses are also called forgetfulness. You know the rules and may or may not have the knowledge, but you simply forget due to your own lapse in memory. Lapses are most common when you must do things in a sequence and you forget a step. For example, I like to cook for my friends, but I am not good at following the steps in the cookbook. Often, I start the recipe and then get distracted by a joke that a friend was sharing. I forget where I am in the recipe, and instead of concentrating on the recipe, I dump the rest of the ingredients into the pot and finish cooking. It rarely turns out well, and I find myself ordering pizza and gaining a reputation as a terrible cook. My frequent lapses resulted in cooking failures. There are additional error categories and finer descriptions of these types of errors. With the purpose of understanding error research in general, we will use this simplified version of how they are categorized.

RELIABILITY ANALYSIS

If a system has a known reliability, the prediction of error is possible. For example, if it is known that the system functions perfectly for 90% of the operations performed, then 10% of the operations are prone to error. Given this function, if two systems must interact to

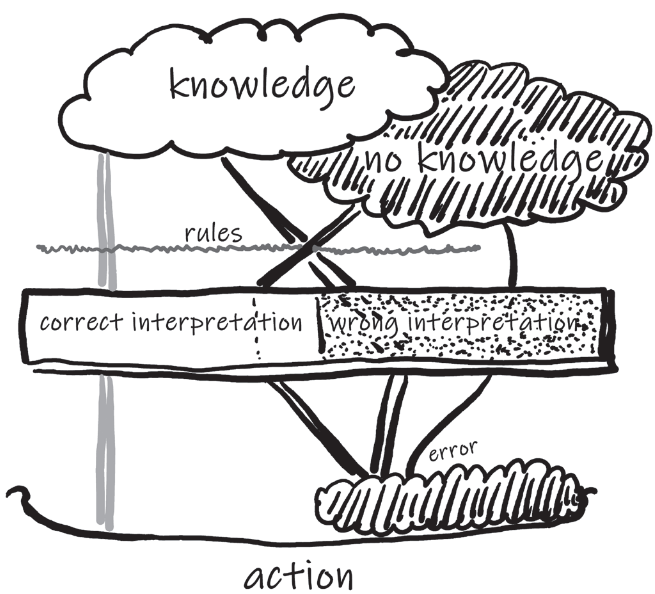

Figure 23.2. How Rules and Knowledge Contribute to Erroneous Actions

produce an operation and both systems operate perfectly 90% of the time, the probability that the two systems together will be successful is 0.90 × 0.90 = 0.81, or an 81% probability that all will be perfect. The probability of error in these two systems will be the inverse, or 1.0 – 0.81 = 0.19, or 19% (Wickens & Hollands, 2000). According to Swain (1990), a similar measurement can be calculated for human error, with the number of actual errors divided by the opportunity for human error.

The opportunity for human error can exist in the system, the system's environment, the organizational context, the procedures, the rules, the training, or the operators themselves. Reason (1990) would call these opportunities latent errors. Latent errors are often mediating factors that can take a human intention to produce the right action and apply the right knowledge into an unsafe act through no fault of the human. This observation prompted Reason to create the Swiss cheese model (SCM).

REASON'S SWISS CHEESE MODEL

The view that no one single human or factor created the error is described in the **Swiss cheese model** (**SCM**). The SCM is just as it sounds, based on the holes in Swiss cheese. Each opportunity for error is a hole. When the holes line up, disaster strikes. As in the medical example, pilot error can also be described using the SCM.

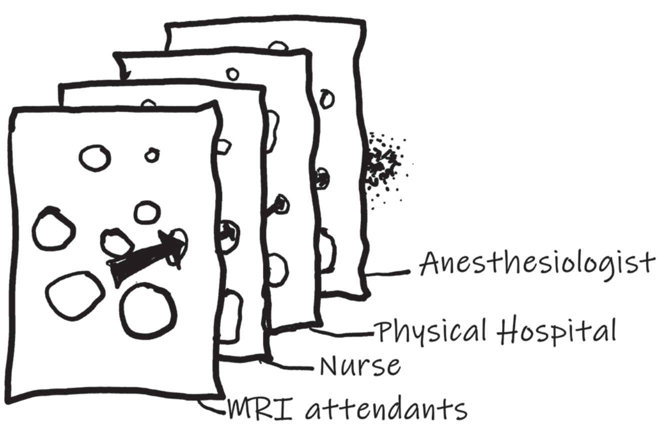

Figure 23.3. Swiss Cheese Model (SCM)

The SCM was developed when Reason needed additional material for a book (Reason et al., 2006). The Swiss cheese model describes previously unknown conditions that have the possibility of developing into disasters or accidents when they are combined. These are called latent conditions. Latent conditions are described by Reason (2000) as "They can translate into error-provoking conditions within the workplace (for example, time pressure, understaffing, inadequate equipment, fatigue, and inexperience), and they can create long-lasting holes or weaknesses in the defenses (untrustworthy alarms and indicators, unworkable procedures, design and construction deficiencies). Latent conditions—as the term suggests—may lie dormant within the system for many years before they combine with active failures and local triggers to create an accident opportunity" (p. 395).

There are other models that focus on specific aspects of error, such as root-cause analysis, Human Factors Analysis and Classification System (HFACS), and others. Reason (2000) states that one of the problems with investigations of errors lies in the perspective that the organization or investigator takes during the investigation. There are person-centered approaches and system-based approaches to error investigation. In the person-centered approach, it is assumed that the error was caused by a person's failure in either an unsafe act or a mental fault that resulted in the error. The mental fault could be "forgetfulness, inattention, poor motivation, carelessness, negligence, and recklessness. The associated countermeasures are directed mainly at reducing unwanted variability in human behavior" (p. 393). Proponents of this approach are likely supporters of the "just world" hypothesis and attribute the error to a moral failure. Supporters of a "just world" hypothesis often support punitive measures for the error and attempt to reduce error by either reducing the human component in the system ("we'll just automate the human out") or creating fear in the human component ("I have to be careful or I will be fired"). Either approach is problematic. In the "we'll just automate the human out" of the system approach, the automation creates additional opportunities for error and problems of its own. In the "I have to be careful or I will be fired" approach, the organization creates a stressful environment, which increases the likelihood of an error.

There are also system-centered approaches to error investigation. These approaches assume that humans are a fallible part of the system and that errors are to be expected as a consequence of "error traps" that are systematic in the system rather than a cause of the error (Reason, 2000). In this approach, the effort is toward changing the organization and the conditions under which the human interacts with the system. In other words, systems should have safeguards that prevent humans from operating the system in an unsafe manner. For example, it is extremely difficult to put a forward-moving car into reverse without making the car come to a complete stop. The automobile designers have put safeguards in the transmission that negate an unsafe act. The human has no choice but to engage the transmission properly.

As Reason (2000) states, "First, it is often the best people who make the worst mistakes—error is not the monopoly of an unfortunate few. Second, far from being random, mishaps tend to fall into recurrent patterns. The same set of circumstances can provoke similar errors, regardless of the people involved" (p. 394). Several models are based on Reason's Swiss cheese model (SCM), shown in Figure 23.3.

The **Human Factors Analysis and Classification System** (**HFACS**; Shappell & Weigmann, 2001) and **root-cause techniques** are applications of the SCM principles (Reason et al., 2006). In these models, the latent factors are failures of the system. While SCM has many proponents, it has many critics as well. Critics state that the **latent variables** (holes) are unclear, and they may change over time. Their interrelation is ill defined, and the method by which the holes align is unclear. The link between the latent variables and the error is available only in hindsight, suggesting that hindsight bias may play a part. In addition, these latent variables can only be identified if there is an accident. Finally, the SCM model is theoretical and has no experimental support (Reason et al., 2006). However, many researchers and practitioners find explanatory value in the SCM model.

Another approach to the modeling of human error is the **Functional Resonance Analysis Method** (**FRAM**; Hollnagel, 2012). This approach is qualitative and seeks to define how individuals' actions separately and together impact a complex system through modeling the categories of behaviors, functions, and structures that the individuals interact with and contribute toward. It begins with the persons involved stating their real and optimal behaviors when reaching a goal and focuses on how these vary throughout the scenario. The model describes where complexities happen, environmental factors, risky behaviors, optimal behaviors, and latent variables in the system (Bridges et al., 2018).

These and other models of human error focus on making systems safer by reducing error and the possibility of error in human-machine systems. In the coming chapters, we'll discuss how these and other models address problems within specific domains. The Human Factors and Ergonomics Society has additional resources and technical groups focused on error and safety. For more information, please see their website: https://www.hfes.org/resources/educational-and-professional-resources/educational-resources/new-item2.

REFERENCES

Bridges, K. E., Corballis, P. M., & Hollnagel, E. (2018). "Failure-to-identify" hunting incidents: A resilience engineering approach. *Human Factors, 60*(2), 141–159.

Hollnagel, E. (2012). *FRAM, the functional resonance analysis method: Modelling complex sociotechnical systems*. Ashgate.

Li, G., Baker, S. P., Grabowski, J. G., & Rebok, G. W. (2001). Factors associated with pilot error in aviation crashes. *Aviation, Space, and Environmental Medicine, 72*(1), 52–58.

Reason, J. (2000). Human error: Models and management. *BMJ, 320*(7237), 768–770.

Reason, J., Hollnagel, E., & Paries, J. (2006). Revisiting the Swiss cheese model of accidents. *Journal of Clinical Engineering, 27*(4), 110–115.

Shappell, S. A., & Wiegmann, D. A. (2001). Applying Reason: The human factors analysis and classification system (HFACS). *Human Factors and Aerospace Safety, 1*, 59–86.

Spencer, F. C. (2000). Human error in hospitals and industrial accidents: Current concepts. *Journal of the American College of Surgeons, 191*(4), 410–418.

Swain, A. D. (1990). Human reliability analysis: Need, status, trends and limitations. *Reliability Engineering & System Safety, 29*(3), 301–313.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance*. Prentice Hall.

Wiegmann, D. A., & Shappell, S. A. (2001). Human error perspectives in aviation. *International Journal of Aviation Psychology, 11*(4), 341–357.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18*(5), 459–482.

CHAPTER 24

ALARMS

GOALS OF THE CHAPTER

- Learn about the main research questions in alarms
- Understand how alarm types interact with human response

ASSESSMENT

- How do alarms trigger a "cry-wolf" effect?
- What is signal detection theory?
- What is redundancy gain?

MULTIPLE MODALITIES

Alarms are all around us. From the alarm on your car or house to the smoke alarm in your kitchen. When an alarm uses more than one sense to alert the operator, the alarm "speeds up processing." This is also called **redundancy gain**. For example, a smoke alarm uses both a loud beep and a flashing light. When the alarm senses smoke, it sounds, and the human responds. If the smoke alarm used two different senses, let's say smell and touch, it would not be as effective.

FALSE ALARMS

In the case of smoke alarms, there are many times when the alarm sounds and there is little danger of a serious fire. This is a **false**

alarm and is triggered when the alarm senses a change in the density of the surrounding air. As you may have experienced, smoke alarms emit false alarms for many reasons, from an overdone dinner to a low battery. In response, some people take the batteries out of the smoke alarm. This is called **alarm fatigue**. People have grown tired of the false alarms and rendered the alarm useless.

There are other cases of **alarm fatigue** where people are unable to disable the alarm and instead ignore it. This type of alarm fatigue can be more serious, as people will ignore hurricane warnings and tornado warnings, putting themselves and their families in danger. This contributes to the "cry wolf" phenomena, where humans either ignore the alarm or produce a delayed response (Wickens et al., 2009).

Detecting a true emergency and determining a real alarm from a fake or false alarm can be modeled by **signal detection theory**. Signal detection theory is a useful way of categorizing the types of signals from the noise and is used in many different disciplines. In the case of human behavior, we can model when the human thinks that the signal is present and when it is actually present, as shown in Figure 24.1.

This theory applies to all instances where a human must determine if something exists in the environment or not. It applies to alarms. When an alarm sounds, the human must determine if the alarm alerted them to a real failure or emergency. In order to determine the hit rate, the number of hits can be divided by the number of misses. In other words, if you find that your smoke alarm sounds whenever you cook, you might calculate the hit rate of the alarm and move it to a better location where it can be more accurate. You don't want

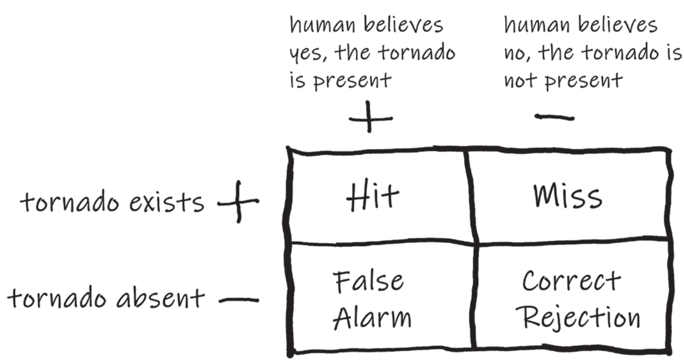

Figure 24.1. Signal Detection Theory

to learn to ignore the alarm over time. These calculations can also be used to determine the optimal threshold for the alarm, the sensitivity of the human in detection, and the relationship between the sensitivity of human detection and response bias. More details can be found in Wickens, Hollands, Banbury, and Parasuraman (2015) and in Green and Swets (1966).

Once an alarm is detected, the human must diagnose what response is needed. Humans may prioritize safety, economy, or fault correction (Wickens & Hollands, 2000). Often, these goals are congruent, but not always. When there is a major abnormal interruption in the system, preprogrammed procedures rarely anticipate abnormal events. This is because rules are written in the "if X then Y" fashion. When the rules are written in this way, both X and Y are assumed. Often humans will need to diagnose the source of the fault in order to address it and match it to X. Often, it does not match, as the instructions are unclear. Here is one example: https://www.youtube.com/watch?v=K_vZk1f353I.

Another diagnosis challenge occurs when multiple alarms signal progressing failure, as in a nuclear power plant. Here is an example from a re-creation of the Chernobyl nuclear disaster: https://www.youtube.com/watch?v=XFEXlqD0Us8. In the video, multiple alarms are signaling, and it is unclear to the human operators what is happening or what actions they need to take to mitigate the disaster. In this case, one failure prompted many others. The sound of several alarms for the same error compounds confusion and leads human operators to rediagnose, thereby wasting valuable time. If the operators' mental model of the system is not consistent or accurate, the multiple alarms can obfuscate the situation.

To address these situations, automation is incorporated with alarms. On a base alarm level, intelligent alarm systems compare current sensor data with past sensor data to diagnose the probability of a true failure/emergency. This lessens the likelihood of a false alarm and helps the human operator when diverting their attention from a specific task to mitigating the emergency. We like to think of visual alarms as binary—either off or on. In the case of intelligent alarm systems, they may display three or more levels of warning to communicate the probability of an emerging disaster. These can be visual, with green, yellow, and red being congruent with cultural norms, or they can be audible, with varying loudness or increasing sound pitch. Sorkin, Kantowitz, and Kantowitz (1988) tested to see if this approach to alarms interfered with a human's response to an impending failure. They found that there was no decrease in performance or increase in human response time. Intelligent alarm systems allow the human operator to take more time to diagnose the situation before it becomes dire.

REFERENCES

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley.

Sorkin, R. D., Kantowitz, B. H., & Kantowitz, S. C. (1988). Likelihood alarm displays. *Human Factors, 30*(4), 445–459.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance*. Prentice Hall.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology and human performance*. Prentice Hall.

Wickens, C. D., Rice, S., Keller, D., Hutchins, S., Hughes, J., & Clayton, K. (2009). False alerts in air traffic control conflict alerting system: Is there a "cry wolf" effect? *Human Factors, 51*(4), 446–462.

CHAPTER 25

ACCIDENTS AND INCIDENTS

GOALS OF THE CHAPTER

- Learn about the types of accident investigations
- Understand the role of humans in the organizational system

ASSESSMENT

- Is there a unified site where accident data can be viewed?
- What are the two types of accident investigation models?
- Which model refers to information processing in a human?
- What does the acronym HFACS stand for?

There are two categories of **deleterious events**: **accidents** and **incidents**. An incident is classified as an event with an unfavorable outcome. An accident is an event with a more severe outcome that can result in personal injury or property damage. The Occupational Safety and Health Administration (OSHA) has a handy guide to classifying either, which is here: http://www.jmcampbell.com/tip-of-the-month/wp-content/uploads/2012/10/figure-4.png. The main webpage it is taken from is here: http://www.jmcampbell.com/tip-of-the-month/2012/10/should-unplanned-maintenance-jobs-be-recorded-as-near-misses/.

Both incidents and accidents are serious unplanned events. However, once an accident happens, regardless of the cause or outcome, there is an investigation. In certain domains, such as nuclear power, aeronautics, and law enforcement, accident investigations are required by law. During accident investigations, personnel will collect eyewitness reports and operator reports, take measurements, and collect evidence, which will form a report. In some **domains**, such as aeronautics, there is a central location where the report is filed and available to the public (e.g., https://asrs.arc.nasa.gov/ and https://www.ntsb.gov/investigations/AccidentReports/Pages/aviation.aspx).

These investigations are rich sources of information that are later analyzed as additional tools are developed by human factors professionals. Anecdotally, Sully Sullenberger stated at a human factors conference that one of the reasons that he could land his plane on the Hudson River so successfully was because he had studied thousands of accidents that were similar to his (https://en.wikipedia.org/wiki/Chesley_Sullenberger). While accident reports can be difficult to construct and review, they are key to reducing future accidents.

According to Reinach (2006), there are additional models of how accidents occur: Edward's (1972) **software, hardware, environment, and liveware (SHEL) model**; Rasmussen's (1983) **skills, rules, and knowledge model**; Wickens and Flach's (1988) **four-stage model**; Reason's (1990) **generic error-modeling system** (**GEMS**); O'Hare's (2000) **"wheel of misfortune" taxonomy**; Shappell and Weigmann's (2000) **Human Factors Analysis and Classification System** (**HFACS**), and Moray's (2000) **sociotechnical model of error**. These models take one of two approaches. The first approach is the human-centric information-processing approach, which asks, "Why did the human make these choices?" (Rasmussen, 1983; Wickens & Flach, 1988). The second approach is the systems approach, which asks, "What system conditions were in place that allowed this error to happen?" (Edwards, 1972; Moray, 2000; O'Hare, 2000; Reason, 1990). Let's take a look at one example from each category: Rasmussen's (1983) skills, rules, and knowledge model and O'Hare's (2000) "wheel of misfortune" taxonomy. Then we'll end with a discussion of the currently most prominent model, Shappell and Weigmann's (2000) HFACS model.

