**Brian Peacock**

**The Laws and Rules of**

**Ergonomics in Design**

by

**Brian Peacock**

reprinted from

Ergonomics in Design

courtesy of

The Human Factors and Ergonomics Society

Copyright © 1993-2009 Human Factors and Ergonomics Society, Inc.

Reprinted with permission from HFES, P.O. Box 1369, Santa Monica, CA 90406-1369 USA, http://hfes.org.

****

**ISBN** 978-0-615-29682-1

**Table of Contents**

**Page**

Acknowledgements 5

Chapter 1 Tight Targets Take Time, Blind Ones Sometimes Take

a Little Longer: Fitts' Law

(Winter 2001) 6

Chapter 2 The More I Practice, The Better I Get

(Spring 2001) 10

Chapter 3 My Arms Are Getting Tired

(Summer 2001) 15

Chapter 4 Gas Happens

(Fall 2001) 21

Chapter 5 Murphy’s Law: If It Can Happen, It Will

(Winter 2002) 25

Chapter 6 Wrong Number: They Didn’t Listen to Miller

(with Georgina Peacock Goebel, Spring 2002) 30

Chapter 7 JNDs, SDs, HSDs, and DNDs: The Weber-Fechner Law

(with Barrett Caldwell, Summer 2002) 34

Chapter 8 Bias in Human Judgment: Is Your Halo Slipping?

(Fall 2002) 39

Chapter 9 What Kind of Shape Are You In? Anthropometry

and Appearances

(Winter 2003) 44

Chapter 10 Stand (Sit) Straight Up: The Functional Anatomy

of Posture

(Spring 2003) 49

Chapter 11 Pay Attention

(Fall 2003) 55

Chapter 12 Newtonian Moments

(Winter 2004) 62

Chapter 13 Expectancy and Compatibility

(with Robert Schlegel, Spring 2004) 67

Chapter 14 Hotter Than Houston: Body Temperature Control

Summer 2004) 73

Chapter 15 Rule-Based Ergonomics

(Fall 2004) 79

Chapter 16 Remember Hawthorne

(Winter 2005) 85

Chapter 17 Unwanted Energy – Vibration

(with Geoffrey Chase, Robert Fox and Richard Parker,

Spring 2005) 91

Chapter 18 “Warning: Do Not Use While Sleeping” – The Role

of Facilitators

(with Lila Laux, Fall 2005) 98

Chapter 19 Whose Agenda Is It Anyway? – Robert’s Rules of Order

(Winter 2006) 106

Chapter 20 Boomer, Sooner and Donders

(with Max Fogleman, Summer 2006) 112

Chapter 21 “You’ve Got to Attend to Everything”

(with Gary Northam, Fall 2006) 118

Chapter 22 A Look Back: The Ergo is More Difficult than Nomos

(Spring 2007) 125

Chapter 23 Just a Moment

(Summer 2007) 131

Chapter 24 Ethics and Ergonomics: Customer Satisfaction

(with Gary Northam and Erica Diels, Summer 2008) 136

Chapter 25 Product and Process Evaluation Using the Six Us

(with Marc Resnick) 140

Chapter 26 Time for bed: Shiftwork 151

**Acknowledgements**

These chapters are all reprinted from the Laws and Rules column published in *Ergonomics in Design*, a journal produced by the Human Factors and Ergonomics Society. While this Society retains copyright for the articles they have kindly given permission for them to be self published is this book form. Special thanks are given to the Society and particularly to Lois Smith, the communications manager, and her assistants who have contributed considerable editorial assistance as these articles were first published.

The author also expresses gratitude to those outstanding pioneers of our profession who conceived and validated these powerful general laws and rules that are widely applicable in the practice of our profession today. They include Paul Fitts, J R de Jong, Walter Rohmert, Jacques Charles and Robert Boyle, George A. Miller, Captain Edward A Murphy, Ernst Heinrich Weber and Gustav Theodor Fechner, Isaac Newton, Henry A. Landsberger (the Hawthorne effect) and General Henry M Robert. A number of the articles were co-authored with professional colleagues including Bob Schlegel, Lila Laux, Max Fogleman, Robert Fox, Geoffrey Chase, Richard Parker, Gary Northam and Erica Diels. Their experience and insight are greatly appreciated. Other, less formal, contributors to these articles include many colleagues in the Ergonomics and Human Factors profession whose interest in the application of our knowledge to make the world a better place has always inspired me.

Finally, I would like to thank Serena Lackey, a brilliant student, meteorologist, pilot and runner, whose organizational skills and help actually brought this project to fruition.

**Chapter 1**

****

**Tight Targets Take Time, Blind Ones Sometimes Take a Little Longer**

**Fitts' Law**

**(Winter 2001)**

**Origins and Development of Fitts' Law**

In 1954, Paul Fitts described the relationship between hand movement time (MT) and the conditions of amplitude (A) and target width (W):



where *a* and *b* are empirical constants that account for other aspects of the task. Fitts' law is easily demonstrated. Every student of human factors/ergonomics goes through the exercise of drawing two sets of parallel lines on a piece of paper and recording the time taken to make 10 dots, counting the number of dots made in 10 seconds, or some variation on this theme. The results of a controlled investigation in which the amplitude and target width are varied clearly demonstrate Fitts' law.

Since Fitts' (1954) demonstration of the law, many researchers (e.g., Drury, 1975; Graesser, 1993; Hoffman, 1996; Poulton, 1958; Stelmach, 1993; Welford, 1968) have explored modifications, extensions, and applications. They dealt with the varying conditions of the task, either in well-controlled laboratory settings or with reference to human performance in applied settings, such as driving and human gait. Of particular interest have been the characteristics of the object being moved and the target itself. For example, a hammer is usually bigger than the nail at which it is aimed, whereas (one hopes) the gap in traffic is wider than the car that is aiming at it.

One complication of empirical studies is that the designed target is not necessarily the de facto target – some participants may be over precise, whereas others may be too casual. Training is another confounding factor. With highly trained people or very easy tasks, the target may become less of a challenge than the sheer mechanical constraints of the amplitude of movement.

The law also breaks down when the amplitude is very large and a constant velocity dictates the movement time.

The adventurous instructor can find many examples in everyday life. For example, arithmetic competitions on conventional and micro keyboards demonstrate some shortcomings of the trend toward miniaturization. For young parents, babies' toys that require assembly make good instructional media. How big are the buttons on your car's dashboard? Recently, the question of how big the diameter of an emergency egress hatch on the International Space Station should be has drawn attention from both anthropometrists and human factors engineers interested in performance.

**Empirical Confirmation**

In 1966, I did an undergraduate student project on the subject. I fastened 2-inch-high by .25-inchdiameter pegs on a board at 4 and 8 inches apart. I also turned out washers on a lathe to create different inside diameters. Like the law, this equipment was very robust. I carried it around the dorm rooms to find my 80 subjects. Each person performed 10 sets of 10 washer transfers with various target and amplitude differences. The experiment also addressed the training effect and visually controlled and blind transfers in a counterbalanced design.

Analysis of variance and regression were carried out with the help of a simple hand calculator (which didn't even have a squaring function!).

The results confirmed Fitts' predictions. I also observed that when "blind" followed "blind," performance was quicker than when "blind" followed "visual"- I discovered then why I am still no good at touch-typing. Finally, I learned that people get better with practice.

Does Fitts' law apply to blind movements? Yes, when they rely on (tactokinesthetic) feedback for movement control. But ballistic movements are different. A target is selected, then the hand or foot is accelerated and is stopped either when it reaches the target or when feedback control takes over. So perhaps Fitts' law does not apply to touch-typing, normal walking, or handwriting.What about swinging a golf club? The actual target size does not change, but what is the relationship between movement time - the speed of the down swing - and accuracy? The application of Fitts' law to ball games could be the subject of a thousand theses, with large rewards if theory could be successfully applied to practice.

**Case Studies from Manufacturing**

In the 1970s, I had my first opportunity to apply what I had learned while working as a consultant for a factory in Hong Kong that made clothes and accessories for Barbie dolls. The operators had to place a dress, hat, shoes, and a bag on closely defined areas of a card that was then covered with clear plastic. Managers were very concerned with "quality" - the items had to be arranged neatly, but we discovered that some latitude was acceptable. So we increased the target size and doubled productivity. Management was happy, my daughters didn't complain about the quality, and I was ecstatic about my simple success as a consultant.

Over the past 10 years, my responsibilities have focused on manufacturing ergonomics. any practitioners in this area limit their interests to the prevention of work-related musculoskeletal disorders. However, there are also many opportunities to make productivity gains through the application of Fitts' law. These include the loading of machines, such as presses or welding machines; the assembly of components to make subassemblies; and the insertion of subassemblies to make vehicles. Standard time data address these opportunities to some extent, but often a few seconds creep in here and there because of tight targets.