SKILLS, RULES, KNOWLEDGE MODEL

Previous to Reason's work on the **Swiss cheese model** (**SCM**; 1990), which was discussed in Chapter 23, Rasmussen (1983) quantified human performance in terms of **skills, rules, and knowledge** and how these items interacted with a person's mental model of a system. A **mental model**, as mentioned in previous chapters, is a person's understanding of how something functions, as in Figure 25.2. The mental model is similar to a conceptual model. As an operator of a system moves toward something that they perceive in the system (i.e., an alarm) with the goal of resolving the system error, they move through a three-rung ladder. Some simple procedures, such as seeing a stopped car and pressing the brakes, exist only in the skill-based rung of the ladder. Other procedures, such as disarming a smoke alarm, may exist only in the rule-based rung of the three-rung ladder. The person may hear the alarm, recognize the sound, associate it with a smoke alarm, locate the smoke alarm, and take the battery out or open a window to lessen the smoke. Other procedures require more diagnosis, such as helping a student who is failing a course. The identification of the problem, the decision of what to do next, and planning the next action to help the grade are required. This would be the knowledge rung. An adaptation of Rasmussen's (1983) diagram describing the three rungs of the ladder is in Figure 25.1. In the diagram in Figure 25.1, signs refer to states of the environment, while symbols refer to internal reasoning

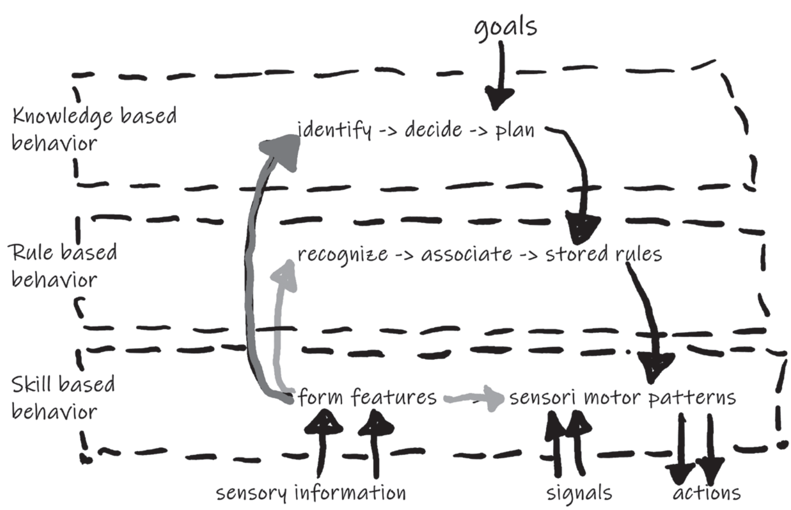

Figure 25.1. Skills, Rules, Knowledge Model
Source: Rasmussen (1983).

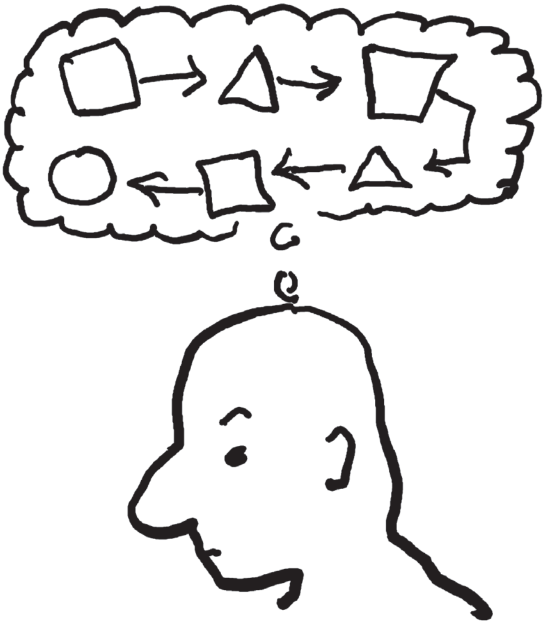

Figure 25.2. A Mental Model

between relationships and properties of the system. As the person organizes information, they may arrange the information into larger chunks (aggregate), they may think about the properties in terms of abstracting it to other categories (abstraction), or they may find structures in the representation that resemble something else (analogize). This qualitative model is a precursor to Reason's SCM (1983) model and many current SCM-based models that describe the conditions under which unsafe acts occur and the information-processing errors that happen as a result of a misperception or misunderstanding in the skills, rules, or knowledge application.

"WHEEL OF MISFORTUNE" TAXONOMY

O'Hare (2000) discusses a unique taxonomy of error in order to provide investigators with heuristic guidelines during an investigation. O'Hare's taxonomy uses Rasmussen's (1983) skills, rules, knowledge approach. There are a series of three concentric circles, with the center circle representing the frontline operators or personnel, the next circle out representing local conditions, and the outside circle representing global conditions. The central operators decide which level of control is appropriate for the task at hand: skill, rule, or knowledge. A well-practiced task may move from stimulus to response immediately and be termed a skill-based task. A less practiced task that has some unique characteristics may demand a rule-based approach. Finally, a unique task that is unanticipated would require a knowledge-based approach. A unique task that requires a knowledge-based approach with novice personnel is a likely error scenario. O'Hare equates Reason's (1990) SCM terminology and Rasmussen's levels of control and internal functions as creating a more comprehensive approach to classifying errors and the contributing factors.

HUMAN FACTORS ACCIDENT CLASSIFICATION SYSTEM (HFACS)

The Human Factors Accident Classification System (HFACS), developed by Shappell and Weigmann (2000), remains the most widely used model when investigating accidents and incidents. It is based on the SCM model by Reason (1990). HFACS categorizes conditions existing in the organization, system, and operator that contribute to accidents. The role of the accident investigator is to uncover the latent factors within four categories: (1) unsafe acts by the operator(s), (2) preconditions in the system (unsafe crew conditions), (3) organizational weaknesses (organizational influences), and (4) supervisory weaknesses (unsafe supervision). Within each of these categories are additional categories that quantify the latent variables in a system, organization, and human information processing. Shappell and Weigmann (2009) state that rarely is the human the sole cause of the system failure; the human is only one component in a complex interaction.

RESILIENCE

Safety systems are not something that is an inherent principle of a system once it is created and before an accident happens. Many system designers fail to realize that they have built-in weaknesses and that it is simply a matter of time before the system weaknesses produce an error (Dekker et al., 2008). Instead, it is easier to blame humans who are part of the system for the system's weakness. It is easiest to blame the human and implore the human to "not exploit the system" (https://www.mysafetysign.com/must-be-worn-on-this-site-safety-sign/sku-s2-2058#S2-2055). However, a stronger approach is to build systems that have few weaknesses. This is called resiliency engineering. Dekker has a presentation on this topic: https://www.youtube.com/watch?v=o3L_TQG-xBs. In the presentation, it is stated that viewing the human as the only part of the system that can anticipate and mitigate risk can create a safer system. The human considers the balance between safety and risk to create a safe outcome. During the activity, the human may consider past experiences and incorporate additional measures to increase safety. Dekker also mentions in the video that organizations that incorporate a resilience engineering approach follow four guidelines. These are (1) they do not assume that the procedure will be safe, (2) they keep the description of possible risks ongoing throughout the procedure, (3) they invite criticism and explore it, (4) there is someone who has the authority to stop a procedure if the risks outweigh the probability of a safe outcome.

There are strengths and weaknesses to each of these models. The information-processing approach of the skills, rules, knowledge model suggests that remediation with training or incentives can address the issues of safety. The systems-based approaches of SCM (Reason, 1990), HFACS (Shappell & Weigmann, 2000), and "wheel of misfortune" taxonomy (O'Hare, 2000) suggest that resilient systems are a better approach. Ultimately, the use of a repository of information such as the Aviation Safety Reporting System (ASRS) should remain public (https://asrs.arc.nasa.gov/). Future engineering psychologists and human factors practitioners are developing emerging models of error and prevention.

REFERENCES

Dekker, S., Hollnagel, E., Woods, D., & Cook, R. (2008). *Resilience engineering: New directions for measuring and maintaining safety in complex* *systems*. Lund University School of Aviation. https://www.researchgate.net/profile/Erik_Hollnagel/publication/238687807_Resilience_Engineering_New_directions_for_measuring_and_maintaining_safety_in_complex_systems/links/0046353148e497e765000000.pdf

Edwards, E. (1972). Man and machine: Systems for safety. In *Proceedings of British Airline Pilots Association Technical Symposium*. British Airline Pilots Association.

Hollnagel, E., & Speziali, J. (2008). *Study on developments in accident investigation methods: A survey of the "state-of-the-art"* (SKI Report 2008:50). Ecole des Mines de Paris.

Moray, N. (2000). Culture, politics and ergonomics. *Ergonomics, 43*, 858–868.

O'Hare, D. (2000). The "wheel of misfortune": A taxonomic approach to human factors in accident investigation and analysis in aviation and other complex systems. *Ergonomics, 43*(12), 2001–2019.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-13*(3), 257–266. https://doi.org/10.1109/TSMC.1983.6313160.

Reason, J. (1990). *Human error*. Cambridge University Press.

Reinach, S., & Viale, A. (2006). Application of a human error framework to conduct train accident/incident investigations. *Accident Analysis & Prevention, 38*(2), 396–406.

Shappell, S. A., & Weigmann, D. A. (2000). *The Human Factors Analysis and Classification System—HFACS* (Report #DOT/FAA/AM-00/7). Federal Aviation Administration Report.

Shappell, S., & Wiegmann, D. (2009). A methodology for assessing safety programs targeting human error in aviation. *International Journal of Aviation Psychology, 19*(3), 252–269.

Weigmann, D. A., & Shappell, S. A. (1997). Human factors analysis of post-accident data: Applying theoretical taxonomies of human error. *International Journal of Aviation Psychology, 7*(1), 67–81.

Wickens, C. D., & Flach, J. M. (1988). Information processing. In *Human factors in aviation* (pp. 111–155). Academic Press.

CHAPTER 26

AERONAUTICS

GOALS OF THE CHAPTER

- Learn about the different domains within aeronautic study
- Understand the underlying problems in each area

ASSESSMENT

- What initial approach did human factors researchers use to control pilot error during World War I and World War II?
- One model of categorizing error is through a systematic-failure taxonomy. What is the other model of categorizing error?
- What is the primary cause of UAS accidents?
- How many feet above ground level is considered to be "space," as in "space flight"?

There are at least four areas of concentrated aeronautical studies that examine safety and error. These are **air traffic management** (**ATM**), aerial vehicles with a human pilot (planes), **unmanned aerial systems** (**UAS**), and space flight. The opportunities for failure are different in each. For example, in ATM, the challenge is balancing human perception and **information processing** with a **high workload and time pressure.** For each area, the investigation methods are designed to address these different key challenges.

AIR TRAFFIC MANAGEMENT (ATM)

In Chapter 1, we discussed that some human factors psychologists work in aviation psychology with pilots. After the focus on selection and training in airplane pilots for World Wars I and II, the focus in aviation psychology shifted from piloting airplanes to managing airplane traffic in the national airspace through ATM. ATM encompasses air traffic control (ATC) as well as other systems to manage the air traffic. Matheny developed the first **ATM simulator** at Ohio State University (Koonce, 1984). This simulator was key in modeling the effects of **working memory**, as air traffic controllers managed to keep track of several "planes" and "flight tracks" at once. This led to recommendations for the optimal lighting and on working memory aids based on Miller's work (Miller, 1956).

While the human error rate in ATM is high (90%), the contribution of ATM to plane accidents is relatively low, according to Isaac, Shorrock, and Kirwan (2002). Most ATM tasks are information-processing tasks, such as **judgment**, **situation awareness**, planning, and communication. The European Union member states addressed the challenge of **human error** in ATM through the **Human Error in Air Traffic Management (HERA)** project. In the HERA project, they compared **models of human performance** and **human-error taxonomies** in order to map the right combination to its **operational context**. Their final model of ATM errors settled on a **mental-model update process**, where the ATM operator's flow of information along with the information quality is constantly being updated by the ATM operator while the ATM operator engages in judgment (i.e., "Are these two planes going to collide?"), planning (i.e., "Who should land next?"), and decision-making (i.e., "Which runway would be the best?"). The mental-model update process is shaped by **performance-shaping factors**, such as the amount of air traffic, the ATM procedures, ATM systems that the controller must use, social factors between the ATM operators, team factors, and internal personal factors. Then they identify key ATM tasks and classify the information needed to perform those tasks and then the types of errors and a classification system of where these errors happen in the information-processing system. Finally, the team identifies several systems of classification, or **taxonomies**, to use in a systematic process to identify the factors contributing to error. This combination of taxonomies was named **TRACEr** (Isaac et al., 2002). In order to address some of the information-processing challenges with human controllers, many ATM systems are moving toward automation that predicts where the plane will go along with the optimal takeoff and landing scenarios. The aim is to decrease error, but it may result in a loss of **situation awareness** in controllers. According to Chiappe, Vu, and Strybel (2012), preliminary studies have suggested that if the **automation** is **adaptive** and if the automation allows the human controllers to override the suggested course of action, then it can be beneficial in reducing **workload**, allowing the controllers to maintain **situation awareness**, and decreasing error.

AIRPLANES

Aviation psychology began with a study of the best characteristics for an airplane pilot and the best way to design the controls in the cockpit (Koonce, 1984). As aviation psychology grew, the issue of communication between pilots and between ATM and pilots became a focus (Koonce, 1984). One of the latest issues is the focus on error and how to reduce crashes. According to Li, Baker, Grabowski, and Rebok (2001), pilot error contributing to crashes in major airlines has decreased 21% between 1983 and 1996. Some of the reasons for this decrease is the increase in technology assisting the pilot with information processing, situation awareness, and workload along with intervention programs such as **crew resource management** (**CRM**). However, many errors remain undetected and below the threshold of causing a crash. A crash would be a necessary condition for an investigation, and the error would be discovered. In a crash, there would be a review of the black box recordings of all pilot actions during flight. To investigate the effectiveness of the increase in technology and the impact of CRM, researchers now use flight simulators to study pilot error and create better models of error.

Aviation crashes can be viewed as systematic failures through the **Swiss cheese model** (**SCM**) (Reason, 1990) or **Human Factors Analysis and Classification System** (**HFACS**; Shappell & Weigmann, 2000) or as operator information-processing error through the "wheel of misfortune" taxonomy (O'Hare, 2000) or **skills, rules, knowledge**

model (Rasmussen, 1983). When viewed as systematic failures, the **latent variables** in the design of the plane, the design of the airport, the communication and design of ATM, the design of passenger management systems, and the weather must all be considered as layers with latent opportunities for error. When considering the humans involved in the process, the issue is simplified as one of communication from, to, and within the humans, as described here: https://flightsafety.org/asw-article/failure-to-communicate/. This encompasses the communication of the system to the humans and of humans to each other. Better communication contributes to better **mental model updates**, increased **situation awareness**, shared **situation awareness**, better judgments, and better decisions/actions. If we know that communication has such an impact, then why do people not communicate? It takes effort, increases workload, and increases stress, and there are **sociocultural barriers** that impede communication. While efforts have been made to standardize communication and normalize when and what to communicate, when unique situations arise, it is not clear what to communicate to whom and when. While the classification of errors and automation of the system will increase efficiency, human communication is the next frontier in averting aviation crashes.

UNMANNED AERIAL SYSTEMS (UAS)

There are many different types of UASs that are classified based on their **maximum take-off weight** (**MTOW**)—or in other words, the most that they can weigh (with cargo) in order to successfully take off and land: https://en.wikipedia.org/wiki/U.S._military_UAS_groups. As the MTOW increases, the automation and the number of operators required to fly it increase as well. For example, in Groups 4 and 5 with an MTOW of over 1,000 lbs., the systems are highly automated and require at least three operators to fly. In Group 1 with an MTOW of less than 20 lbs., the systems are lower on the automation taxonomy. Several UASs can be flown by a single operator. With the exception of **loss-of-signal errors**, the errors produced are different in different groups. In Groups 4 and 5, the errors resemble those of airplane errors, with communication as a primary root cause of many accidents. In Groups 1, 2, and sometimes 3, weather/wind and system failure/unpredictability are the primary root causes of many accidents. All groups must interact with ATM. ATM can be unfavorable to UAS. These aircraft have a reputation for unpredictability and causing disruption in the air traffic flow. When permission is granted for a UAS to operate in the national airspace, the ATM will issue a **notice to airmen** (**NOTAM**). The NOTAM will advise pilots to give the UAS a large space to operate: https://en.wikipedia.org/wiki/Airspace. In addition, the ATM will authorize UASs to operate only in sparsely populated airspace and within a particular class of airspace or a specific area denoted in feet above ground level.

Loss of Signal

Loss of signal is the primary cause of UAS crashes and a top concern of the ATM. Loss of signal occurs when the UAS loses radio contact with the **ground control station**. In Groups 1 and 2, a wind over 5 mph can easily take the UAS out of range of the ground control module. When loss of signal occurs, no one has control of the UAS. Most UAS manufacturers program the system's automation to either **loiter in place** or return to the take-off point (or home) when loss of signal occurs. Loiter in place means that the UAS will remain where the signal loss occurred and fly in a set pattern over the last known signal location. Sometimes the operator knows where the signal loss occurs and sometimes the operator does not. When signal loss occurs, it is similar to talking to a friend when the cell phone call drops. You might talk for a bit before you realize that your friend isn't there. A second problem happens when loss of signal occurs and the wind takes the UAS several miles off course. Then, if the UAS is programmed to **loiter in place**, finding the UAS can become a challenge. If the UAS is programmed to return home, finding the UAS is easier, but home may be several miles from where the UAS experienced the loss of signal. In that time, the UAS may experience a loss of power from battery depletion or incur damage, causing it to crash. During key UAS operations, the airline regulators may require that the smaller Groups 1 and 2 UASs have chase planes so they can be found when loss of signal occurs.

The second cause of loss of signal is **flying beyond the box**. A UAS flies within an imaginary signal box, as shown in Figure 26.1. With rapid acceleration, many UASs can accelerate rapidly and exceed the ground control station's transmission. This is referred to as "flying

Figure 26.1. The Box

beyond the box." When this happens, the UAS loses contact with the ground control station and the UAS operator. In Groups 1 and 2, this is a human operator challenge to control the UAS speed. In Groups 3 and greater, the human navigator can set coordinates for the UAS to fly toward. These coordinates or waypoints are carefully mapped to the mission and within the control area. This can be a human navigator challenge to manage the waypoints properly. Groups 1 and 2 UAS also may use waypoint navigation and possibly experience a similar challenge with novice operators.

SPACE FLIGHT

Flight beyond what is normally agreed upon as airspace is called space flight. In most countries, this is agreed to be approximately 62 miles vertically above ground level (https://en.wikipedia.org/wiki/Airspace); however, there are few defined rules of what does and does not constitute space flight. One of the reasons is clearly that our capability to fly more than 62 miles above ground level is limited by technology. As our capability increases, concerns about error and safety for humans increase as well. White and Averner (2001) suggest that the psychological and social challenges of confinement for a long period of time, the reduction in body weight, and radiation will be the top challenges to address outside of spacecraft operation. In the current pandemic climate, many people are experiencing the psychological and social effects similar to a long spaceflight, as they are isolated in their homes and fearful of contracting a virus if they leave their homes. White and Averner (2001) state that spending an extended amount of time with the same people poses challenges to the psychological and motivational effectiveness of a space crew and passengers. Many people reported that during the pandemic, they were experiencing a disruption in sleep and appetite along with interpersonal problems compounding preexisting internal propensities to anxiety and depression. To address these problems during space flight, measures of performance, stress, emotional distress, boredom, anxiety, and depression will be key.

REFERENCES

Chiappe, D., Vu, K. P. L., & Strybel, T. (2012). Situation awareness in the NextGen air traffic management system. *International Journal of Human-Computer Interaction, 28*(2), 140–151.

Isaac, A., Shorrock, S. T., & Kirwan, B. (2002). Human error in European air traffic management: The HERA project. *Reliability Engineering & System Safety, 75*(2), 257–272.

Koonce, J. M. (1984). A brief history of aviation psychology. *Human Factors, 26*(5), 499–508.

Li, G., Baker, S. P., Grabowski, J. G., & Rebok, G. W. (2001). Factors associated with pilot error in aviation crashes. *Aviation, Space, and Environmental Medicine, 72*(1), 52–58.

Miller, G. A. (1956). The magic number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 91–97.

O'Hare, D. (2000). The "wheel of misfortune": A taxonomic approach to human factors in accident investigation and analysis in aviation and other complex systems. *Ergonomics, 43*(12), 2001–2019.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions* *on Systems, Man, and Cybernetics, SMC-13*(3), 257–266. https://doi.org/10.1109/TSMC.1983.6313160.

Reason, J. (1990). *Human error*. Cambridge University Press.

Shappell, S. A., & Weigmann, D. A. (2000). *The Human Factors Analysis and Classification System—HFACS* (Report #DOT/FAA/AM-00/7). Federal Aviation Administration Report.

White, R. J., & Averner, M. (2001). Humans in space. *Nature, 409*(6823), 1115–1118.

CHAPTER 27

MEDICINE

GOALS OF THE CHAPTER

- Learn more about the human factors research areas in medicine
- Understand how the principles of alarms apply to medicine
- Understand the dangers of health-based applications

ASSESSMENT

- What do the terms EMR and EHR refer to?
- What are the three types of checklists?
- What is the first step in gauging situation awareness according to these researchers?
- Why are medical professionals concerned about patients' use of apps?

ELECTRONIC MEDICAL RECORDS

According to Moacdieh, Ganje, and Sarter (2017), most of the hospitals in the United States have adopted information-processing systems to gather, organize, store, and share patient data. These systems are called electronic health records or **electronic medical records** (**EMRs**). EMRs have been a great advancement in patient care, as patients are no longer required to get their medical records from the physician and then transfer them to the new physician. The records are already in place and show the etiology of the current health conditions based on past visits. Medical personnel have mixed opinions on their usefulness, with many physicians and pharmacists complaining about poor interface design and automated behavior. Moacdieh, Ganje, and Sarter (2017) investigated poor interface design in terms of visual clutter. In a survey of medical residents, they found that visual clutter in terms of poor placement/design and extraneous information about a patient was distracting enough to contribute to medical errors. In an unpublished study done with a medical center, Elliott (2021) found the EMR used in the pharmacy was producing many false alarms. The false alarms were creating alarm fatigue in the pharmacists and attending physicians. The concern was that the alarm fatigue was prompting physicians and pharmacists to overlook potentially harmful drug interactions. The medical center addressed the problem with an updated EMR.

Alarm fatigue is not only an EMR problem; it is a problem throughout the medical community (Sendelbach & Funk, 2013). Sendelbach and Funk estimate that between 72%–99% of medical alarms are false. The alarms come from various sources, such as intercom alerts, patient call buttons, and informational announcements. The primary source of the alerts is medical devices. According to Sendelbach and Funk, one hospital recorded about 350 alarms per patient per day. Devices will alert medical personnel when the patient has run out of intravenous medication, there is a device malfunction or inconsistency, and the patient has had a significant change in status. Sendelbach and Funk note that nurses grow accustomed to **false alarms** and are more likely to conclude an alarm to be false before checking. A single patient could be monitored by six different medical devices, all emitting different-sounding alarms at different intensities at different points with little opportunity to differentiate which alarms indicated a critical need and which could be ignored for a short period of time. Sendelbach and Funk state, “Of the 33 different sounds, they correctly identified only 50% of the critical alarm signals and 40% of the noncritical alarm signals” (2013, p. 379). The website Boston.com writes that patients have died as a result of alarm fatigue (http://archive.boston.com/news/local/massachusetts/articles/2010/04/03/alarm_fatigue_linked_to_heart_patients_death_at_mass_general/?page=1).

Morano (2014) states that medical alarm fatigue has resulted in at least 80 deaths between 2009 and 2012. Alarms fall into four categories: crisis, warning, advisory, and message. Most medical device false alarms are due to placement of the sensors, problematic algorithms, and patient movement. Both Morano and Sendelbach and Funk recommend having personnel whose only role is to monitor the alarms of several patients and alert the proper personnel to the most important issues. While many medical organizations have introduced systems that monitor the alerts, these systems seem to have compounded the issue.

HOSPITAL PRACTICE

Concerns over alarms and the quality of health care increase when a patient enters a hospital. This is because the number of health-care professionals and medical devices used to address the health problem exponentially increases the number of potential adverse events. According to Subbe and Barach (2019) and colleagues, up to 10% of patients suffer unwanted consequences and complications due to surgery. These events are termed as "quality-of-care" issues. Up to 40% of all health spending is to address the "quality-of-care" issues. Hospitals and physicians are aware of these challenges and are working to address them.

Human factors professionals have defined at least two of the factors leading to most errors in patient care. These are communication between medical professionals and **shared team cognition / situation awareness**. Within the communication realm, patients should feel comfortable talking to their physicians and nurses, and the physicians and nurses should feel comfortable talking to each other and sharing concerns. Prior research in surgery suggested that the surgeon held a position of authority on the team and discouraged others from sharing concerns. If the surgeon would fail to notice something, other team members would be afraid to point this out. Eventually, the number of errors due to this lack of communication led to the adoption of checklists before the patient enters surgery, during surgery, and during the postsurgical procedure. There are checklists for surgery team members who know each other well and for those who are meeting for the first time. These checklists include some additional unusual procedures, such as addressing each team member by their first name and marking the part of the patient's body receiving the procedure with a black permanent marker.

Subbe and Barach (2019) describe the adoption of a crisis emergency checklist to address challenges for common emergency situations. There are three types of checklists used: normal checklists, nonnormal checklists, and emergency checklists. As it would seem, normal checklists are for standard operating procedures, nonnormal are for nonstandard procedures, and emergency checklists are for unusual and crisis situations. Subbe and Barach (2019) used the templates provided by Ariadne Labs, where additional resources are available (https://www.ariadnelabs.org/), to develop their checklist. While checklists have improved communication and team cognition / situation awareness, they have not eliminated all problems. According to Subbe and Barach (2019), some medical personnel worry that the use of checklists will limit their autonomy and judgment or change professional relationships. However, there is little evidence to support the viewpoint.

SITUATION AWARENESS

Much of the human factors work in aviation has translated well to medicine. One of these areas is **situation awareness**. Situation awareness, as discussed in previous chapters, applies to medicine in the same way. The medicine domain has much in common with the aviation domain, as both are complex and dynamic and have intense and vast information needs, a changing workload, and considerable risk if all does not go as planned (Gaba et al., 1995). Situation awareness is the awareness of what has happened, what is presently happening, and then what may happen in the future. Situation awareness relies on cues, evolving situations, and expertise (Gaba et al., 1995). One of the first steps in gauging situation awareness is to define the tasks through an analysis, measuring the workload, and then measuring the spare capacity to process data through a secondary task (Gaba et al., 1995). One of the key questions Gaba and colleagues addressed was if a training module would help medical professionals with situation awareness challenges. They suggest that indeed situation awareness contributes to information processing in medicine, shared mental models, and shared cognition.

WORKPLACE INJURIES

According to Thompson, Stock, and Banuelas (2017), nurses have the second-highest nonfatal workplace injury rate. This is due to a combination of intense physical demands, stress, and **shift work**. Communicating these safety hazards to nurses through workplace training, addressing shift work through prohibiting back-to-back shifts, and giving nurses tools to alleviate stress are measures that human factors practitioners have recommended to address the issues. Understanding these factors and how they impact nursing helps increase functional capacity.

PATIENT-CENTERED APPS

Patient-centered apps on smartphones have increased exponentially, along with fitness apps. While many patients embrace the apps to monitor blood pressure, diabetes, or advice on children's milestones, few apps are developed with a medical professional on the team (Lazzara et al., 2017). This leads to insufficient or poorly constructed advice for the consumer. Many medical professionals are concerned about users' perception of serious illness and the use of apps to replace professional medical advice. It is unclear what the effect of health apps will be in the future.

REFERENCES

Elliott, L. J. (2021). *The effect of false alarms in pharmacy applications*. [Manuscript in preparation]. Department of Psychology, Pennsylvania State University.

Gaba, D. M., Howard, S. K., & Small, S. D. (1995). Situation awareness in anesthesiology. *Human Factors, 37*(1), 20–31.

Lazzara, E. H., Baker, A. L., Abts, N., Nathan-Roberts, D., Ranton, J., Fouquet, S., & Chaparro, B. (2017, September). Going mobile: Guiding the development of safer and more effective mobile access in healthcare. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 578–582). SAGE.

Moacdieh, N. M., Ganje, T., & Sarter, N. (2017, September). Physician perceptions of clutter in electronic medical records. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 572–572). SAGE.

Morano, J. (2014). *The growing concern surrounding medical "alarm fatigue."* ABM Healthcare Support Services. Retrieved June 3, 2020, from http://

www.abm.com/WhyABM/white-papers/Documents/Healthcare_AlarmMgmt_Whitepaper.pdf

Sendelbach, S., & Funk, M. (2013). Alarm fatigue: A patient safety concern. *AACN Advanced Critical Care, 24*(4), 378–386.

Subbe, C. P., & Barach, P. (2019, September). Crisis checklists in hospital: Translating from complicated to complex care environments. In *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care* (pp. 157–160). SAGE.

Thompson, B. J., Stock, M. S., & Banuelas, V. K. (2017). Effects of accumulating work shifts on performance-based fatigue using multiple strength measurements in day and night shift nurses and aides. *Human Factors, 59*(3), 346–356.

CHAPTER 28

FIRST RESPONDERS

GOALS OF THE CHAPTER

- Learn about the types of human factors safety challenges in first responder professions
- Understand the types of tools developed to explore these challenges
- Understand the impact and consequences of safety challenges in first responder professions

ASSESSMENT

- What types of things does a police officer carry on their equipment belt?
- What is one thing that a police officer can do to detect a lie?
- Why are firefighters at risk for slips and falls?
- What is an example of a question that you might ask an EMT that would indicate their understanding of team structure and leadership?

Law enforcement officers, firefighters, and emergency medical technicians are referred to collectively as **first responders**. They are the people who are first on the scene of an accident or emergency. The public relies on them to make decisions and avoid errors. As in aviation, the human factors problems vary based on the nature of the job, the environment, and the training factors. For example, law enforcement officers typically have an undergraduate degree and then are trained in an academy for at least 9 months. They then go into a mentorship program, where they are paired with a training officer for an extended period of time. Many firefighters (who may be volunteers or paid) earn a fire science degree before attending a fire academy for 600 hours or more. Volunteer firefighters work in small communities as part-time firefighters and have the same duties as career/paid firefighters. The volunteer firefighters may have other full-time jobs, such as rancher or farmer. Many firefighters are also emergency medical technicians (EMTs). They have additional medical training and certification with state-based licensure and exams.

Different first responders experience different levels of predictability in their day-to-day work. In the case of firefighters, their task is fairly predictable. Fire behavior is well known. Firefighters may rely on external cues and the environment to select their next course of action. In the case of law enforcement and EMTs, their task is unpredictable; human behavior or medical problems change rapidly and are less well controlled. The signals of what to do next, the environment, and the social-cultural climate may change quickly. Both law enforcement officers and EMTs must adapt immediately. This may require a different course of action. In the case of law enforcement, people lie to law enforcement officers.

Lies become a danger to both the law enforcement officer and the person who is lying. A failure to represent the situation truthfully impairs the officer from doing their job properly. Detecting lies is critical in law enforcement. Yet we are still discovering more about how people work in an environment where lying is normal. Driskell, Salas, and Driskell (2012) asked pairs of law enforcement officers and firefighters who had previously worked together to tell each other stories that were either true or false. Then they recorded their stories and reactions. They measured synchrony in terms of time spent in mutual gaze, speech transitions, and frequency/type of word usage. They found that truthful exchanges had more mutual gaze, more transition words, first-person-plural pronouns, negations, social processes, and words that indicated tentativeness, certainty and, inhibition. Truthful pairs asked more questions and used language indicating agreement.

Lying demands **cognitive resources**, as the liar must monitor their behavior and the lie so that it remains plausible and undetected. Vrij, Granhag, Mann, and Leal (2011) suggest that law enforcement officers demand that the interviewee maintain eye contact and ask the interviewee to tell the story in reverse order. Vrij and colleagues suggest that the reverse-order technique disrupts the reconstruction of the event from a schema or standard paradigm. This technique makes the subtle cues in the lies more noticeable and makes it more difficult for the person lying to maintain the plausibility in a story. When the reverse-order technique was employed in their study, 42% of the lies were identified. However, the assumption of lying as a normative behavior adds stress to a law enforcement officer's job.

STRESS

Stress impacts working memory, which impacts information processing. As stress increases, working memory decreases, taking a portion of the available attention pool and hijacking it as the officer determines if the person is lying or not. With decreased working memory, law enforcement officers rely on well-known routines and standard operating protocols. However, these are imperfect, as nearly every situation is unique. In Elliott (2011), law enforcement officers and firefighters were compared on their ability to select their next course of action in a scenario constructed by their senior personnel. Their next **course of action** was based on their ability to determine which indicators in the story were key to determining the correct course of action. Then signal detection theory was used to determine which group was more accurate. After reviewing the results, both parties noted that the firefighters used heuristic rules to determine the next course of action based on cues. In other words, once the fire does this, then we do this. Law enforcement officers noted that while they have heuristic rules or standard operating procedures, humans are unpredictable. A single cue may lead to multiple correct courses of action. This revealed the stress and uncertainty involved in law enforcement.

This uncertainty, along with time pressure and severe consequences for a mistake, can lead to anxiety, stress, and information-processing decrements. Baumann, Gohm, and Bonner (2011) found that exposing firefighters to extreme training simulations repeatedly lessened the effect of stress, uncertainty, and pressure. However, the repeated training in one simulation did not **transfer** well to a slightly different simulation. Firefighters who participated showed an increase in anxiety and **cognitive load** when the simulation changed, even though it was similar.

ERGONOMICS

Both law enforcement officers and firefighting personnel use heavy equipment. For example, in law enforcement, officers must wear an equipment belt. This belt must carry a pistol, extra ammunition, a large flashlight, handcuffs, pepper spray, and sometimes a baton or radio (https://www.americancityandcounty.com/2010/02/05/ergonomics-and-police-duty-belts-easing-their-load/, https://www.ncjrs.gov/pdffiles1/nij/grants/229710.pdf, and https://missionlocal.org/2017/10/heavy-gun-belts-are-giving-sfpd-officers-back-problems/). Low back pain is reported in 43% of law enforcement officers, with some officers' gait changing due to the weight of their equipment belts (Ramstrand et al., 2016).

Officers access a laptop or **mobile data terminal** (**MDT**) when they are in their vehicles. The MDT allows nearly real-time access to databases listing previous offenses, license plate information, and applications to file their reports. The MDT itself causes distractions, with 37% of law enforcement accidents attributed to distracted driving (https://www.policeone.com/police-products/vehicle-equipment/mounts/articles/is-your-patrol-in-car-setup-cluttered-distracting-and-possibly-preventing-your-airbags-from-saving-your-life-z5c4hD9E14YzvOEJ/). The position of the laptop requires that officers twist in their seat, which may increase issues when wearing an equipment belt (https://blog.thebackschool.net/uncategorized/police-officer-ergonomics-pt-1/ and https://crossfire.com.au/ergonomics-and-police-duty-belts-easing-the-load/).

A similar issue happens in firefighting, as firefighters must wear breathing equipment when they enter burning structures. The equipment can hamper movement. The heavy smoke in burning structures impairs vision. For these two reasons, firefighters are at risk for slips and falls as they navigate unfamiliar buildings. EMS personnel are at risk for back injuries when lifting heavy patients and repetitive-motion injuries (https://www.work-fit.com/blog/5-ergonomic-best-practices-for-emergency-first-responders).

While some human factors professionals have advocated for safer equipment, manufacturers are challenged to produce safer equipment, as departments replace equipment infrequently. Newer equipment that meets enhanced ergonomic guidelines is becoming available slowly (https://www.policeone.com/police-products/duty-gear/articles/dragon-skin-ergonomic-duty-belt-on-display-at-shot-show-2019-ncoGtiD0QH8xvfXV/). While the Occupational Safety and Health Administration (OSHA) implemented guidelines in the 1970s to protect workers from injury in their workplace, injuries inflicted over long-term use are difficult to attribute to a poor design. The redesign of law enforcement vehicles has been suggested, but departments lack funds to update vehicles on a regular basis (Koder, 2005). While the work on how equipment should be improved and the impact of the current equipment has been documented in law enforcement and firefighting, until the public provides the means to update equipment to the latest standards, officers will make do with what is available.

EMERGENCY MEDICAL TECHNICIANS

In EMT work, **teaming** creates rapid emergency resolution. As Patterson and colleagues (2012, p. 98) note, the factors of "1) team orientation, 2) team structure and leadership, 3) partner communication, team support, and monitoring, 4) partner trust and shared mental models, 5) partner adaptability and back-up behavior, 6) process conflict, 7) strong task conflict, 8) mild task conflict, and 9) interpersonal conflict" contribute to a new measure of teamwork in EMTs. Since the development of this model, other researchers have used it to explore team member familiarity and relationships to injury, communication, and outcomes. Hughes, Patterson, Weaver, Gregory, and colleagues (2017) found that a positive perception of teamwork was prevalent. When administering the survey, they found that "5) partner adaptability and back-up behavior" was the most prevalent factor when the teams knew each other well.

Other factors in EMT work include fatigue, sleep, and alertness. It is common for EMTs to work long shifts or double shifts, as there are typically only a few EMTs per station. In addition, EMTs may have particular attitudes toward fatigue and alertness that make it difficult to work as a good team member. Patterson, Buysse, Weaver, Suffoletto, McManigle, Callaway, and Yealy (2014) created a survey tool to help identify poor attitudes toward fatigue and alertness. In the future, this tool can be used to identify attitudes that would adversely impact the alertness of the team members and the performance of the team.

As in aerospace crews, medical crews such as EMTs use crew resource management (CRM). Hagemann, Kluge, and Kehren (2015) point out that first responders are also **high responsibility teams** (**HRTs**) because of unpredictable working conditions where the consequences for a mistake are potentially life changing. CRM training was provided to a group of emergency doctors, which are similar to EMTs in the United States. Hagemann and colleagues found that the training increased the doctors' awareness of teamwork, speaking up, and stress management. Many of the participants indicated that they would use the knowledge gained in their future teams.