Like pennies, a few seconds saved can add up to enormous savings over the lifetime of a vehicle design. Moving a large or heavy object toward a target is more difficult than moving a small or lightweight one. Consequently, time allowances based on movement amplitude and target size must also include allowances for size and weight. Physics may come first, but finesse needs Fitts!

It also turns out that robots behave like people – they obey the law. Manufacturing engineers make big targets for automatic devices whenever possible. Large transfer presses are remarkably precise and fast, but smaller presses are faster; this is only partly attributable to the laws of physics. Mechanical devices to assist operators fight gravity often add inertial challenges to the control task. In this case, compliant end effectors are added to allow the human operator to finesse the large items into the tight spaces. Here again, engineers intuitively seek larger targets. Unfortunately, the competition for space under the hood of a contemporary automobile brings back the tight targets, and often the pathways to the target are convoluted.

**Other Cases**

Fitts’ law is all around us. Have you ever observed the very precise task of a driver applying makeup while ignoring the fact that the signal has turned green? Of course, there are dangers in stretching the application of the law too far – unless it is a very robust law, like gravity or Fitts’ law. Speaking of gravity, why not try docking the space shuttle to the International Space Station – it’s almost as difficult as backing an 18-wheeler between two buildings!

If you are designing an interactive Web page, you will no doubt design buttons for the point-and-click process of selection. This activity is controlled by such input devices as a mouse, touch pad, or trackball. In more recent forms of Internet terminals, remote pointing devices are used from the comfort of your armchair. Again, Fitts’ law applies: The time taken to navigate a Web site can be shown to be affected by target (button) size, as wells as information content. Perhaps elections for future presidents will be conducted using the same media.

**Rules**

Human motor skills and Fitts’ law have proven relevant in games, manufacturing and assembly, consumer product design, and Internet interface design. But does the law simply describe performance, or are there rules that can be derived to assist designers? One possible rule would be to suggest that the ratio of object to target size should be less than 0.9, 0.75 or 0.5. Another would be that the subtended angle from the starting point to the target should be greater than, say, 5°, 10° or 20°. Perhaps buttons on a Web page should be at least 0.25, 0.5, or 1.0 inches in diameter.

Remember that people improve with practice; thus, derived rules might change. But people vary and grow old and slow, so designing one’s own target might be the ultimate customer satisfier. We could design targets based on performance rather than spatial factors such as anthropometry, with the criterion being measured in seconds per transaction.

**Conclusions**

Fitts’ law is robust and intuitive. It stands up to empirical confirmation in many situations. It is easy to demonstrate in the classroom and provides important lessons for designers.

**References**

Drury, C.G. (1975). Application of Fitts’ law to foot pedal design. *Human Factors*, 17, 368-373.

Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*. 47, 381-391.

Grasser, A.C., & Marks, W. (1993). Models that simulate driver performance with hand controls. In J.B. Peacock & W. Karwowski (Eds.), *Automotive ergonomics* (pp. 383-399). London: Taylor & Francis.

Hoffman, E.R. (1995). Effective target tolerance in an invented Fitts task. *Ergonomics,* 38, 828 836.

Poulton, E.C. (1958). Measuring the order of difficulty of visual motor tasks. *Ergonomics*, 1, 234 239.

Stelmach, G.E. & Nahom, A. (1993). The effect of age on driving skill cognitive motor capabilities. In J.B. Peacock & W. Karwowski (Eds.), *Automotive ergonomics* (pp. 219 236). London: Taylor & Francis.

Welford, A.T. (1968). Fundamentals of skill. *In Movement* (chapter V, pp. 137-160). London: Methuen.

**Chapter 2**

**The More I Practice, the Better I Get**

**(Spring 2001)**

Human factors engineers know all about the differences between people and advise their design colleagues accordingly. But we often fail to address the sometimes greater amount of variability that occurs within individuals due to learning or aging. This article addresses some of the general theories of these changes and goes on to identify some design opportunities.

**General Background**

In 1957, Dutch psychologist J. R. de Jong reported on a study of the motor skills of people involved in rolling cigars. He observed that they got faster with practice and that this improvement was detectable after millions of cycles over many years. He described the familiar logarithmic learning curve. Not only did the cigar rollers get faster, they also became less variable in their times. Other students of learning in general, and motor learning in particular, observed kinks or plateaus in the learning curves of individuals, which indicated that there were quantum jumps in performance following periods of apparently no change. Changes occur following "internal re-organization" as well as overt practice.