While there is much human factors work that is needed in first responders' environment, we do know that stress, unpredictable working environments, ergonomics, and teamwork can affect a person's ability to function in these positions. While we cannot change the nature of the job in law enforcement (i.e., people will still lie), we can provide tools that help officers better control their own information-processing levels and lessen the stressful effects. As we cannot change the dim lighting conditions of a burning building, we can provide ergonomically sound equipment recommendations for future equipment purchases. Finally, the importance of individuals' perceptions of teamwork, their willingness to participate as alert and functioning team members, and their understanding of how well-functioning teams work together cannot be emphasized enough.

REFERENCES

Baumann, M. R., Gohm, C. L., & Bonner, B. L. (2011). Phased training for high-reliability occupations: Live-fire exercises for civilian firefighters. *Human Factors, 53*(5), 548–557.

Driskell, J. E., Salas, E. & Driskell, T. (2012). Social indicators of deception. *Human Factors, 54*(4), 577–588.

Elliott, L. J. (2011). *On the role of analogical reasoning in operational decision making: Positive and negative correlations for decision-based cue evaluation* [Doctoral dissertation, New Mexico State University].

Hagemann, V., Kluge, A., & Kehren, C. (2015). Evaluation of crew resource management interventions for doctors-on-call. In D. de Waard, J. Sauer, S. Röttger, A. Kluge, D. Manzey, C. Weikert, A. Toffetti, R. Wiczorek, K. Brookhuis, & H. Hoonhout (Eds.), *The Human Factors and Ergonomics Society Europe Chapter 2014 Annual Conference* (pp. 237–253). Retrieved April 23, 2021, from https://www.hfes-europe.org/human-factors-high-reliability-industries-2

Hughes, A. M., Patterson, P. D., Weaver, M. D., Gregory, M. E., Sonesh, S. C., Landsittel, D. P., Krackhards, D., Hostler, D., Lazzara, E. H., Wang, X., Vena, J. E., Salas, E., & Yealy, D. M. (2017). Teammate familiarity, teamwork, and risk of workplace injury in emergency medical services teams. *Journal of Emergency Nursing, 43*(4), 339–346.

Koder, G. (2005). *Ergonomics and the future of the police vehicle*. Florida Department of Law Enforcement. Retrieved July 13, 2020, from http://www.fdle.state.fl.us/FCJEI/Programs/SLP/Documents/Full-Text/Koder-george-paper.aspx

Park, K., Hur, P., Rosengren, K. S., Horn, G. P., & Hsiao-Wecksler, E. T. (2010). Effect of load carriage on gait due to firefighting air bottle configuration. *Ergonomics, 53*(7), 882–891.

Patterson, P. D., Buysse, D. J., Weaver, M. D., Suffoletto, B. P., McManigle, K. L., Callaway, C. W., & Yealy, D. M. (2014). Emergency healthcare worker sleep, fatigue, and alertness behavior survey (SFAB): Development and content validation of a survey tool. *Accident Analysis & Prevention, 73*, 399–411.

Patterson, P. D., Weaver, M. D., Weaver, S. J., Rosen, M. A., Todorova, G., Weingart, L. R., Krackhardt, D., Lave, J. R., Arnold, R. M, Yealy, D. M., & Salas, E. (2012). Measuring teamwork and conflict among emergency medical technician personnel. *Prehospital Emergency Care, 16*(1), 98–108.

Ramstrand, N., Zügner, R., Larsen, L. B., & Tranberg, R. (2016). Evaluation of load carriage systems used by active duty police officers: Relative effects on walking patterns and perceived comfort. *Applied Ergonomics, 53*, 36–43.

Ricanek, K., & Boehnen, C. (2012). Facial analytics: From big data to law enforcement. *Computer, 45*(9), 95–97.

Vrij, A., Granhag, P. A., Mann, S., & Leal, S. (2011). Outsmarting the liars: Toward a cognitive lie detection approach. *Current Directions in Psychological Science, 20*(1), 28–32.

CHAPTER 29

PRIVACY AND SECURITY

GOALS OF THE CHAPTER

- Understand the misinformation effect
- Understand the challenges that security analysts face
- Know how ownership changes perception of responsibility

ASSESSMENT

- How is misinformation spread on social networks?
- Why is misinformation so effectively spread through friends?
- What contributes to understanding privacy and security in users?

You may be wondering why the chapter on privacy and security is in the "Safety and Error" section of this book. Many computer users are not aware of the damage that can be done and how accidents can happen when they employ a liberal privacy policy for their online accounts. In addition, organizations seem to have a poor understanding of security. Privacy, security, and **misinformation** create significant damage to individuals, organizations, and communities.

MISINFORMATION

Memory is malleable and prone to reconsolidation or reremembering. The work of Loftus has demonstrated how memories can be unreliable and changeable. When memories are remembered, they are most vulnerable to being changed by new information or information that has been replaced. The misinformation effect is compounded by retroactive interference or new information that "rewrites" the old information. Memory is open to suggestibility, or the influence of others on our memories, and it is open to misattribution, or the influence of forgetting the original memory and instead using something that is likely and close to the original. This is called the misinformation effect (https://en.wikipedia.org/wiki/Misinformation_effect).

For example, during the 2016 U.S. presidential election, many ads appeared online and on social media making untrue claims about a candidate. A subsequent analysis found that these ads may have persuaded people who would have voted for this candidate to vote for the other candidate. The movie *The Great Hack* (https://www.imdb.com/title/tt4736550/?ref_=fn_al_tt_1) tells the story of Cambridge Analytica as it relates to this and other events. As a result, social media organizations have built algorithms that screen ads and messages for this type of content (https://money.cnn.com/2017/05/09/technology/facebook-fake-news/index.html).

Social media is populated by the users' friends and the friends of friends. According to Chong, Farquharson, Choy, Lukman, and Mokhtar (2011) and Binzel and Fehr (2013), friends create bonds between them. The strength of these bonds creates social capital with the relationships between the people creating a framework or the social capital framework. Increased social capital increases trust and reciprocity. According to this premise, people who have strong bonds with others, such as those on a social network, are more likely to trust friends than other sources for critical information. The social capital framework, along with retroactive interference, is thought to have contributed to social media's influence.

The untrue claims about the candidate spread rapidly through social networks. One such claim was "Pizzagate," which is described here: https://en.wikipedia.org/wiki/Pizzagate_conspiracy_theory. The claims grew into conspiracy theories and proliferated social media at the time. In the Pizzagate example, an individual took action on these false claims. Misinformation, promoted by social networks, can be reduced by the removal of the misinformation and educating the social media users on how false claims spread, the security of their networks, and a better understanding of privacy and how it can impact their lives.

SECURITY

There are two parts to security. The first is the security professional's side of helping organizations keep a system operable and safe from intrusion. The second is a combination of security and privacy from the user's point of view. First, let's consider the design professional's challenges as they impact the ability to create a safe and usable system. According to Borghetti, Funke, Pastel, and Gutzwiller (2017), security professionals are tasked with determining intrusions during constantly changing and fluid environments. The work is time sensitive and stressful, with a high degree of uncertainty. The tools that they use are new as well, with design inconsistencies that may not uniformly deal with operational challenges.

SECURITY ANALYSTS

Vieane, Funke, Greenlee, Mancuso, Borghetti, Miller, Menke, Brow, Foroughi, and Boehm-Davis (2017, p. 375) state that "as reliance on cyber systems increases, so too will adversarial efforts to attack, exploit, and disrupt. . . . Maybury [a top U.S. Air Force scientist] encouraged research and technological development to improve cyber defense capabilities." Network analysts monitor internet traffic, and when they are alerted to violations of security policies, acceptable use policies, or standard security practices, the system will send an alert to the analyst. Then the analyst determines if the event actually happened or not. Most of the alerts generated by the system are false alarms. However, as Vieane and colleagues (2017) state, a standard protocol is followed that consists of the following steps:

1. Determine the threat and the evidence.

2. Investigate sources that the threat exists in firewall logs, network packets, and so on.

3. If enough evidence is found, the information is forwarded to a threat analyst who decides if the quality and amount of evidence are enough to warrant further investigation.

According to Vieane and colleagues (2017), analysts report that some of the challenges of the job include chronic job and task stress and high workload. Interruptions in the form of phone calls, emails, and coworkers create a significant portion of the stress, as the analysis task requires continuous concentration. Prior research in human factors and cognitive psychology has indicated that errors increase during repeated interruptions during a critical task, as people de-engage from the task to address the interruption and then must mentally reengage in the original task. During the process of de-engagement and reengagement, critical information stored in working memory is waiting to be processed to long-term memory. This information may be lost during the interruption. Delaying the interruption until it occurs between investigations rather than during an investigation is thought to be key to reducing stress. Indeed, in a simulated cybersecurity investigation task, Vieane and colleagues found that interruptions were detrimental to complete, accurate, and thorough investigations, and limiting interruptions helped the participants engage fully in the task with fewer errors.

In both Vieane and colleagues (2017) and Borghetti, Funke, Pastel, and Gutzwiller (2017), it was noted that the study of security analysts is critical. Much work is still needed to understand the challenges presented to analysts and to develop recommendations on better tool design. The domain resembles other time-sensitive and judgment-sensitive disciplines that require considerable expertise and carry considerable consequences of poor decisions.

STAKEHOLDERS AND SECURITY

The research is sparse in terms of security and individual users. In Stanhope and Elliott (2015), researchers asked participants to arrive at the lab with their laptop computers. Then they accessed a website using either their own computers or a university computer. During their use of the website, an alert box appeared. The amount of time that it took for the participant to react to the alert box was measured, and the action that they took to continue the study was measured. If a participant used a university computer, they took longer to react than if they used their own laptop. Subsequent to the study, users were asked why they took longer and if ownership factored into their decision to act. Their responses indicated that they considered ownership when weighing security risks.

In Chin, Felt, Sekar, and Wagner (2012), a survey found that people who owned both a smartphone and a laptop had greater concerns over security and privacy on their smartphone than on their laptop. They were less likely to use their smartphone for financial and other sensitive tasks. The researchers found that participants had a poor conceptual model of how wireless networks function.

PRIVACY

Privacy and security are typically together in application settings. Users seem to have a poor understanding of both and an even poorer understanding of how they interact. Usability experts and development teams note that with any increase in the usability of a system, the security of a system suffers. Organizations struggle to find a middle ground that satisfies security requirements, protects users' privacy, and allows tools to support normal human behaviors, such as password recovery.

However, users will unintentionally subvert the system by using the same password for multiple sites, writing the password down on sticky notes near their computer, using names posted publicly on social media as passwords, posting their actual birthdate on social media, and choosing an email name that includes their birth year. Users fail to address privacy concerns. Until the European Union decided to address issues of privacy in 2018 with the General Data Protection Regulation (GDPR; https://digitalguardian.com/blog/what-gdpr-general-data-protection-regulation-understanding-and-complying-gdpr-data-protection), privacy was largely dependent on an individual user's understanding of the risks. Since the GDPR, organizations such as Google or Facebook have strengthened policies and educated users on how their data are collected, stored, used, and exposed.

Development teams have historically addressed user privacy and security issues promptly and worked to develop tools to address breaches and weak codes. However, as Acar, Fahl, and Mazurek (2016, p. 3) note, "Historically, the huge gap between the theoretical strong security offered by these mechanisms and actual low security in practice is often caused by the poor usability of security solutions." Researchers have taken a two-pronged approach to the problem: defining the vulnerabilities by creating taxonomies and creating potential threat analyses for development teams. Antón, Earp, and Reese (2002) and then Meis, Wiertz, and Heisel (2015) have created taxonomies that list key vulnerabilities in many applications. Hatamian, Serna, and Rannenberg (2019) created an analysis program, MARS, for exploring user reviews of privacy and then quantifying the top concerns. Both approaches have been useful, but additional approaches are needed. Privacy, security, and misinformation are three areas of research development that will continue into the next century.

REFERENCES

Acar, Y., Fahl, S., & Mazurek, M. L. (2016, November). You are not your developer, either: A research agenda for usable security and privacy research beyond end users. In *2016 IEEE Cybersecurity Development (SecDev)* (pp. 3–8). IEEE Computer Society.

Antón, A. I., Earp, J. B., & Reese, A. (2002, September). Analyzing website privacy requirements using a privacy goal taxonomy. In *Proceedings IEEE Joint International Conference on Requirements Engineering* (pp. 23–31). IEEE Computer Society.

Benenson, Z., Gassmann, F., & Reinfelder, L. (2013). Android and iOS users' differences concerning security and privacy. In *CHI'13 extended abstracts on human factors in computing systems* (pp. 817–822). Association for Computing Machinery.

Binzel, C., & Fehr, D. (2013). Social distance and trust: Experimental evidence from a slum in Cairo. *Journal of Development Economics, 103*, 99–106.

Borghetti, B., Funke, G., Pastel, R., & Gutzwiller, R. (2017, September). Cyber human research from the cyber operator's view. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 350–350). SAGE.

Chin, E., Felt, A. P., Sekar, V., & Wagner, D. (2012, July). Measuring user confidence in smartphone security and privacy. In *Proceedings of the Eighth Symposium on Usable Privacy and Security* (pp. 1–16). Association for Computing Machinery.

Chong, S. T., Farquharson, K., Choy, E. A., Lukman, Z. M., & Mokhtar, M. K. (2011). Enhancing youth civic engagement and generalized trust through bonding social capital among friends. *Pertanika Journal of Social Science and Humanities, 19* (spec. issue), 57–64.

Endsley, M. R. (2018). Combating information attacks in the age of the internet: New challenges for cognitive engineering. *Human Factors, 60*(8), 1081–1094.

Hatamian, M., Serna, J., & Rannenberg, K. (2019). Revealing the unrevealed: Mining smartphone users privacy perception on app markets. *Computers & Security, 83*, 332–353.

Meis, R., Wirtz, R., & Heisel, M. (2015, September). A taxonomy of requirements for the privacy goal transparency. In *International Conference on Trust and Privacy in Digital Business* (pp. 195–209). Springer.

Stanhope, J. L., & Elliott, L. J. (2015, April). Reactions to internet security warnings: Personal versus university computers. In *Human Factors and Applied Psychology Student Conference*. Embry Riddle Aeronautical University Scholarly Commons. https://commons.erau.edu/cgi/viewcontent.cgi?article=1029&context=hfap

Vieane, A., Funke, G., Greenlee, E., Mancuso, V., Borghetti, B., Miller, B., Menke, L., Brown, R., Foroughi, C. K., & Boehm-Davis, D. (2017). Task interruptions undermine cyber defense. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*(1), 375–379. https://doi.org/10.1177/1541931213601576

UNIVERSAL DESIGN

CHAPTER 30

UNIVERSAL DESIGN

GOALS OF THE CHAPTER

- Learn the definition of universal design
- Understand how universal design is explored
- Understand some of the challenges to testing products

ASSESSMENT

- What are two modifications that a researcher would need to make to their research materials or plan in order to test differently abled participants?
- What is the purpose of universal design?
- What types of disabilities are covered under universal design?

WHAT IS UNIVERSAL DESIGN?

Universal design, or **inclusive design**, is an approach to making products (including software) that accommodates the full range of human abilities throughout the human life span (D'Souza, 2017). This means that the design accommodates people who have cognitive, perceptual, and physical challenges due to occupational factors, age-related factors, or other factors. The move toward universal design started in the 1990s, with a small group of people across several continents. At the time, manufacturers considered universal design from a people-centric point of view: people were disabled by physical, sensory, and mental impairments. This group wished to change that viewpoint to a product-centered viewpoint: "People are disabled by designs and environments that do not consider the full range of human abilities" (Clarkson & Coleman, 2010, p. 127). Guidelines on how to best design different types of products as universal design products and federal guidelines can help manufacturers and development teams produce products and services that are available to everyone.

For the internet, there are a set of guidelines called the **Web Content Accessibility Guidelines** (WCAG 2.2; https://www.w3.org/TR/WCAG22/). These guidelines specify how media are to be displayed so that blind and deaf internet users can access the content, how colors should be used so that color-blind individuals can easily read, and how websites should be adaptive to "no keyboard" use. These adaptations are useful to all users. There are many instances where I wish to look at a webpage and have no keyboard, so the "no keyboard" adaptation is one that I use frequently. There are other adaptations that you may use as well without knowing that these are implemented because of universal design.

Until recently, many manufacturers and development teams have ignored universal design, as they felt that the approach added time and extra expense to the project. As Schmutz and colleagues (2016) note, 85% of the world's population is not disabled, and most web content creators focus on designing for the majority—95% of websites are classified as nonaccessible and not incorporating universal design principles. Manufacturers state that there is no demand for accessible sites and little financial incentive to undertake this added expense. Recent studies have explored the usefulness of implementing the guidelines and found that when even one or two of the WCAG guidelines are implemented, the website becomes easier for everyone to navigate and is rated higher on aesthetics and trustworthiness (Schmutz et al., 2016). However, manufacturers and developers are correct in asserting that universal design takes extra care. Song and Lee (2008, pp. 447–448) explored a set of universal design guidelines to help development teams, as seen in Table 30.1.

One of the challenges in creating accessible sites is in user testing on **impaired users**. Impaired users can be difficult to find, and gaining

Table 30.1. Universal Design Guidelines

Principle	Explanation
Equitable use	The design attracts people with diverse abilities.
Flexibility in use	The design allows people with different preferences to customize and adapt it to their abilities and needs.
Simple and intuitive use	Beginning users and users with low-level language skills or impaired thinking can use it easily.
Perceptible information	It is easy to see in all light conditions and communicates to those with sensory impairments.
Tolerance for error	The potential for unintended consequences is minimized.
Low physical effort	It does not cause fatigue and is low effort.
Size and space for approach and use	A person of any size can use this, whether they use a wheelchair or not.

Source: Song and Lee (2008, pp. 447–448).

their trust can be challenging. Stonewall, Fjelstad, Dorneich, Shenk, Krejci, and Passe (2017) suggest some strategies to engage **underserved populations** that can help when testing special user populations. Stonewall and colleagues suggest that the research team "earn trust through partnership" with local organizations through endorsements and cobranding the advertisements for the user testing. They also suggest that researchers create an alliance with an existing authoritative or community figure who has respect and power within the population of interest. Researchers should "be multilingual and inclusive." This may mean learning the challenges, language, and cultural norms of the community of interest. For example, in the blind community, there are different levels of blindness with different terms for when vision is lost and how much vision remains. A person who has had vision all their life and then loses their central foveal vision due to macular degeneration has a different perspective than a person who never had vision. Knowing these terms and knowing how people prefer to be addressed is important for gaining trust.

Finally, Stonewall and colleagues (2017) suggest that researchers strive to respect participants' different time schedules and norms, offer something of use, and understand the participants from their viewpoint. This means making the advertisements and the user testing experience itself accessible for all participants. For example, if the user-testing study sought blind participants, advertising on radio or another auditory medium would be preferred. Respecting the participants also means accommodating travel challenges for the differently abled who may be unable to drive. It may also include designing the research experience so that the participants benefit as well, with a stipend for their time, a useful product, or helpful information. Stonewall and colleagues describe a study about energy use where the participants were given caulk and instructed how to seal gaps in windows and doorways. Including accessible user testing into a testing plan may include providing materials in braille or closed-captioning video, having a sign language interpreter on the research team, and making sure that the physical environment is wheelchair accessible (Lazar et al., 2013).

The training of new human factors professionals on how to incorporate universal design into their testing plans includes an understanding of what universal design entails in its entirety. This may mean specific courses on disability and design. D'Souza (2017) suggests that new professionals understand **ergonomics** as it relates to obesity and physical disability, demographic and social trends in universal design, and how to measure the extent of the different types of impairment. In practice, the use of field studies to explore the interaction of the product in the actual environment can reveal subtle design flaws that an average person may miss. These flaws are glaring when a differently abled person uses the product. For example, a person with cognitive impairments may be catching a city bus. This person is using an application on their phone to estimate when the bus will arrive at the stop and how long they must wait. The application should consider the time of day and how likely it is that the user will be distracted or experiencing a high cognitive workload. If this is considered, the application would allow the user to reduce the amount of information presented during cognitive overload. This would help not only the cognitively compromised user but the average user, as both would experience cognitive overload, with increased noise, traffic, and additional riders waiting for the bus.

One of the best ways to grow a new area of research is through conferences. There are a few conferences and organizations focused on promoting and sharing research about universal design: the International Association for Universal Design (INCLUDE) and the Cambridge Workshop on Universal Access and Assistive Technology (CWUAAT), among others. CWUAAT highlights tools, guidelines, and best practices to help development teams design for all potential users (Lazar et al., 2013). They include a wider range of disabilities than other conferences, as they consider users with perceptual disabilities, motor disabilities, age-related disabilities, low literacy, socioeconomic disabilities, and cognitive disabilities. Researchers and attendees of CWUAAT hope that by characterizing the environments and the struggles of different groups, the research in these areas will increase and provide guidelines to professional practitioners adopting universal design.

In addition to guidelines and best practices, Clarkson and Coleman (2010) advocate the use of a public database so that all teams may access data that address a practical issue such as biomechanics. A database could also create a taxonomy of sensory and cognitive ranges with minimums and maximums. A public, open database such as this would allow manufacturers to have enough information to develop solutions without user testing. This approach may create interface and design solutions, such as using sound to display weather data for the blind (Clarkson & Coleman, 2010). While manufacturers may state that accessibility is not needed for their product because of few complaints or the perception that their population does not have these difficulties, the universal design guidelines address best design practice for all users regardless of ability. As my differently abled friend says, “Those of you who think of yourselves as ‘normal’ are simply people who do not have deficits yet.”

REFERENCES

Clarkson, J., & Coleman, R. (2010). Inclusive design. *Journal of Engineering Design, 21*(2), 127–129.

D'Souza, C. (2017, September). Topics in inclusive design for the graduate human factors engineering curriculum. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 403–406). SAGE.

Lazar, J., Langdon, P., & Heylighen, A. (2013). Designing inclusive systems. *Journal of Usability Studies, 8*(4), 90–92.

Schmutz, S., Sonderegger, A., & Sauer, J. (2016). Implementing recommendations from web accessibility guidelines: Would they also provide benefits to nondisabled users. *Human Factors, 58*(4), 611–629.

Song, K., & Lee, S. (2008, July). Mapping user accessibility needs systematically to universal design principles. In *Asia-Pacific Conference on Computer Human Interaction* (pp. 446–456). Springer.

Stonewall, J., Fjelstad, K., Dorneich, M., Shenk, L., Krejci, C., & Passe, U. (2017, September). Best practices for engaging underserved populations. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 130–134). SAGE.

CHAPTER 31

AGING

GOALS OF THE CHAPTER

- Learn about the range of disability that can happen in aging adults
- Understand the challenges of different everyday activities
- Learn some approaches to categorizing decrements

ASSESSMENT

- How are adults categorized in aging research?
- Where are falls most likely to happen in the home?
- What contributes to motor/locomotion problems?
- What types of problems with driving contribute to automobile accidents?

Older adults are often maligned for their deficiencies, ideas that seem outdated, and grouchiness. As Mark Twain said, "Do not complain about growing old. It is a privilege denied to many." Indeed, as modern medicine has helped us live longer and fuller lives, it has increased the number of elders. The National Institutes of Health (NIH) states that in 2015, 8% of the population was age 65 or older. They estimate that by 2050, this will be 17% (https://www.nih.gov/news-events/news-releases/worlds-older-population-grows-dramatically). Within the United States, the population of those age 65 and older is projected to triple in three decades, with those who are 80 or older increasing dramatically as well. While this is good news for longevity, researchers, and those who attain these milestones, human factors practitioners expect that this will have a dramatic impact on product and service design.

From a psychological point of view, the deficiencies brought about by aging are unlike other **universal design** challenges. Most of the aged previously had abilities within the normal range. Beginning in our mid-20s, muscular strength and perception begin to wane (http://www.scholarpedia.org/article/Touch_in_aging). This combines with slower reaction and information-processing time to make the older brain seem like it runs on molasses to people in their late teens and early 20s, when the brain is at its processing peak. Charness (2008) notes that the **processing difference** can be between 50% and 100% when comparing the young brain of 20-year-olds and the older brains of those above 50 years old. This is part of the normal aging process.

According to Charness (2008), there are **three categories of older adults**: "The young-old (65–74), middle-old (75–84), and old-old (85+)" (p. 548). These bands define the ranges for noticeable decline. However, Charness notes that decrements are not universal or constant. Two similarly aged individuals living in the same household may have different patterns of impairment, and they may have "good days" where the impairment is barely noticeable or "bad days" where the impairment is profound.

Fiske and Rogers are some of the top human factors researchers in a field called **applied gerontology**. In 1998, Rogers, Meyer, Walker, and Fiske explored the types of challenges older adults have in everyday activities. They categorized these into three types of activities: **activities of daily living** (**ADL**; e.g., bathing, eating, and toileting); **instrumental activities of daily living** (**IADL**; e.g., preparing meals, maintaining a house, managing medications); and **enhanced activities of daily living** (**EADL**; e.g., learning and adapting to new technologies and procedures; p. 111). They interviewed 59 independently living adults from the age of 65–80 years old and discussed what challenges they were having. They coded the challenges according to the following list of factors (Rogers et al., 1998, p. 111): "(a) the locus of the problem (motor, visual, auditory, cognitive, external, or health limitations); (b) the activity involved (e.g., transportation, leisure, housekeeping); (c) whether the problem was attributable to task difficulty or the perception of risk; and (d) response to limitations (perseverance, cessation, compensation, or self-improvement)."

The interview answers were transcribed, and the researchers found that challenges in motor limitations were the most frequent followed by cognitive limitations related to different types of memory tasks. Auditory and visual challenges were the least common. Within the types of activities, leisure activities were the most frequently mentioned, followed by transportation, housekeeping, and locomotion. Most of the problems were perceived to be in task difficulty rather than in the risk associated with the task.

Rogers, Meyer, Walker, and Fiske (1998) reported that participants' knee and back problems made it difficult to walk, bend, climb stairs, and carry heavy things. The participants also reported that if they had arthritis, they could not use buttons, open medicine bottles, write, sew, or open certain bags (p. 119). There were also problems with balance. Participants felt that they might fall when exiting their car, getting up from a soft chair, or using stairs or ladders.

In addition to knee and back problems, Rogers and colleagues (1998) found that **cognitive limitations** impacted participants' lives as they aged. Participants reported problems with recalling facts or when completing complex procedures. They listed things such as operating their home security, operating a lawnmower, filling out Medicare forms, using a computer, and using complex audio-visual equipment. Participants also had problems remembering where they put something. Participants reported having trouble remembering that something was on the stove and unintentionally burning the pan.

Rogers and colleagues also found other frustrations not related to personal limitations. These were "fear of crime, financial limitations, frustrations 'with things that break or do not work', loneliness, and dependence on others" (p. 117). In aging adults without disease, there will be less dramatic changes. Gradually, other faculties diminish, as older adults may have difficulty driving, managing their medications, climbing stairs, and paying bills. Family members may notice an increase in dangerous driving habits or accidents. Older adults may forget to pay bills or take medications on time. They may be noticeably slower when walking or climbing stairs. These are areas where design can make a difference.

Currently, there are some design recommendations and guidelines for websites at https://nnlm.gov/mar/guides/making-your-website-senior-friendly and https://www.w3.org/WAI/older-users/developing/. Product design guidelines can be found at https://aginginplace.com/universal-design/ and at https://www.aarp.org/home-family/personal-technology/info-2014/is-this-the-end-of-the-nursing-home.html. Numerous environmental guidelines can be found at https://www.aarp.org/livable-communities/info-2014/aarp-home-fit-guide-aging-in-place.html and https://www.architecturaldigest.com/story/aging-in-place-guide. There is also an emerging market of 3D-printed adaptive items to help people stay in their homes and cope with disability: https://www.arthritissupplies.com/, https://www.thingiverse.com/thing:1020816, and https://grabcad.com/library/3d-printable-jar-opener-1#!

DRIVING

According to Smith (1990), some of the difficulties with driving for aging adults are attributable to seeing at night, seeing signage, physically turning to see what is behind them while backing up, exiting highway traffic, reading the instruments, and reaching the seat belts. Accident data suggest that backing up, merging on the highway, changing lanes, and yielding to the right of way are the most common difficulties.

PREVENTING FALLS AT HOME

Most elderly people live in their own homes, and most would prefer this over a group home or nursing home. Yet one of the primary challenges in keeping the middle-old and the old-old in their homes is the prevention of falls. Smith and colleagues (1990) note that as physical disability and cognitive impairment increase due to age, medication side effects, or alcohol use, falls are the most hazardous and likely accident to occur in the home, resulting in a hospital stay or entrance to long-term care facility. The most likely behavior resulting in a fall is associated with the use of stairs.

MEDICATIONS

As age increases, many adults find that medication is necessary. **Hypertension** is very common among older adults, with 70% of older adults taking prescription medication to manage hypertension (Blocker et al., 2017). Adhering to the prescribed regimen in chronic illness where there are few symptoms can be challenging for older adults. Among these are hypertension and diabetes. Blocker and colleagues surveyed older adults with hypertension to discover adherence strategies that could be shared with prescribing physicians for future recommendations. They found that an association strategy and a location strategy to be the top two strategies participants used to remind themselves to take the medication. In the association strategy, participants paired the medication with part of an existing routine, such as dinnertime or brushing their teeth. In the location strategy, participants kept the medication in a visible spot where they frequently engaged in the associated activity (e.g., next to their coffeemaker in the kitchen or toothbrush). Many participants in the study reported using a combination of both the association and the location strategy to help them remember to take the medication. Simple strategies such as these help people overcome the limits of working memory.

CONCLUSION

As Fiske and Rogers (2002) note, several literature reviews exist to help designers understand the challenges that aging presents to design (Craik & Salthouse, 2011). They also note that in reviewing typical household products, 72% of them posed usability challenges to normally functioning adults in either the products' functioning or their instructions (Hancock et al., 2001). Even more disturbing is that older adults use **complex technology** to monitor chronic diseases such as diabetes and hypertension. The use of these tools is ill specified, and design considerations are largely ignored, with the manufacturers blaming the user for error (Colagiuri et al., 1990). As in other areas of engineering psychology, applying the principles of psychology and known methodologies, such as **task analysis**, reveals design impediments and better design options. When time or financial considerations make a task analysis impossible, Charness (2008) notes that there is a wealth of research on aging as it relates to design in various areas: workplace performance (Sharit et al., 2004), aircraft piloting (Taylor et al., 2005), training principles (Jamieson & Rogers, 2000), technology interaction (Czaja et al., 2006), and medication adherence (Morrell et al., 1990; Morrow et al., 1998).

REFERENCES

Blocker, K. A., Insel, K. C., Koerner, K. M., & Rogers, W. A. (2017). Understanding the medication adherence strategies of older adults with hypertension. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 11–15). SAGE.

Charness, N. (2008). Aging and human performance. *Human Factors, 50*(3), 548–555.

Colagiuri, R., Colagiuri, S., Jones, S., & Moses, R. G. (1990). The quality of self-monitoring of blood glucose. *Diabetic Medicine, 7*, 800–804.

Craik, F. I., & Salthouse, T. A. (Eds.). (2011). *The handbook of aging and cognition*. Psychology Press.

Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., & Sharit, J. (2006). Factors predicting the use of technology: Findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and Aging, 21*, 333–352.

Fisk, A. D., & Rogers, W. A. (2002). Psychology and aging: Enhancing the lives of an aging population. *Current Directions in Psychological Science, 11*(3), 107–110.

Hancock, H. E., Fisk, A. D., & Rogers, W. A. (2001). Everyday products: Easy to use . . . or not? *Ergonomics in Design, 9*, 12–18.

Jamieson, B. A., & Rogers, W. A. (2000). Age-related effects of blocked and random practice schedules on learning a new technology. *Journal of Gerontology: Psychological Sciences, 55B*, 343–353.

Morrell, R. W., Park, D. C., & Poon, L. W. (1990). Effects of labeling techniques on memory and comprehension of prescription information in young and old adults. *Journal of Gerontology, 45*, 166–172.

Morrow, D., Leirer, V., Carver, L. M., & Tanke, E. D. (1998). Older and younger adult memory for health appointment information: Implications for automated telephone messaging design. *Journal of Experimental Psychology: Applied, 4*, 352–374.

Rogers, W. A., Meyer, B., Walker, N., & Fisk, A. D. (1998). Functional limitations to daily living tasks in the aged: A focus group analysis. *Human Factors, 40*(1), 111–125.

Sharit, J., Czaja, S. J., Hernandez, M., Yang, Y., Perdomo, D., Lewis, J. E., Lee, C. C., & Nair, S. (2004). An evaluation of performance by older persons on a simulated telecommuting task. *Journal of Gerontology: Psychological Sciences, 59B*, 305–316.

Smith, D. B. D. (1990). Human factors and aging: An overview of research needs and application opportunities. *Human Factors, 32*(5), 509–526.

Taylor, J. L., O'Hara, R., Mumenthaler, M. S., Rosen, A. C., & Yesavage, J. A. (2005). Cognitive ability, expertise, and age differences in following air-traffic control instructions. *Psychology and Aging, 20*, 117–133.

CHAPTER 32

CHILDREN AND EDUCATION

GOALS OF THE CHAPTER

- Learn about the types of technology products developed for children
- Understand the unique challenges of automation in children's technology
- Learn how automation and artificial intelligence can be used to enhance learning

ASSESSMENT

- How would a human factors researcher assigned to develop a product for 6-year-old children begin?
- What is one challenge for automation in edtech?
- What is the best use of social education robots?

Engineering psychology and **ergonomics** for children have not received enough attention. The current literature in this area focuses on two areas: ergonomics and education. The need for both is self-evident. Children are smaller; their environment should be as well. Education is the primary job of most children; emphasis on scaling their workspaces so they fit properly makes sense.

Developing guidelines for design and determining best practices for children's design in ergonomics have been simple, with the emphasis on the safety of car seats, toys, and playgrounds. The development of guidelines and best practices for educational products and cognitive toys/robots has not been as clear or well regulated. One of the challenges is how to determine how well something is functioning for a person with limited communication ability or developing cognitive abilities. While usability testing is present, the results are often misinterpreted or nonexistent. This problem has emerged recently in education technology.

The COVID-19 virus outbreak and subsequent quarantine have highlighted the need for computers and online services that are safe and easy for children to use. Beginning in the second half of the spring 2020 semester, millions of students attended classes by computer (https://www.weforum.org/agenda/2020/04/coronavirus-education-global-covid19-online-digital-learning/). The result was that millions of children either did not have the ability to connect or failed to understand how to connect and how to use the online learning platforms. Millions of frustrated parents tried to help their children (https://www.devex.com/news/ed-tech-research-hub-launched-95129). Researchers worry that many children in lower socioeconomic conditions have fallen behind academically because of the challenges with technology use. This is a concern for human factors practitioners.

Traditionally, **education technology** (**edtech**) began with quizzes and links to videos or with tutoring systems that helped students learn to read, do simple math, or practice critical decisions. These systems arrived on disks that instructional technologists would install in a classroom full of computers. Students would spend an hour a week using a single program that focused on a single subject.

Edtech today has evolved to provide tutoring based on different skill levels detected through artificial intelligence, complex learning management platforms, advanced online testing platforms that grade essays automatically, and platforms for gathering and assessing knowledge, providing teachers feedback, and nearly everything in between. As the platforms have evolved, so has our understanding of how automation mediates the relationship between the system and the human user to create trust and interaction. Yang and Dorneich (2016) found that different attributes of automation can affect motivation, confidence, satisfaction, performance, and frustration in young learners. When the automation uses predictable etiquette strategies and varies its interaction with the student and when the automation considered the emotional states of the learner, the students were less frustrated and more effective.

Artificial intelligence (AI) incorporated into **social robots** is an emerging area of research and development. Social robots can be companions or tutors. Kulik and Fletcher (2016) have an excellent review of intelligent tutoring systems that are the precursor to today's AI tutors and AI tutoring robots. There are social robots that simply accompany children as allies, social robots that deliver a learning experience, and social robots that deliver science, technology, engineering, and math (STEM) curriculum. Belpaeme, Kennedy, Ramachandran, Scassellati, and Tanaka (2018) reviewed social robots. Social robots are the best choice when the learned skill has a physical component. However, social robots must understand the social context in which they are engaging as well as the learner's aptitude, progress, and motivation. In addition, the technological barriers are still in place for speech recognition and fluid artificial intelligence based on adaptive automation. Prognosticators that envision a world where human teachers are replaced by social robots will be disappointed to learn that this world is still many robotic generations in the future. Robots, like their computer-based brethren, are still best at teaching single skills within a domain, as programming the broader skills is very resource intensive. However, social robots as companions to human teachers in the classroom to train specific skills or to be used for classroom monitoring and assistance could be a closer reality.

One of the newer approaches to making edtech more accessible to children is the use of **participatory design**. Druin (1999) suggests that a **cooperative inquiry approach** is used to elicit design suggestions and usability findings from children. Druin recommends that **contextual inquiry** about the environments in which the devices will be used should be focused on first children as well as human factors professionals as note takers or researchers. Druin notes that children who are similar to the target population can reveal findings that the adult professionals will miss. Then Druin suggests **prototyping with low-fidelity materials**, such as markers, crayons, paper, clay, and string, with children who are slightly older than the target population. Children who are slightly older still retain the memory of what it was like to be the age of the target group but have enhanced communication skills to describe key findings to the adult professionals. Finally, Druin suggests **technology immersion** so that child users may choose which prototype or technology platforms they prefer. Instead of presenting the options serially, the professional should present them all at once so children can compare them.

Edtech and usability for children is an emerging field. As Druin suggests, working with children to create a **child-centered design** requires specific **accommodations**. For the success of the design for children and accommodating the coming changes for edtech, this emerging field is necessary.