Thorndike (1931) was one of the earliest psychologists to study learning formally. He described the process of trial-and-error learning, "the law of effect" or positive reinforcement, that indicated the importance of feedback as laboratory animals and people acquired knowledge and skills. Annett (1969) also described in detail the role of feedback in the modulation of human behavior, although students of the maturation process suggest that some changes occur without the benefit of overt practice.

Learning curves have also been seen to occur in large organizations such as found in the aircraft (Hirschmann,1964) and automobile manufacturing (personal experience) arenas. "When a new model is being launched, it takes time for all the kinks of machine setup, materials delivery, operator skills, and line balance to be worked out. Another characteristic of automobile assembly and similar operations is that operators may have to "unlearn" the procedures that were ingrained over previous years on similar tasks. Learning builds on previous knowledge and skills, and transfer of training is greater when components of the new task are similar to those of the old task.

Perhaps there are other forces at work - the starting age or choice of parents! Competitive gymnasts sometimes become too old at 20, and professional basketball players retire at 30. There comes a time when the ability to learn as indicated by performance improvement appears to be overtaken by the deterioration of key bodily functions associated with aging and disuse. This is also borne out by road accident statistics. As teenagers, we learn how to drive fast and close to the vehicle in front and fine-tune our reaction skills, until we hit a slick spot; we slow down a bit in our 30s and 40s, with occasional middle-age-crisis lapses with Corvettes. Usually, learned strategic behavior keeps us out of trouble until our 60s and 70s, when nature takes over and the accident rate curve begins to shoot up again.

**Personal Background**

Like most of you, my first experience of systematic learning was the rote repetition of my times tables. It was well understood that repetition, coupled with feedback, brought improvement in the accuracy of verbal learning. Around the same period, I had to learn poems and passages from the Bible. Again, practice was supposed to make perfect, but I was never sufficiently keen on practice, and the negative feedback from parents and teachers didn't always work.

Sometime later in my career (I will skip the Shakespeare and Wordsworth period), I was required to learn Morse code. A is di-dah, B is da-di-di-dit, and so on. Part learning was the preferred method. We learned five letters until we made no or few mistakes, then another five. Then we put the two groups together, and so on, until we had learned the alphabet, the numbers, and a variety of other signals, including "short numbers," which were differentiated from letters by the context. It took us about 12 weeks to get to the lightning-fast speed of 25 words per minute. Another observation of this learning period was that, with practice, we could distinguish the Morse signals within a noisy environment. Learning had resulted in robust skills. Bryan and Harter demonstrated the learning curve associated with telegraphy in 1899!

I belong to a group of expatriate British-trained ergonomists and have discussed with them the challenges of going home and driving on the wrong side of the road. The general opinion is that it does not take very long to revert to the old contextual cues so one can avoid testing the crashworthiness of one's rental car. These rental cars have the steering wheel on the right, and the gear shifter requires the use of the left hand, but, fortunately, the accelerator is to the right of the brake. The hand controls generally conform to a very loose international convention for location and direction of motion.

The greatest difficulties occur when leaving a one-way street or negotiating a roundabout. I recommend that learning theorists visit the "magic roundabout" in Hemel Hempstead. This traffic control device consists of eight mini-roundabouts and about six exits. The simple rules are to give way to traffic on your right and keep your eyes open. It is amazing to observe the adept commuters dancing through this junction. When a fender-bender occurs, it is simply treated as another mini-roundabout and life goes on. I'm glad I don't have learning challenges like this one every day!

**Learning and Design**

The purpose of this short article is not to discuss the intricacies of human learning theory, but rather to indicate why human factors engineers must recognize learning changes whenever they evaluate a human-machine interface. There are entire college courses on learning theory, and the education and training industries have been applying these concepts for years. Rats

have been learning their way around mazes, actors have been learning their lines and movements, piano players have learned their scales, and rock and roll stars have been improving their skills, often without the ability to read music. Even babies learn how to pile blocks on top of one another and how to read the pictures in books – the words are for the parents. We provide "crutches" and job aids for older people and hope that they do not develop a dependency to the detriment of the real learning challenge.

The challenge to the human factors/ ergonomics engineer is to design systems to match the capabilities and expectations of the users. We often recognize the normal distribution that

describes the variability among members of the user population, but the recognition of within-individual variability attributable to learning is not often addressed explicitly in design. Designers of almost any human-machine interface and the human fac*tors* evaluators of that interface must realize that with use of the system, learning will occur. Thus, the designer should accommodate learning and, if necessary, anticipate the need for unlearning.