REFERENCES

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics, 3*(21), 1–9.

Druin, A. (1999, May). Cooperative inquiry: Developing new technologies for children with children. In *Proceedings of the ACM CHI '99 Conference on Human Factors in Computing Systems* (pp. 223–230).

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research, 86*(1), 42–78.

Yang, E., & Dorneich, M. C. (2016, September). Evaluation of etiquette strategies to adapt feedback in affect-aware tutoring. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 393–397). SAGE.

CHAPTER 33

AUTISM SPECTRUM DISORDER (ASD) AND DEVELOPMENTAL DIFFERENCE

GOALS OF THE CHAPTER

- Learn about the developmental differences that characterize autism and Down syndrome
- Understand what types of products are needed
- Understand how these products are developed for people with different communication skills

ASSESSMENT

- How does a person's sensory system differ with ASD?
- What is the prevalence of ASD in children?
- What are adaptive products?
- Why should a development team consider including an ASD participant in all usability tests for products for a typical population as well as the ASD population?

WHAT ARE DEVELOPMENTAL DIFFERENCES?

Developmental differences are changes in cognitive, sensory, and physical abilities as a result of either a genetic abnormality or another effect that happens either before or shortly after birth. Some developmental differences can be determined before a child is born. Other differences reveal themselves as the child matures.

DOWN SYNDROME AND FRAGILE X

According to the National Institute of Health, Down syndrome includes both physical and mental delays (https://www.nichd.nih.gov/health/topics/down/conditioninfo/symptoms). These can include poor muscle tone, thinking and learning problems, attention problems, impulsivity and judgment problems, sensory loss, and delayed language development. In addition, they are at higher risk for developing other conditions, such as autism spectrum disorder.

Fragile X syndrome is a genetic disorder that affects a person's ability to behave and learn. It affects communication skills and sensory-processing information, according to the National Institute of Health (https://www.nichd.nih.gov/health/topics/fragilex/conditioninfo). There are other intellectual and developmental disabilities that affect learning, communication, information processing, and memory. Social interactions, behaviors, and emotions can be affected as well (https://www.nichd.nih.gov/health/topics/idds).

AUTISM SPECTRUM DISORDER: WHAT IS IT?

One of the most striking developmental disorders is autism spectrum disorder (ASD). There is much yet to learn about ASD. Its impact on the population is increasing each year. In 2020, the prevalence of ASD is 1%–2% of the population, according to the Centers for Disease Control and Prevention (CDC; https://www.cdc.gov/ncbddd/autism/data.html). In children, 1 in 54 has been diagnosed with the disorder. This rate has increased in every 2-year period for the last decade and shows no sign of decrease.

Persons with the disorder have problems with social, emotional, and communication skills. They may have repetitive behavior and demonstrate an inflexibility in routine. Persons afflicted with ASD show heightened touch, sensory, and perceptual sensitivities. They may not want to be hugged or find sound and light grating. They may have difficulty communicating this to others and instead seek solitude, have outbursts, or engage in repetitive behavior to soothe themselves. People with autism have been known to bolt, are attracted to water, and may exhibit aggression and self-injury. The disorder occurs on a continuum, from mildly to profoundly impaired persons (https://www.nichd.nih.gov/health/topics/autism), with some symptoms occurring in some people and not others. There is no universal treatment for ASD. Ashlea McKay describes it well here: https://medium.com/@AshleaMcKay/the-autistic-uxer-designing-with-and-for-a-different-kind-of-mind-4fc1ccaf327b.

Other developmental disorders include cerebral palsy, attention disorders, and seizures. According to the CDC, the prevalence of all developmental disorders in children is about 1 in 6 (https://www.cdc.gov/ncbddd/autism/data.html). Some children may improve as they get older with special therapies, other children may learn to navigate the world with their disorder, and many children do not improve and need accommodations in order to work and live. The medical community, including the CDC, recommends that parents watch for developmental milestones and alert their physician if a child is failing to meet a milestone or their abilities suddenly change in order to treat the disorder early and avert its progression. For those who do not improve with treatment, accommodations can include specially designed environments and products that adapt to their abilities.

ADAPTIVE PRODUCTS

Some products are developed to help a person adapt to work, school, and social situations. Other products help make a person more comfortable at work or home. These accommodations are necessary and mandated by law, as about 2% of adults have ASD in the United States (https://www.cdc.gov/ncbddd/autism/features/adults-living-with-autism-spectrum-disorder.html#). Developing products in this realm can be challenging, as communication between the user and development team is key to a successful design. Also, the type and level of impairment differ between individuals, making some designs suitable for only some of the users. As with many universal design problems, technological solutions for developmental differences are evolving as we learn how to best design for the different abilities.

SENSORY DIFFERENCES

In both ASD and fragile X, a person's sensory system may overreact to a seemingly normal setting. Classrooms, treatment rooms, and home environments require special care to reduce clutter/chaos, noise, and textures in the room. Autism Speaks is an advocacy group for individuals afflicted with ASD. Their website occasionally posts businesses that cater to ASD clients. Architects discuss the importance of designing for ASD through sensory design: https://architizer.com/blog/inspiration/stories/sensory-design/.

One of the key challenges in ASD is the stress that comes from overstimulation, or sensory overload. Products that apply pressure on the skin can help produce a calming and focus effect. This swing is one such product: https://www.nationalautismresources.com/airwalker-therapy-swing/. There are other products, such as this vest: https://www.nationalautismresources.com/weighted-pressure-vest/.

Sensory challenges in other developmental diseases have been accommodated by products that help develop a sense of balance and help people adapt to motor dysfunctions and neurological impairments. Sites such as these help market these therapeutic and adaptive products: https://www.autismenabled.com/, https://shop.autismspeaks.org/, and https://www.autismthings.com.

COGNITIVE DIFFERENCES

As with designing for neurologically typical children, designing for atypical children poses challenges in communication, establishing preferences, and ensuring satisfaction with the design decisions. In typical children and adults, verbal or symbolic questions can be used. In atypical children and adults, other accommodations must be employed. Some usability designers refer to this type of design as neurodiverse design.

The challenge with this type of design approach is trying to understand how to best accommodate a condition that may be completely foreign to your sense of being. As humans, it can be difficult to imagine what it is like to be another person. Shahid, ter Voort, Somers, and Mansour (2016) reveal their approach, which is (1) know your user and their challenges with a task analysis or competitive analysis; (2) imagine your user with use cases, user journey maps, and interviews; (3) design for this user and test and retest with low-fidelity and then higher-fidelity prototypes, using participatory design processes as much as possible. Shahid and colleagues' work can be found here: https://www.nohitch.nl/projects/research-ui-autism/. Varnagy also suggests guidelines when user testing: https://www.uxmatters.com/mt/archives/2015/10/usability-testing-with-people-on-the-autism-spectrum-what-to-expect.php.

As McKay describes, persons afflicted with ASD have issues with social communication and interaction with others. This includes imagining how interactions take place (https://uxmastery.com/researching-designing-for-autism/). McKay mentions some exemplary designs that accommodate ASD, including the New Struan School in Scotland. This school designed its physical environment to accommodate ASD challenges: https://www.scottishautism.org/about-autism/research-and-training/design-autism.

Within the New Struan School design described at https://scottishautism.org/, they consider the "sensory distortion (visual, auditory, tactile, olfactory), perceptual distortion (of light, shade, color, movement), executive function difficulties (personal organization and planning), central coherence difficulties (arranging the details/making connections), inflexibility of thought and action (leading to difficulties in making transitions), and social interaction difficulties." To accommodate these differences, they have included curved walls, omitted as much overhead fluorescent light as possible and used wall lighting instead, and avoided radiators or other visual distractions in the walls, ceilings, and floors (https://www.scottishautism.org/about-autism/research-and-training/design-autism).

DESIGN GUIDELINES

Other resources include these general design guidelines: https://www.designmantic.com/community/designing-for-autistic.php. Web design guidelines can be found here: https://www.forbes.com/sites/robertszczerba/2016/09/08/is-your-new-app-autism-friendly-probably-not/?sh=6db1da684deb.

Interface design guidelines can be found here: https://www.researchgate.net/publication/276495184_User_Interface_for_People_with_Autism_Spectrum_Disorders. Those published by a blog can be found here: https://www.matrixgroup.net/snackoclock/2017/05/designing-users-autism/.

CONCLUSION

As with design for children or aging adults, inclusive design promotes good design practices for all users regardless of their ability. Developers and designers as well as human factors practitioners will note that the design recommendations are similar across all groups. Testing users with developmental differences can yield important findings for adapting not only the current design but also a better design for all users. As in prior inclusive design chapters, differently abled populations should be sought for user testing, as they will identify problems quicker than typical users. Typical users have capacities to compensate for poor design, which they may unconsciously use due to demand characteristics in the user testing paradigm. Differently abled users rarely try to compensate for difficult designs and make the errors obvious to researchers.

REFERENCES

Pavlov, Nikolay. (2014). User interface for people with autism spectrum disorders. *Journal of Software Engineering and Applications, 7*, 128–134.

Shahid, S., ter Voort, J., Somers, M., & Mansour, I. (2016, September). Skeuomorphic, flat or material design: Requirements for designing mobile planning applications for students with autism spectrum disorder. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (pp. 738–745). ACM.

CHAPTER 34

DESIGNING FOR PHYSICAL IMPAIRMENTS

GOALS OF THE CHAPTER

- Learn the range of physical impairments
- Understand how the onset of the impairment affects a person's ability to successfully use accommodations

ASSESSMENT

- How does the onset of the impairment affect a person in blindness?
- In the Be My Eyes app, who are the eyes for the visually impaired?
- Which computer operating systems have screen readers?

VISUAL IMPAIRMENTS

Individuals with visual impairments vary in the amount and type of sight that they retain. Some users have **macular degeneration** and have lost sight in the central part of their visual field. They were once fully sighted when disease reduced their sight gradually over their lifetime. Others never had sight. These groupings are important when understanding the challenges that people have with sight. Those who previously had sight have a **mental model** of the world. Those who were born without sight have no mental model of what the world looks like or how it works.

There are many tools to help people navigate the world without sight. Computer users can use the JAWS screen readers found here: https://www.freedomscientific.com/products/software/jaws/ and https://en.wikipedia.org/wiki/JAWS_(screen_reader). There are screen readers for other devices (https://en.wikipedia.org/wiki/Category:Screen_readers). Nearly all operating systems incorporate a screen reader in the accessibility options in the system. Screen readers are incredibly useful. I like to use them to read manuscripts to me as I edit. I also like to listen to journal articles when I travel to save time in the office.

Braille, books on tape, and new apps are available to help people navigate without a caregiver. Funded contests allow developers to compete for the benefit of the disabled community: https://www.fastcompany.com/3059158/crowdsourced-knowledge-is-helping-people-with-disabilities-navigate-life. One of my favorite apps is Be My Eyes (https://www.bemyeyes.com/). This app pairs sighted volunteers with blind users. The blind user can call someone with the app, point the camera, and the sighted volunteer will tell the blind person what the camera sees. According to their website, there are over 3 million sighted volunteers who have downloaded the app to help the blind users.

AUDITORY IMPAIRMENTS

As with a visual disability, a hearing disability can begin at birth, the loss can be gradual throughout a lifetime, or the loss can happen suddenly. Persons who have had hearing and lost it have different adaptations than those who never had hearing. Hearing aids and cochlear implants are available. However, for many people, a hearing aid or implant does not restore hearing. In these cases, there are many accessible features in computer operating systems, televisions, and some video game consoles. Sign language is widely accepted and available in addition to captions.

PHYSICAL IMPAIRMENTS, MOBILITY, AND COORDINATION

Physical impairments can include any physical condition that causes a person to function differently from others. This could be differently sighted, different hearing, different mobility or dexterity, different stamina, or a different sense of being conscious. These impairments could be caused by a birth event, a genetic event, aging, accident, or illness. The cause of the impairment is unimportant; the future functioning and accommodation are the primary design considerations. There are several advocacy groups that promote legislation designed to protect differently abled persons. Advocacy groups provide resources and can be found here: https://en.wikipedia.org/wiki/Disabled_Peoples%27_International, https://en.wikipedia.org/wiki/Disability_culture, and https://en.wikipedia.org/wiki/Disability.

When people must use a wheelchair, they become physically barred from locations outside of their homes. Many nonprofit organizations are advocating for change: https://meril.org/ and https://www.reach.ca/. Organizations such as these match current regulations to an organization's existing compliance and provide recommendations. These organizations also provide resources for individuals to address problems with accommodations in the United States and in Canada. Some of the regulations can be found in the Americans With Disability Act: https://www.ada.gov/pubs/adastatute08markscrdr.htm. The American Association of Retired Persons (AARP) estimates that over 12% of the population will be over 65 years old after the year 2000 (Maisel et al., 2008). As mentioned in prior chapters, aging often brings physical impairment, with stairs as a top impediment.

While government buildings and schools must provide entrances and bathrooms that are accessible to wheelchairs, private homes do not. In the 1980s, visitability began to change this approach: https://visitability.org/. The idea, which began in the United Kingdom and Canada, was that all newly constructed homes should have an outside door that does not have stairs. It should have doorways that are at least 32 in wide to accommodate a wheelchair. There should be at least one bathroom that is large enough for a wheelchair-bound person to be able to use.

Good design provides benefits to disabled and nondisabled users (Schmutz et al., 2016). Enabling inclusive design helps the human factors professionals find the weaknesses in any design. As in prior chapters, accommodating differently abled users in all user experience studies is recommended.

REFERENCES

Maisel, J. L., Smith, E., & Steinfeld, E. (2008). Increasing home access: Designing for visitability. *AARP Public Policy Institute, 14*, 1–34.

Schmutz, S., Sonderegger, A., & Sauer, J. (2016). Implementing recommendations from web accessibility guidelines: Would they also provide benefits to nondisabled users. *Human Factors, 58*(4), 611–629.

CHAPTER 35

CONCLUSION

GOALS OF THE CHAPTER

- Learn the key points in each of the chapters
- Understand how the book works together
- Understand the relationship between the chapters

Now that I have spent 2 years writing this book, I realize that it was to answer the question "What is an engineering psychologist?" It is imperative that our discipline survives and thrives as reliance on technology will increase in the coming years. With this reliance comes challenges to our safety, our performance, and our information-processing abilities brought about by the integration of technology into all aspects of our lives. I hope that this book has given you a good foundation in the key concepts and that you might consider joining the profession. In summary, the different parts of the book had key concepts that are foundational to the profession.

THE DESIGN OF COGNITIVE WORK

This section discussed 11 basic research areas of engineering psychology that overlap with cognitive, sensation, and perception research in experimental psychology. In Chapter 3, we discussed **attention**, **vigilance**, and **fatigue**. In this chapter, the **bottleneck theory** and **task switching theories of attention** were covered along with **inattentional blindness** and **change blindness**. The **SEEV model**, **visual search**, **vigilance**, and **signal detection theory** were also discussed.

In Chapter 4, information processing and memory were discussed. This included **the stages and types of memory**, **cognitive fatigue**, and **information processing**. The chapter concluded with a discussion of **cognitive engineering**.

In Chapter 5, the training and automaticity chapter, **conditioning** and **practice** were discussed. In Chapter 6, the stress and workload chapter, **information processing**, **cognitive workload**, and **multiple resource theory** were discussed. In Chapter 7, the displays, monitors, and screens chapter, the **proximity compatibility theory**, **Fitts law**, and the **Hick-Hyman law** were discussed.

In Chapter 8, the usability chapter, the **product development life cycle**, how the constructs of usability are measured, and **design guidelines** were discussed. In Chapter 9, the teams chapter, **trust, shared mental models**, **situation awareness**, and **team communication** were discussed. In Chapter 10, the situation awareness and metacognition chapter, the **levels of situation awareness** and their measurement were discussed along with the relationship between situation awareness and metacognition.

In Chapter 11, the emotion, motivation, and boredom chapter, the **technology acceptance model**, the **rational actor theory** and its prediction of engagement with technology were discussed along with the types of **boredom** and the roles of **emotion** and **motivation**. In Chapter 12, the decision-making and expertise chapter, the development of **expertise** and **recommender systems** were discussed along with **expert systems**. In Chapter 13, the language and artificial intelligence chapter, **natural language processing** (**NLP**) was discussed along with **latent semantic analysis** (**LSA**), **semantic and syntactic structures**, and the **McGurk effect**.

THE DESIGN OF PHYSICAL WORK

This section discussed nine basic areas of research that relate to how we physically operate in the world. Some of these sections overlap with basic research in **cognition**, **sensation**, and **perception**. In Chapter 14, **anthropometry**, **ergonomics**, and the measurement of the human body are discussed, along with the importance of **design for variation** in measurement.

In Chapter 15, the automation chapter, **the levels of automation**, **out-of-the-loop unfamiliarity** (OOTLUF), **overconfidence**, **complacency**, **autonomy**, and **adaptive automation** are discussed. In Chapter 16, human-robot interaction, the **three Ds of automation use** are discussed along with **trust**, **anthropomorphism**, **human-robot teaming**, **measuring trust**, and the **social aspect of robots**. In Chapter 17, **virtual environments**, **immersion**, **VR sickness**, and the **triad of VR** are discussed.

In Chapter 18, vision and visual search, **serial search**, **parallel search**, and **gestalt psychological theories** in search are discussed. In Chapter 19, audition and noise, the **cocktail party effect**, **just noticeable difference**, and **masking** are discussed. In Chapter 20, haptic control and vibration, the differences between a **graphical user interface** (**GUI**) and a **tactile user interface** (**TUI**) are discussed along with **proprioception**.

In Chapter 21, spatial sense and maps, there was a discussion of **two-dimensional space**, **FORT**, **egocentric/exocentric views**, and **spatial cognition**. In Chapter 22, controls and control panels, the **different types of controls**, **gain and lag**, **speed-accuracy trade-off**, **stimulus-response compatibility**, and when to use which controls were covered.

SAFETY AND ERROR, UNIVERSAL DESIGN

In these two sections, application and the unique problems encountered in applying human factors to different areas were covered and organized by topic. In Chapters 23, 24, and 25, **safety**, **alarms**, **accidents and incidents**, and **human error** were discussed. In these chapters, you will find a discussion of the **Swiss cheese model** and **HFACS** as well as topics in Chapter 29 on **privacy** and **security**. In Chapters 26, 27, and 28, you will find a discussion of three domains in which safety and error are of primary concern. These are **aeronautics**, **medicine**, and **first responders**.

In the universal design section, the problems unique to different common design challenges were discussed as well as an overview of **universal design** with many resources for developers and designers. The challenges unique to aging are in Chapter 31. The challenges unique to children are in Chapter 32. The challenges unique to **developmental impairments** are in Chapter 33, and the challenges unique to **physical impairments** in Chapter 34.

I hope that this overview will help you locate your favorite topics or help you study for an exam. The references for each chapter offer deeper and more complete information on the topics discussed. I sought references that would be easy for most readers to find using a common scholarly articles search engine. I also encourage you to attend a conference in usability, human factors, or ergonomics. These conferences and organizations are discussed below.

WHERE TO CONNECT TO OTHERS, FIND JOBS, AND SEEK FURTHER EDUCATION

The largest organization for human factors professionals and educators in the United States is the Human Factors and Ergonomics Society (HFES). Their website has a list of education resources in the United States and Canada: https://www.hfes.org/Resources/Education-Resources. HFES has a yearly conference with several smaller regional and specialty conferences such as a conference devoted to human factors in health care. HFES publishes several journals (https://journals.sagepub.com/home/hfs).

Interaction design is another term for how people interact with products and services. The Interaction Design Association (IxDA) can be found here: https://ixda.org/.

The Association for Computing Machinery (ACM) is the preeminent organization for persons interested in computing. They have several special interest groups (SIGs). The one that human factors professionals join most often is SIGCHI, or the Special Interest Group on Computer–Human Interaction. SIGCHI has its own yearly conference, affiliated journals, and design competitions.

The User Experience Professionals Association (UXPA, https://uxpa.org/) invites human factors professionals who are interested in usability and interaction to participate. The UXPA sponsors an open-access journal focusing on usability studies (https://uxpajournal.org/), a newsletter, a blog, local groups, and a yearly national conference.

Each year, the Applied Human Factors and Ergonomics Conference (AHFE) attracts human factors practitioners, researchers, psychologists, ergonomists, and engineers from across the globe. AHFE publishes the proceedings in a multivolume set of books that begin with the title *Advances in* . . . (https://www.springer.com/gp/book/9783030513689).

As in the words of my mentor, be well and grow the profession, as it will benefit us all.

ASSESSMENT QUESTIONS

The following are the assessment questions listed at the beginning of each chapter:

01 A HISTORY OF ENGINEERING PSYCHOLOGY

- What are the differences between engineering psychologists and human factors engineers?
- How did engineering psychology begin according to psychologists?
- Does an engineering psychologist conduct research or provide therapy?
- Why are engineering psychologists often called engineers when they are not required to pass the ABET or take engineering courses?

02 METHODS OF ENGINEERING PSYCHOLOGY

- What are the names of the types of tests that are done before and after a product is created?
- What are some examples in each category?
- What are three surveys typically used and what do they measure?

03 ATTENTION VIGILANCE AND FATIGUE

- How is attention studied experimentally, and can you give an example of an experiment?
- What is inattentional blindness?
- What is a self-terminating search and what is the difference between a self-terminating and an exhaustive search?

04 INFORMATION PROCESSING

- How does information go from the initial perception to final storage in long-term memory?
- What is an example of a retrieval problem and in what stage did it occur?
- What is cognitive fatigue, and can you give an example in your own words?
- What is cognitive engineering?

05 TRAINING AND AUTOMATICITY

- How are operant and classical conditioning the same, and how are they different?
- What steps would you use to teach a dog to go to the refrigerator and get you a can of Pepsi?
- What are the two types of knowledge, and can you give an example of each?
- What are the two theories of automaticity, and can you explain them in your own words?

06 STRESS AND WORKLOAD

- How does the information-processing pipeline break down when cognitive workload is high?
- What is an example of a high cognitive workload?
- How can stress be measured, and can you name three ways?

07 DISPLAYS, MONITORS, AND SCREENS

- What is a card sort, and why would you do one?
- What is the proximity compatibility principle, and why do you need to use it?
- What is Fitts law, and how is it different from the Hick-Hyman law?

08 USABILITY

- What are the three constructs measured in usability?
- What is the ISO?
- What is an MVP?

09 TEAMS AND PERFORMANCE

- How do surgery teams address mental model differences?
- What is implicit communication in teams?
- What is an outcome of punishing the weakest link in a team?

10 SITUATION AWARENESS

- Which cognitive processes compose situation awareness?
- What are the situation awareness processing levels, and can you give a real-life example of each?
- How is SA measured?
- Why would people say that you cannot have SA without metacognition?

11 EMOTION, MOTIVATION, AND BOREDOM

- Why did most human factors practitioners view the theories of emotion and motivation as simply theories unrelated to technology use?
- What is the technology acceptance model, and how does it influence technology use?
- What is the rational actor model, and how does it influence technology use?
- What are the two types of boredom, and how do they occur?

12 DECISION-MAKING AND EXPERTISE

- How is criticism vital to developing expertise?
- What are the three ways that expert status is determined?
- What is a recommender system, and how is that different from an expert system?

13 LANGUAGE AND ARTIFICIAL INTELLIGENCE

- What is the difference between syntax and semantics in language?
- Why should the people who write error messages be concerned about conversational implicature?
- How are both incorporated in modern AI?
- What is latent semantic analysis, and how does it work?
- What is the McGurk effect, and why is it a problem for AI?
- What is Searle's Chinese room, and how is it a critique of AI?

14 ANTHROPOMETRY

- Why are humans measured, and who measures them?
- What organization ensures that workplaces follow best practices?
- What is the difference between an anthropometricist and an ergonomist?

15 AUTOMATION

- What is OOTLUF?
- What happens as the level of automation increases to OOTLUF?
- Why is situation awareness important?

16 HUMAN-ROBOT INTERACTION

- What are the three Ds?
- How might a human factors professional measure trust in human-robot teams?
- How are telerobots and teleoperators different?

17 VIRTUAL ENVIRONMENTS

- What three items make up the triad of a virtual environment?
- What is virtual reality sickness?
- What are two things that you can do if a person has virtual reality sickness?

18 VISION AND VISUAL SEARCH

- What are three things that influence a person's visual search?
- What is the difference between a serial search and a parallel search?
- What do the gestalt principles do for us in terms of search?

19 AUDITION AND NOISE

- What are 5 of the 11 instances where an auditory signal is preferred over a visual signal?
- What is JND, and how is it used?
- What are the four tasks that sound conveys?

20 HAPTIC CONTROLS AND VIBRATION

- How are a GUI and a TUI the same, and how are they different?
- What is one use of a haptic interface?
- What is one use of a proprioceptive interface?

21 SPATIAL SENSE AND MAPS

- What is the theory of FORT?
- What are the differences between egocentric and exocentric views?
- What is spatial cognition, and why is it important to understanding maps and displays?

22 CONTROLS AND CONTROL PANELS

- What is gain?
- What is lag?
- What is Fitts law?
- What is the speed-accuracy trade-off, and why is it important in design?
- Why is an eye-controlled device impossible for a human to use?

23 HUMAN ERROR

- What is the SCM model?
- What are latent variables?
- How are stress and performance related?
- What does a reliability analysis demonstrate?

24 ALARMS

- How do alarms trigger a “cry-wolf” effect?
- What is signal detection theory?
- What is redundancy gain?

25 ACCIDENTS AND INCIDENTS

- Is there a unified site where accident data can be viewed?
- What are the two types of accident investigation models?

- Which model refers to information processing in a human?
- What does the acronym HFACS stand for?

26 AERONAUTICS

- What initial approach did human factors researchers use to control pilot error during World War I and World War II?
- One model of categorizing error is through a systematic-failure taxonomy. What is the other model of categorizing error?
- What is the primary cause of UAS accidents?
- How many feet above ground level is considered to be "space," as in "space flight"?

27 MEDICINE

- What do the terms EMR and EHR refer to?
- What are the three types of checklists?
- What is the first step in gauging situation awareness according to these researchers?
- Why are medical professionals concerned about patients' use of apps?

28 FIRST RESPONDERS

- What types of things does a police officer carry on their equipment belt?
- What is one thing that a police officer can do to detect a lie?
- Why are firefighters at risk for slips and falls?
- What is an example of a question that you might ask an EMT that would indicate their understanding of team structure and leadership?

29 PRIVACY AND SECURITY

- How is misinformation spread on social networks?
- Why is misinformation so effectively spread through friends?
- What contributes to understanding privacy and security in users?

30 UNIVERSAL DESIGN

- What are two modifications that a researcher would need to make to their research materials or plan in order to test differently abled participants?
- What is the purpose of universal design?
- What types of disabilities are covered under universal design?

31 AGING

- How are adults categorized in aging research?
- Where are falls most likely to happen in the home?
- What contributes to motor/locomotion problems?
- What types of problems with driving contribute to automobile accidents?

32 CHILDREN AND EDUCATION

- How would a human factors researcher assigned to develop a product for 6-year-old children begin?
- What is one challenge for automation in edtech?
- What is the best use of social education robots?

33 AUTISM SPECTRUM DISORDER (ASD) AND DEVELOPMENTAL DIFFERENCE

- How does a person's sensory system differ with ASD?
- What is the prevalence of ASD in children?

- What are adaptive products?
- Why should a development team consider including an ASD participant in all usability tests for products for a typical population as well as the ASD population?

34 DESIGNING FOR PHYSICAL IMPAIRMENTS

- How does the onset of the impairment affect a person in blindness?
- In the Be My Eyes app, who are the eyes for the visually impaired?
- Which computer operating systems have screen readers?

GLOSSARY

Above the fold: The most important items on the screen should be on the first screen.

Absolute identification: The presence of a specific sound signals an alert.

Accident: An event that results in property destruction or personal injury or death (e.g., a tree falls on the metal building and damages it beyond repair).

Accommodation: Changing the design to allow persons with different abilities to use the design efficiently.

Activities of daily living: Bathing, eating, and toileting.

Adaptive automation: Automation that senses the environment and changes for that environment.

Air traffic management (ATM): The people, systems, and procedures that organize air traffic in the national airspace and at airports.

Air traffic management (ATM) simulator: A simulated ATM system.

Alarm fatigue: People ignore an alarm because they have heard too many false alarms and believe the alarms to be mostly false. The frequency of false alarms has conditioned the human to ignore alarms.

Algorithm: A set of procedures that give a particular output with a given input.

Amplitude: Perceiving sound waves as loudness.

Anthropometry: The study of how people fit into physical spaces across all sizes of humans.

Anthropomorphism: A machine or system that looks, acts, or is perceived to have human qualities.

Applied gerontology: Using research to discover the needs of older adults in an effort to design better products to ease their pain points.

Arduino computer: A microcontroller or computer using a single board that is open source and manufactured by Arduino (https://en.wikipedia.org/wiki/Arduino).

Area of interest (AOI): Place you are paying attention to in the environment.

Arrangement: Where or how a target will appear.

Associative learning: Learning how to use something by associating it with a positive outcome or practice. For example, I like to listen to the *BG Podcast* when I exercise because I work out longer when I do. I have associated the *BG Podcast* with running on the treadmill through practice or repeated use.

Auditory memory: A longer and more permanent memory-storage place for sound.

Automaticity: When a person does a task so often that it takes up very little of the attentional bandwidth.

Automation: When a single computerized or mechanical device or series of devices (machines) do something that may be normally done by a human.

Autonomy: A machine, system, or robot that is capable of controlling itself independent of a human.

Bayesian decision-making: A theory that describes how people make decisions based on uncertainty and prior probability of occurrence.

Below the fold: Lesser items should be shown below the first screen when the user scrolls down.

Brain-computer interface (BCI): Brain wave elicited at a certain frequency as a signal to control a computer or robotic device.

Card sort: When a person chooses where it would be easiest for them to find items in different categories.

Central executive: Collates the information and decides what to do next with it.

Change blindness: Happens when a person has simultaneous tasks and misses something that changes in their environment.

Child-centered design: Design that takes into consideration the needs of children's cognitive, physical, and sensory differences from adults.

Chinese room: The idea that if a monolingual English speaker is in a room of Chinese symbols and is given the instructions of what to do with the symbols in English when the symbols come through a door slot, the person can return an answer in Chinese based on the instructions.

Chunking: This refers to combining items to be remembered into a single item that is easier to remember and thereby lessens the effort needed to remember it.

Classical conditioning: May happen intentionally or unintentionally and is a learned process when a stimulus automatically creates a response because it was previously paired with a stimulus that created an automatic physiological response. For example, when you see a picture of snow on a hot day, you feel cooler because you previously associated the experience of snow with being cold.

Cocktail party effect: This effect refers to the phenomenon of standing and talking at a party and then hearing your name being mentioned across the room.

Cognitive biases: Errors in judgment that happen during decision-making.

Cognitive engineering: A new branch of human factors concerned with how people think and make decisions in real situations.

Cognitive fatigue: When the amount of cognitive workload has been high for a certain amount of time and your brain is tired. This differs by person and task.

Cognitive limitations: The inability to process different sensory information, produce language, or process information at a typical pace.

Cognitive load or cognitive workload: The hypothetical amount of cognitive resources used by a person suggested by the bottleneck theory of attention. A person has a limited amount of attention, working-memory capacity, and information-processing capacity. When lots of stimuli happen at the same time and the stimuli are diverse, the situation demands a lot of attention, working memory, and information processing. At times, this can overwhelm a person, and when it does, their cognitive load is high. The amount of effort expended in thought, attention, memory, sensory processing, and decision-making at any given time or for a given task differs by person and task.

Cognitive map: The internal map in our mind that we have of an environment or place.

Cognitive resources: Information processing that requires attention, working memory, and perceptual processes to be engaged.

Cognitive work analysis: In this method, the analyst observes or transcribes what someone does to operate or within a system.

Competence: An ease-of-use concept that suggests that as the system is easier for a user to use, they increase in expertise or their ability to use the system.

Competitive analysis: A method where the analyst compares similar products or services to the one that they are helping develop.

Complacency: When a human attributes an automation error to another cause because the human is unaccustomed to having the automation fail.

Complex technology: Technology that requires many steps to memorize and may require a complex manual to operate.

Computational ways: Methods of representing knowledge in a computer only.

Connected-word speech recognition: The system responds to the utterance of a series of words that are predefined.

Construct: A set of measurable behaviors that represent a larger behavior (in-group) or thing (love) so that we can quantify it and understand it better. Isolating something that is intangible and then describing it in a way that can be studied. When studying the construct of love, the researcher measures the amount of time spent together, the number of positive words exchanged, and the number of valuable items shared or exchanged. As these observable items increase, we assume that love increases.

Contextual inquiry: The researcher immerses themselves in the environment with the person to discover the challenges that the person is having when trying to use the technology or accomplish their goals.

Continuous: Operates along a continuum.

Continuous speech systems: The system responds to a continuous series of words as in natural language.

Convergent problems: A problem set with a single best answer.

Cooperative inquiry approach: The researcher works alongside the user to discover design problems or design solutions that were not visible to the development team.

Course of action: What the person plans to do next based on training, environmental cues, and cultural norms within the profession.

Crew resource management: A training method that focuses on creating norms of increased communication, strong leadership skills, and better judgment tactics for professionals who work in environments that have severe consequences of human error. More information can be found here: https://en.wikipedia.org/wiki/Crew_resource_management.

Critical decision method (CDM): An interview method that was an outgrowth of the critical incident technique in which an analyst asks a decision-maker questions about the information they had prior to the decision and the options that were available to them at different points in the decision-making process.

CRT monitor: Cathode-ray tube monitor screen.

Declarative knowledge: You can explain it and declare it; this is knowledge for facts or "things that you know."

Deleterious events: Things that happen that are unfavorable or unwanted, not to the level of an incident (e.g., a tree branch over a metal building).

Deliberate practice: When a person works to gain expertise by attending to an expert's criticism.

Dependent variable: Something that is measured and thought to be an outcome of an independent variable measured at two or more levels. In the example of a person's comfort in wearing sandals on a snowy day or a rainy day, the person's rating of comfort on both days is a dependent variable.

Detection: Determining if a sound is present or not.

Development life cycle: The time that it takes to create a system or application from idea to physical reality and then through revisions until it is no longer useful.

Discrete: Operates as either on or off.

Divergent problems: A problem set with many best answers.

Divided attention: This is when you were trying to attend to more than one item at once.

Domains: An area of practice of a specific profession.

Drive theory: Based on the idea that when we have a deficit or a need, this need motivates us to solve it.

Drury's model: A two-stage model in which a person first gauges the probability that they will find a target and then decides whether or not to continue the search.

Dual-task paradigm: If we were researching how much attention was left over while doing a primary task, we would assign the secondary task at varying levels of difficulty.

Echoic memory: A temporary and very short memory storage place for sound.

Educational technology (edtech): Created to deliver educational content to students.

Effectiveness, efficiency, and satisfaction: The three constructs that are usually measured in a system, product, or software.

Egocentric view: A frame of reference in a map from the user's point of view, as if a camera were perched on the user's head while they traveled.

Electronic medical records: Information-processing systems made to gather, organize, store, and share patient data.

Enhanced activities of daily living (EADL): Learning and adapting to new technologies and procedures such as learning how to use a new cell phone.

Ergonomics: The study of how people interact with physical items and physical spaces.

Error messages: When a computer has a problem fulfilling the user's request, it returns a message stating that the request cannot be fulfilled. This message is an error message.

Event rate: How frequently or when something that seems like a target or the target item will appear.

Exhaustive search: When the person searches all items in the field.

Exocentric: The viewpoint from anyone's point of view.

Expertise: The proficiency with which a person can function, perform, and make decisions as it builds over time.

Expert systems: Computer systems that mimic the decision-making of experts in a field. These systems are trained with data and then an algorithm that mimics a human expert.

External environmental information: When your body has an effect from outside of itself that affects the entire body's functioning, such as being very cold.

External sensory input: When your body perceives a temporary effect from outside of itself such as hearing an alarm or seeing a flashing red light.

False alarms: When an alarm signals an event that is not urgent, is not true, or does not require a human's immediate attention.

Feedback loop: When a change in a control produces a change in the tool and the feedback given to the human as a result of that change in the control position.

First responders: People who arrive at an emergency first (e.g., firefighters, emergency medical technicians, police officers).

Fitts law: The time that it takes to move from one point to another point is determined by the distance between these divided by the width of the point that someone is attempting to select.

Flying beyond the box: The imaginary zone in which the unmanned aircraft can receive the frequency from the ground control station. This varies by maximum takeoff weight (MTOW) classification and by manufacturer. With small aircraft, it is a few feet. With large aircraft, it can be a few miles. If the operator sets the aircraft to fly too fast or a wind takes the aircraft farther than anticipated, the aircraft can lose the ground control communication, and the operator loses control of the aircraft. If this happens, the operator must find the aircraft physically and reestablish communication.

Focused attention: When you select a particular item and attend only to it.

Focus group: Whenever a product is in the initial design phase, the design team will call on previous and potential users to discuss the product.

Formative tests: Tests of usability before the application and the interface are finished.

Frame of reference: Viewpoint on a map.

Frame of reference transformation (FORT): The mental rotation that people must make to translate a two-dimensional map into a three-dimensional idea of where things are in their minds.

Frequency: Sound waves that are perceived as pitch by our ears.

Functional product specification document: This document is the blueprint or plan for what the product, application, or device can and cannot do. Sometimes it includes specific functionality for a group of stakeholders such as users of the item.

Functional resonance analysis method: This approach is qualitative and seeks to define how individuals' actions separately and together impact a complex system through modeling the categories of behaviors, functions, and structures that the individuals interact with and contribute toward.

Gain: The ratio between the movement of the control and the movement of the system's tool.

Gestalt psychology: A part of psychology that focuses on perception and sensation.

Grammar: The set of rules that govern a language and make it comprehensible to others.