A familiar learning situation encountered by HF/E engineers is the stereotype response in operating controls. McFarlane and Wierwille (1990) pointed out that there are strong and weak stereotypes, as indicated by the relative probabilities of a particular response. For example, we have learned to recognize that movement of a control in a clockwise direction causes an increase in volume, but what about turning on a water faucet? Another conflicting stereotype is found when one visits the United Kingdom, where down turns the light *on.* Sometimes traffic lights are arranged vertically and sometimes horizontally (thank goodness for the redundant color codes!). I remember a traffic island in Canada where the red and green lights were both permanently on and a sign, some 50 yards earlier, explained that this meant that traffic was supposed to merge. So stereotype responses are learned, and their strength is related to the amount of repetition. Where conflicts occur, wrong responses may result in serious situations.

The context of a human response is also important. For example, the arrangement of numbers on a telephone keypad is different from that of a calculator; the conflicts occur only when the two functions coincide – an increasing possibility in the recent convergence of computation and communication technologies. In the computing world, we are familiar with the arrangement of the menu bar across the top of our screens. In the increasingly complex and rapidly expanding world of graphics and database computing, there are many opportunities for standardization in interfaces, but such trends appear to be unlikely. Designers of interfaces assume that users will learn or unlearn as the case may be. A key to the success of common office software is that it is usable by people with a wide variety of knowledge. Beginners may be satisfied with opening, closing, and saving a file and a few editing commands. More experienced users move further across and further down the menus. Eventually, some users learn that the right mouse button opens a whole new world of possibilities. The key to the enormous success of this software is that it does indeed cater to users on various levels of the learning ladder.

So how do we design for the accelerated learning of children or deterioration that accompanies the aging process ("geriating")? For normal people, learning is achieved by presenting a task, having the individual carry out the task, and then providing appropriate feedback - rewards or punishment. In some cases, we provide crutches or facilitators during the learning period. Small children use training wheels as they learn to ride bicycles - to avoid negative feedback! Musicians use sheet music until they've mastered the piece. Students use textbooks until they have gotten their "license" to practice their chosen profession on an unsuspecting public. Car, computer, and household appliance users use their user manuals only as a last resort. Parents, faced with the dreaded "some assembly required or send $10 if you want us to do it in a tenth of the time it will take you because you are too arrogant to use the instructions" statement, often learn by their mistakes.

The annual rite of completing and filing tax returns is another opportunity for learning theorists. More than 50% of tax returns contain errors, and tax advisers complete about 50% of returns. The key learning challenge is that we get to rehearse this complex task only once a year. Research into tax form and instruction design indicates that both time and errors could be reduced through the application of traditional human factors principles to tax form and instruction design, for both novice and professional tax preparers (Theisen, Peacock, & Peacock, 1991).

A recent paper nu Roesseger (2001) emphasizes the importance of “the cognitive interface.” Learning of complex tasks can be aided to a limited extent by attention to the cosmetics of the controls and displays. The key to good design is more strongly related to the understanding of the system’s functionality. The tax and driver instruction manuals go only so far. It behooves the designer and the human factors engineer to address the learning and understanding of the users of how the system works. *Practice makes perfect, old habits die hard, you can’t teach an old dog new tricks, it’s the cognitive interface, dummy!*

**References**

Annett, J. (1969). Feedback and human behavior. In *Motor tasks* (pp. 38-53). New York: Penguin.

Bryan, W., & Harter, N. (1897). Studies in the physiology and psychology of telegraphic language. Psychological Review 4, 27-53. Cited in Singer, R.N. (1968). *Motor learning and human performance*. New York: Macmillan.

DeJong, J.R. (1957). The effect of increasing skill on cycle time and its consequences for time standards. *Ergonomics* 1, 51-60.

Hirschmann, W.B. (1964, January-February). Profit from the learning. *Harvard Business Review*, No. 64107, 105-119.

Roesseger, P. (2001, March). *Applied ergonomics of driver information systems and car PCs*. Presented at the Applied Ergonomics Conference, Orlando, FL.

Theisen, B., Peacock, B., & Peacock, E. (1991, March). *Tax-form redesign*. Presented at the Annual Conference of the American Accounting Association, Nashville, TN.

Thorndike, E. L. (1931). *Human learning*. New York: Apple-Century-Crofts.

Wierwille, W.W. (1990). *Study of direction of motion stereotypes for automobile controls* (ISE Report No. 90-02, work sponsored by the Society of Automobile Engineers).

**Chapter 3**

**My Arms Are Getting Tired**

**(Summer 2001)**

This article is about local and general fatigue - their causes, nature, empirical effects, and possible countermeasures.

**Work, Wheels, and Variation**

A long time ago, before the invention of the wheel, I used to carry my heavy suitcases from the far side of the airport parking lot to the check-in counter. Sometimes I planned ahead and brought equally balanced bags, but more often I had an overweight suitcase in one hand and an overweight briefcase in the other. The sharp pain in my fingers and the dull aches in my arms and shoulders signaled me to change with increasing frequency as I approached my destination. I had just discovered local muscle fatigue. I also intuitively understood that I could initiate recovery simply by alternating hands, if the loads were within my capacity.

On my return home, I found that I had to take out the trash and paint the ceilings. I discovered that the center of gravity of the trashcans was a long way from mine and that the handles on the recycling bin filled with unread newspapers were sharper than the average bread knife. The painting job started with a flourish, until I discovered the double whammy as gravitational resistance to circulation compounded muscle fatigue. Luckily I was able to solve this one by adding an extension onto the paint roller handle. It's amazing how much physiology and ergonomics we can learn by just doing the job!

My experience may have led me to realize that if I were in better shape, I would be able to withstand these everyday stresses. I enrolled for a course of conditioning and weight lifting at the local fitness center, as advertised on TV. My personal trainer instructed me to alternate pushing and pulling and to rotate exercises among arms and shoulders, trunks and legs. She had me push 80% of my maximum voluntary capability for 10 repetitions and sometimes gave me encouragement. (She had probably read the article, "Feedback and Maximum Voluntary Contraction" by Peacock, Westers, Walsh, & Nicholson, 1981.)

As I got to the eighth or ninth repetition, my arms began to shake, and my attentive instructor reminded me of the importance of someone "spotting" when I approached my limits. As I let out a cry of joy when number 10 was pressed, she stressed the importance of breathing out as I pushed and suggested that it would be a good time to have my blood pressure checked. After six months of dedicated training, I realized that I could face the world's biomechanical and physiological challenges with aplomb. I had discovered by looking around the gym that there really was a difference between people and that individuals changed with training.

**The Literature**

A visit to the library explained why medical students are so strong: they have to carry around thick textbooks all day. They're lucky - their work keeps them in shape, and they also discover first hand the importance of sleep. One such book - Best and Taylor's *Physiological Basis of Medical Practice* (West, 1985) – explains clearly how this fatigue thing works, assuming one had paid attention in high school and college chemistry classes. On page 92 it describes the Citric Acid (Kreb's) Cycle and then spends the next thousand or more pages telling about all the other factors than allow people to operate as load-carrying agents.

My empirical instincts led me to Basmajian's *Muscles Alive* (1978). This text deals with the electromyographic approach to the study of muscle activity in general and fatigue in particular (pp. 83-91). My next stop was a little more specific and quite a bit thinner - I discovered Astrand and Rodahl's *Textbook of Work Physiology* (1977). It even has a chapter (13) called "Applied Work Physiology" and a couple of pages (472-473) on general and local muscular fatigue.

I then moved to the classic book collection and discovered a 70-page text entitled *Muscular Work, Fatigue and Recovery* written by G. P. Crowden in 1932! On page 11 are some exponentially decaying curves that represent people's capacity for doing work over time as a function of load. These curves were redrawn from a 1919 article by R. A. Spaeth in the *Journal of Industrial Hygiene.*

Perhaps the most comprehensive treatment of the subject is contained in the collection of papers edited by Hashimoto, Kogi, and Grandjean, *Methodology in Human Fatigue Assessment* (1975). The classic Kodak (1986) books, largely written by Suzanne Rogers, also provide an outstanding exposition of work and fatigue. Another notable fact is that the citations in this article come from the United States, the United Kingdom, Sweden, France, Germany, Canada, and Japan - fatigue is indeed a global phenomenon.

A notable product of the German work physiology tradition is Walter Rohmert. In 1973, he quantified the fatigue curves to show the relationship between load and endurance. Some years later, Rose (1992) developed more conservative curves that also addressed the additional challenges of overhead work and later (Rose, Ericson, & Ortengren, 2000) addressed my first problem of fatigue and elbow loading. Also in the 1980s, researchers from the Institute of Ergonomics at the University of Michigan articulated the remarkably versatile "one-third rule," which has been adapted and maladapted throughout the ergonomics furor of the 1990s. Simply stated, this rule says that one can work all day at a level of effort equal to one-third of one's maximum. An extension of the rule suggests that one can work overhead for one-third of the one-minute job cycles typical of automobile assembly and maintain this work-rest ratio for the duration of the work shift. These useful rules were immortalized in the classic tripartite ergonomics agreements between the United Auto Workers, the U.S. Occupational Safety and Health Administration, and the Big Three automobile companies.