Graphical user interfaces (GUI): A visually oriented interface.

Gricean maxims: *See* **maxims of communication**.

Ground control station/module: The radio, cameras, and communication devices that control an unmanned aerial system vehicle.

These items are usually in the form of a laptop or a workstation or look like a video game controller and are kept on the ground.

Hamburger menu: A series of stacked lines or a square set of dots that indicate there is another menu or series of menus that the user should explore.

Haptic feedback: Underneath your layers of skin, you have neurons that respond to pressure. These neurons tell you when you have grasped the cup of coffee in your hand so you can carry it across the room. The sensation of the cup in your hand is haptic feedback.

Heuristic analysis: A methodology that involves several experts using a checklist of ideal design guidelines to rate the level at which the software or device conforms to the design ideals.

HFACS (Human Factors Accident Classification System): A model by Weigmann and Shappell that categorizes and organizes the system and operator conditions that contributed to the accident or incident being investigated.

Hick-Hyman law or Hick's law: The time it takes for a person to choose is related to the number of items they must choose from, the complexity of the decision, and the consequences of a poor choice.

High responsibility teams: Teams of people who address problems that have severe consequences such as death.

Human error: When an adverse event happens as a result of a human's action or inaction.

Human Error in Air Traffic Management (HERA) project: Certain countries in Europe shared data and research to find the best description of optimal human performance in managing air traffic.

Human error taxonomies: A system of categorizing different types of situations resulting in human error.

Hypertension: High blood pressure that must be controlled or it could lead to diabetes and heart failure. Many hypertensive medications cause other impairments such as the physical impairment of arthritis, sensory impairments, and cognitive changes.

Immersion: The sense that you are actually in the virtual environment.

Impaired users: People who have abilities that are not typical, such as diminished vision, hearing, or cognition.

Implicit memory: A type of memory that relates the sensations that indicate the next steps with the motor movements and what to do next, such as what is used with rock climbing or bike riding.

Inattentional blindness: This happens when we see something but we do not recognize it.

Incident: An event with an unwanted outcome but not to the level of an accident (e.g., a tree branch falls on a metal building and creates a dent in it).

Inclusive design: *See* **universal design**.

Independent variable: Something that is manipulated at two or more levels, and then the reactions are compared at both levels to see if they are different or the same. This could be a person's comfort in wearing sandals on a snowy day or a rainy day. The snowy day or the rainy day would be the independent variables, and comfort would be the dependent variable.

Information processing: Taking sensory information and perceiving it, then deciding what to do next with it: memorize it, act on it, get more information, or ignore it. A series of cognitive tasks that require the person to access knowledge, procedures, memories, and rules to create judgment, situation awareness, planning, and communication.

Instrumental activities of daily living (IADL): Preparing meals, maintaining a house, managing medications.

Interface: The part of the software application or control panel that a person manipulates in order to issue commands to the application. Most interfaces are graphical user interfaces and include that option for a person to select visual items instead of using a command line to use the application.

Internal environmental information: When your body has a temporary internal environmental state that affects the entire body's functioning, such as being fatigued.

Internal sensory input: When your body has a temporary internal state that affects only a part of the body's functioning such as being in pain or feeling confused.

International Organization for Standards: An organization that sets international standards for systems and products.

Isolated word speech recognition: The system responds to the utterance of a single word.

Judgment: The process of selecting from different options or choices.

Just noticeable difference (JND): When 50% or more of the persons tested can detect the signal in a normal noise environment for which that signal is present.

Kinesthetic: The way people move in space.

Lag: The time between when the control position has been changed and the tool reacts.

Lapses: Also called forgetfulness. You know the rules and may or may not have the knowledge, but you simply forget due to your own lapse in memory.

Latent semantic analysis: A way of representing language computationally.

Latent variables: Areas in which the system has the possibility for error or failure.

LCD monitors: Liquid-crystal diode monitor screen.

Learning management system: A system that interacts with educators and students in an educational setting.

Levels of automation: Categories of how much work is done by the machine and how much is done by the human in a sequential ascending order.

Localization: Trying to find the location of something by where it is emitting a sound.

Loiter in place: The term for an unmanned aerial system to fly in a circle or a figure eight at the place that it lost the frequency signal to the ground control station.

Loss-of-signal errors: When an unmanned aircraft system fails to maintain the connection with the ground control station that is controlling it.

Low workload: A task or series of tasks that leave a lot of cognitive resources left for other tasks. This could be because of automaticity or because the task does not need a lot of resources.

Manner: A maxim of communication.

Maslow's hierarchy of needs: Suggests that we have needs but that a single behavior can satisfy more than one need.

Maxims of communication: H. P. Grice's idea that people follow culturally agreed-upon guidelines in order to communicate with each other (https://en.wikipedia.org/wiki/Cooperative_principle #Grice's_maxims).

Maximum take-off weight (MTOW): The maximum amount that an unmanned aircraft system can lift when taking off from the ground. This is the stage of flight that is most prone to collapse if the take-off weight is more than the engines can support. This also defines the different classes of unmanned aircraft systems from a Group 1 up to a Group 4/5.

McGurk effect: When the motion of a person's lips is needed in order to translate the sound that that person is saying.

Mental models: A person's understanding of how something functions.

Mental model update process: When a person updates their understanding of how something works or occurs given new information.

Mental model updates: A better understanding of the situation given new information.

Mental representations in maps: A person's understanding of how close or far things are according to how they are represented in a physical space such as on a map.

Mental rotation: Is the construct that describes how a person translates what they are seeing into a useful spatial thought or cognition.

Mental rotation task: A task where people match a target shape to the corresponding shape rotated in different ways.

Metacognition: The understanding of what is known and not known in one's own knowledge.

Minimally viable product (MVP): A product or software that has the bare minimum of features and functionality.

Misinformation: Information that is misleading or inaccurate. There is purposeful misinformation from people who spread it unknowingly and intentional misinformation from people who spread it knowingly in order to fulfill a hidden goal or agenda.

Mistakes: Happen when the person misinterprets or misdiagnoses that state of the system or misunderstands what they are supposed to do in response.

Mobile data terminal: A technology terminal that connects to central dispatch. More information can be found here: https://en.wikipedia.org/wiki/Mobile_data_terminal.

Mock-up: A fake simulation of the interface. This could be with paper cutouts or with a series of linked pdfs or HTML pages produced by an interface design program such as Invision or Adobe XD.

Models of human performance: Rule-based, computational, and qualitative models of how humans typically behave in different contexts, situations, and given circumstances.

Moore's law: States that technology will increase exponentially.

MRI (magnetic resonance imaging) unit: A large piece of medical equipment that creates an image of a person's internal organs by spinning a large magnet around a cylinder. More information can be found here: https://en.wikipedia.org/wiki/Magnetic_resonance_imaging.

Multiple resource theory: The idea that when workload exceeds capacity, performance will deteriorate.

Multitasking: This is when a person tries to attend to more than one item at a time.

Naturalistic decision-making: The view that people who are experts in a domain decide differently in their expert domain based on pattern matching and other variables gained through their expertise.

Nondisclosure agreement (NDA): An agreement that a person may sign at the start of the development process in which they agree not to disclose information about the product or application they are working on. The agreement can be enforceable in court if it is found that the person shared information about the product or application outside of the team/organization.

Notice to airmen (NOTAM): A written or verbal communication to unmanned and manned aircraft given by the ATM. The NOTAM usually advises on weather conditions, special operations, and recommendations and alerts the airmen to any unmanned aerial activity in the vicinity.

Occupational Safety and Health Administration (OSHA): An organization created to ensure safe working conditions and healthy conditions in the workplace (https://en.wikipedia.org/wiki/Occupational_Safety_and_Health_Administration).

Operant conditioning: Usually, this happens intentionally as one organism (person or animal) rewards or punishes another organism (person or animal) in order to get that organism to behave in a certain way.

Operational context: The work or system environment in which something happens.

Operator: A person who is using an interface or control panel to perform a complex procedure or piece of machinery.

Out-of-the-loop unfamiliarity (OOTLUF): The idea that human skill disappears if it isn't practiced regularly when a machine takes over that procedure. As humans drive less, they will become unfamiliar with driving, and their driving skills will deteriorate. If they need to take the wheel, the OOTLUF will interfere, and they will be less proficient.

Pain points and frustrations: When a person who is using an interface encounters an unwanted interruption in their ability to achieve their goal due to the design of the interface. User testing seeks these interruptions and increases the flow between the user and the application by eliminating these design flaws.

Parallel search: When you are searching for more than one item at a time such as in an airport security scanner.

Participatory design: Where the user contributes to the design of the product.

Pattern library: A file that has all the standard buttons and menus that are used across the organization's projects. Some pattern libraries are open for anyone to use, such as material design pattern libraries by Google, called material.io, which can be found here: https://material.io/design.

Performance-shaping factors: External factors that affect the performance of an air traffic controller.

Perseverate: The tendency of people to follow a course of action that they have used in the past regardless of the previous success or failure.

Phoneme: The smallest meaningful part of a word.

Physiological measure: The person is connected to a machine that measures sweat (galvanic skin response, or GSR), brain wave patterns, heart rate, or any type of physical response to work.

Plain old experiment (POE): Using an experiment with something that varies on two or more levels to measure the change in something else.

Plan view: A frame of reference where the map looks down over the whole scene as if from the sky.

Point of view (POV): When a person looks through one screen and uses only that viewpoint to operate the vehicle or system.

Practice effect: When people become better at something simply through practicing it.

Procedural knowledge: You cannot explain clearly this knowledge; it is for "things that you do."

Processing differences: How quickly or slowly a person processes information.

Processing/search strategies: Ways in which a person tries to search using approaches that worked previously.

Proof of concept (POV): An initial product that shows how the idea works.

Proprioceptive: How people know where their body is in space.

Proprioceptive feedback: A type of haptic feedback that tells a person where their body is in space (https://en.wikipedia.org/wiki/Prospect_theory).

Prototyping with low-fidelity materials: A prototype is a mock design for the user to pretend to use as if it was the real design. When the prototype is created with paper and pencils/markers, it will simulate the actions that the person might choose given specific design options.

Proximity-compatibility principle: Screens that have information that must be used together should be located near each other or grouped together.

Quality: A maxim of communication.

Quantity: A maxim of communication.

Rational actor perspective (RAP): The interaction between technology and goals that assumes that a person will choose the action that gives them the most likelihood of reaching their goals.

Reaction time: The time between a stimulus beginning and when a person reacts to that stimulus.

Real estate problem: As the size of the screen decreases, the amount of room on the screen decreases as well.

Real operators of the system: People who are actually pilots in the plane or nuclear power plant operators who are doing the actual work.

Recommender system: A system that applies an algorithm to a set of data and extracts a future possible ideal choice or set of choices.

Redundancy gain: When an alarm uses more than one sense to alert a person.

Refresh rate: The number of times that the screen redraws the display.

Relation: A maxim of communication.

Relative discrimination: The person must identify when the sound is at a level where they must act.

Resilience: Building a system that has few weaknesses to begin with is resilience engineering.

Return on investment (ROI): The increased sales or savings in person-hours that the organization gains by doing the usability work as the product is developed.

Root-cause techniques: Identifying what contributed to a problem or accident. More information is available here: https://en.wikipedia.org/wiki/Root_cause_analysis.

Salience, effort, expectancy, and value (SEEV) model: A model that predicts when people will attend and when they will not attend to a certain item.

Selective attention: When you change your attention to focus on one area.

Semantic relatedness/neighborhoods: Words that relate to the same concept. For example, the words *fur*, *paw*, *tail*, and *floppy ears* would all relate to the concept of a dog.

Serial search: Looking for items one by one.

Serial self-terminating search: When we put our suitcases through this X-ray machine, a person searches for certain shapes and patterns in our bag.

Servos: An automatic mechanism that uses feedback to change a system state.

Shared team cognition / situation awareness: A shared understanding between team members of the environment, decisions, and information they have as a team.

Shift work: Working in an interval outside of the 9 a.m. to 5 p.m. workday. This might mean that as a nurse, you work from midnight to 8 a.m. Shift work occurs in professions that require continuous expertise be available.

Signal detection theory: A ratio of hits, misses, false alarms, and correct rejections that can measure the sensitivity a person has to an alarm or target. It can also measure bias and criterion.

Situation awareness (SA): A human's understanding of what is currently going on in their environment and what could potentially happen. This includes any system states such as where the planes are located on the air traffic management system.

Skills, rules, knowledge model: A model by J. Rasmussen that categorizes the interaction between the system and the operator as a progression of skills, rules, and knowledge.

Slips: Happen when someone understands but does not execute the right action but fully intends to do so.

Social component: Part of a relationship between a human and robot that refers to their interaction as two beings in relation to how two humans would interact.

Social modeling and learning: Learning how to use something through watching others. For example, most people learn how to use different apps by watching their friends.

Social robots: Robots whose purpose is to fulfill a social role in a human's life such as a companion robot made to listen and talk.

Sociocultural barriers: Cultural ideas and social norms that make communication difficult between people. Cultural ideas that may or may not be true. Social norms that dictate how a person would behave in a social setting but that may not apply to a work setting such as "men initiate the conversation."

Sociocultural factors: Cultural and sociological norms that people follow without thought (e.g., nurses should always respond to every call for help).

Spatial cognition: Where things are located or how they are located in physical space. This can include how a shape is situated or the furniture is placed.

Spatial proximity: How close or far items are from each other or the viewer.

Speed-accuracy trade-off: As a person speeds up in their effort to push a button or select an object, their accuracy decreases and produces more errors.

Stakeholders: Anyone who has an interest in an application or device or the data that it produces. This could be the user of the application, the person who developed it, the organization that funded the development, or the people who benefit from the use of the application. Stakeholders can be primary, secondary, or tertiary based on their access to the product. Primary stakeholders have direct access to the product. Secondary stakeholders have access to the primary stakeholders yet are influenced by the product without direct use. Tertiary stakeholders are influenced by use but may not have any access to the product.

Start-up week: An entrepreneurial event that helps new entrepreneurs get a product started and a business to support it started.

Stimulus-response compatibility or display-control compatibility: The control should be near the tool, and the movement of the control should mimic the change of state of the tool.

Stress: An internal subjective state based on an emotional response to perceived pressure.

Subjective nature: Something that is inherent to that person alone and cannot be assumed to be true of all other people in the same setting.

Subjective self-report: A person answers questions about themselves.

Subject matter expert (SME): A person who is an expert in a given domain or area.

Summative tests: Tests of usability after the application and the interface have been coded and development is nearly finished.

Surveys: In this context, it is a series of questions intended to measure an attribute of the application or product.

Sustained attention: When you continue to attend to a particular item in your environment during a very long task.

Swiss cheese model (SCM): A model of error that suggests that each part of the system is flawed, and when these flaws align, the system fails.

Syntax: The set of rules that govern word order within a language.

Tangible user interface (TUI): Offers features for the visually impaired to create music or interact with a system without the need for vision.

Task: Something that you do on an interface to achieve a goal, such as “Open a document in Microsoft Word.”

Task analysis: When a researcher views how a person would normally do a task either before a technology is designed or after the technology is in place to reveal how difficult or easy different parts of the task are or what the actual steps are in completing the task. This is a methodology that is common in industrial-organizational psychology, where it is used to define a person's role and responsibilities within an organization.

Task switching: Happens when a person has more than one task to perform and they change the primary task to be the secondary task.

Taxonomy: A system of categorization.

Teaming: The act of creating a team or acting in a team manner; groups of people working on the same problem.

Technology acceptance model (TAM): How a person perceives the usefulness and ease of use of a particular system.

Technology immersion: Having all the technology available at once. Giving a person all the possible platforms or ways that they might access the application and letting them choose (e.g., an iPhone, a Chromebook, a Windows laptop, a MacIntosh computer, an interactive large screen).

Teleoperators: Those who remotely control objects or gain access to different types of spaces, such as oceans, interstellar space, airspace, and inside human bodies.

Telerobots: Robots that have sensors that allow them to be aware of their environments with limited mobility and a limited number of actions that they can carry out.

Temporal: Time.

Tethered view: A frame of reference where the map is from a certain corner of the room's point of view.

Three categories of older adults: The three categories of older adults according to the average amount of functioning that they have retained.

3Ds: Dirty, dangerous, and dull tasks.

TRACEr: Taxonomies that list the factors contributing to ATM operators' errors.

Transfer: When the learning in one situation allows a person to respond appropriately to a variety of similar situations.

Triangulation: Using more than one set of empirical measurement or data sets and observing where they agree in order to determine the true nature of the phenomena.

Trust: A sense of confidence that the system or application will do what it is expected to do (https://en.wikipedia.org/wiki/Trust_(social_science)).

Turing test: The Turing test is passed if a machine can converse with human judges so that the humans believe that they are conversing with another human, not a machine.

Underserved populations: People who do not represent the majority. For example, urban farmers would be a group of people who do not represent the majority of farmers in the United States. Usually, the term is used when the idea, product, or service is something that everyone should have equal access to.

Universal design: The idea that any product or software could be used by any person of any age or any ability. The design should have design features to allow people with limited vision, hearing, or physical abilities to use it. It should also accommodate cognitive changes brought on by disability or by age. This includes the young and the very old.

Unmanned aerial vehicles (UAS): Small or large planes that are controlled by a dedicated frequency from a control station on the ground instead of by a pilot in the plane.

Usability testing: A method in which a researcher sits with a user and watches them do key tasks with the interface. The researcher may ask questions or ask the user to "think aloud" during the tasks.

User: Anyone who uses a computer or system interface.

User requirements: This is a document that specifies functions of the application or device to help the user navigate the interface. *See* **functional product specification document**.

Utility theory: A theory that describes how people will weigh different options as they decide on which option to choose based on value. Utility theory suggests that people will choose the option with the most utility (https://en.wikipedia.org/wiki/Utility).

Virtual reality (VR): A way to transport a person to the perception of a different reality that they physically do not exist in.

Visual search: When we look for a relevant target item among other items in an environment.

VR Interaction, immersion, and imagination: The principles of virtual reality that guide the success of full immersion.

Web content accessibility guidelines: A set of guidelines that specify how content is presented on the internet. These guidelines propose to make all internet content available to people of differing abilities and overcome barriers excluding persons whose abilities are not typical.

Working memory: A part of memory that is a temporary storage place where we can combine and manipulate sensory information, knowledge, other memories, emotions, and our perception of the environment to use immediately or store for later use.

Workload: The amount of a cognitive, physical, perceptual, or time-related resource that is consumed by a task or left to do additional work.

World centered: The map with north facing up or at the top of the map.

World Health Organization: The World Health Organization is a special subcommittee of the United Nations that focuses on public health within and between nations of the world (https://en.wikipedia.org/wiki/World_Health_Organization).

Yerkes-Dodson law: The idea that there is an ideal performance curve where stress enhances performance up to a point and then decreases performance. This ideal performance curve differs by individual and task.

You are here (YAH) maps: Maps that designate where the user is located in relation to the immediate environment. Typically, they are large signs or placards on buildings that cannot be moved, but this can refer to virtual maps as well.