These events reminded me of my first lecture in ergonomics in 1964 by the late Ernest Hamley at Loughborough University. It was entitled "How hard should a man work?" There are jobs, hard jobs, and jobs your father used to do.

**Rules and Tools for Physical Work**

I still don't know the answer to Dr. Hamley's question, but I have a lot of information and analysis tools that can lead me in the right direction. Unfortunately, the question is clouded by individual differences, the nature of the work, the political environment, and the influence

of incentives and motivation. I have discovered that professional athletes are prepared to work very hard to achieve goals that lead to large rewards and that amateur athletes, often with less talent, sometimes work just as hard to achieve less public goals. Training for and running marathons brings a wide range of people face to face with fatigue during both training and the races. The chart below shows the fatigue-related curve of one amateur in a recent marathon.

The slow pace during Mile 1 was caused by congestion, and those at Miles 5 and 13 to calls of nature. But the clear decline over the second half of the race was simply physical fatigue as the input effort didn't change. The spurt during Mile 26 indicates the possibility of extra effort when faced with a qualifying time goal that was achieved with one secondto spare! This curve is typical of the majority of marathon runners, although sound strategy and talent occasionally result in so-called negative splits: faster second halves than first halves. Although these results clearly indicate the phenomenon of fatigue, amateur athletes can take a rest - they don't have to run a marathon every day or even every week. On the other hand, manual workers are expected to turn up for work every day in a rested state.



Early industrial engineering approaches to the estimation of job loading used either predetermined time data or stopwatch studies to establish "an experienced worker standard," which eliminated training effects. Fatigue allowances of 10% to 20% (Konz & Johnson, 2000) were added to establish a level of activity that would ensure that a rested worker would show up for work the next day. Recent activities within the ergonomics community have used psychophysical techniques (e.g., Snook & eirello, 1991) and the consensus of experts based on the scientific literature (Waters, Putz-Anderson, Garg, & Fine, 1993). The on again, off again OSHA ergonomics standard has generated a plethora of work measurement tools aimed at the prevention of work-related musculoskeletal disorders and unacceptable levels of muscle fatigue, although there is some debate regarding the precise relationship between these two consequences of physical activity. Unfortunately, these and many similar contemporary techniques have been subject to intense political debate because they not only attempt to address the health of some workers but do so by addressing both the physical elements of the tasks and work (time) standards. A fair day's work may be measurable, with due regard to individual differences, but

a fair day's pay is negotiable and subject to prevailing economic conditions.

**Work Design**

Our dilemma as ergonomists is to find acceptable ways of addressing the questions of fatigue and a fair day's work, given the premise that we expect a healthy, rested worker to show up for work every day. We know that the principles of rotation allow for recovery. We also know that job enlargement and autonomy can also lead, through less formal rotation, to recovery.

Unfortunately, such strategies often conflict with seniority-based approaches to the choice of jobs - a policy that is widely perceived as being fair but that falls down physiologically when the whole cohort of operators reach their fifties.

Finally, we know that there *are* counteractive training and aging effects as well as individual variability to take into account. Even though we have hard scientific evidence of all these phenomena and many useful measurement tools, it appears unlikely that we will ever be able to apply the techniques to address the problem of tired arms without a formal policy/political overlay. Meanwhile, wheels allow us to dash blythely across airport parking lots only to heft a too-heavy suitcase into the overhead bins, professional athletes work very hard to earn and honest dollar, and amateur athletes continue to stretch the limits provided by their parents in pursuit of a few moments of glory and reduction of the inevitable decline associated with aging.

**References**

Astrand, P.O., & Rodahl, K. (1977). Textbook of work physiology. New York: McGraw-Hill.

Basmajian, J.V. (1978). Muscles alive. Baltimore: Williams and Wilkins.

Crowden, G.P. (1932). Muscular work, fatigue and recovery. London: Sir Isaac Pitman and Sons.

Hashimoto, K., Kogi, K., & Grandjean, E. (1975). Methodology in human fatigue assessment. London: Taylor & Francis.

Kodak Ergonomics Group. (1986). Work-rest cycles. In S. Rodgers (Ed.), *Ergonomics design for people at work* (vol. 2, chap. 11). New York: Van Nostrand Reinhold.

Konz, S., & Johnson, S. (2000). *Work design*. Scottsdale, AZ: Holcomb Hathaway.

Peacock, B., Westers, T., Walsh, S., & Nicholoson, K. (1981). Feedback and maximum voluntary contraction. *Ergonomics*, 24, 223-228.

Rohmert, W. (1973). Problems in determining rest allowances. Applied Ergonomics, 4(2), 91-95.

Rose, L. (1992). *Ergo-index – Development of a model to estimate physical load, pause need and production time in different working situations*. Licentiate thesis report R-018, Department of Injury Prevention, Chalmers University of Technology, Goteborg, Sweden.

Rose, L., Ericson, M., & Ortengren, R. (2000). Endurance time, pain and resumption in passive loading of the elbow joint. *Ergonomics*, 43, 403-420.

Snooks, S.H., & Cirello, V.M. (1991). The design of manual handling tasks: Revised tables of maximum acceptable weights and forces. *Ergonomics*, 34, 1197-1214.

Waters, T., Putz-Anderson, V., Garg, A., & Fine, L. (1993). Revised NIOSH equation for the design and evaluation of manual lifting tasks. Ergonomics, 36, 749-776.

West, J.B. (Ed.). (1985). Best and Taylor’s physiological basis of medical practice (11th ed.). Baltimore: Williams and Wilkins.

**Chapter 4**

**Gas Happens**

**(Fall 2001)**

*Boyle's law:* The volume of a given mass of gas, at constant temperature, is inversely

proportional to its pressure.

*Dalton's law:* The pressure of a mixture of gases is equal to the sum of the partial pressures

of its constituents.

*Charles's law (Gay-Lussac's law):* The volume of a given mass of gas, at constant pressure,

increases by 1/273 of its value at 0° C, for every degree Centigrade rise in temperature.

*Henry's law:* The amount of gas dissolved in a solution is directly proportional to the

pressure of that gas over the solution.

*Law of gas diffusion:* Gas molecules will diffuse from an area of higher concentration to

an area of lower concentration until equilibrium is reached.

*Graham's law:* The rate of diffusion of a gas is inversely proportional to the square of its

density.

*Convection and forced convection:* The transfer of heat in a gas or liquid is proportional to temperature difference and can be increased by increasing airflow.

-Nelkon and Parker, 1962

I our early teens, we begin to understand how algebra can be applied to all sorts of things, including gas. Furthermore, these mathematical relationships are easily demonstrated in the laboratory and around the house. Gas and the laws that describe its behavior are both useful and fun. Sometimes when gases misbehave, they can be far from fun - they can be outright dangerous. As ergonomists, we can find many applications of this knowledge.

In this short article, I explore the ramifications of the gas laws in SCUBA diving, spaceflight, oil and gas well drilling, and cooking. I could also explore many others if I had the space: vehicle suspensions, truck brakes, ball games, air travel, powered hand tools, fishing, barbecues, Bhopal, LNG tanker navigation, display design, athletics, rehabilitation, swimming pools, balloons, mountain climbing, anesthetics, weather, sailing, toilets, tornadoes, barometers, global warming, liquid refreshments, central heating, vehicle fuels, greenhouses, chimneys, blowing your hot porridge and cold fingers ...

**Altitude**

The air that surrounds us consists of 78% nitrogen, 21% oxygen, 0.03% carbon dioxide, and 0.97% other gases. We extract about 5% of the oxygen in exchange for some carbon dioxide before breathing it out. The relative partial pressures of oxygen and carbon dioxide between our blood stream and alveoli cause the exchange to take place. The metabolic processes at the other end of the circulation extract oxygen from its carrier - hemoglobin - and dump out the excess of carbon dioxide. Work physiologists have observed these processes for many years and used the information to describe the relationships between gas exchange and physical work. Typically we

use about 1 or 2 liters of oxygen per minute. Unfortunately, things get a bit more complicated when we move away from sea level (Astrand & Rodahl, 1977).

At sea level, the pressure of air is 14.7 pounds per square inch, or 1 atmosphere. This pressure - the combined weight of all the atmospheric gases creating a force on the surface of the earth - will cause a column of mercury in a barometer to rise 29.92 inches (760 mm). At 18,000 feet above sea level, the pressure falls to half an atmosphere, and at 1200 miles, it drops to zero. At just 33 feet below the sea, the pressure is equal to 2 atmospheres.

Humans can breathe quite well up to about 10,000 feet, but between 10,000 and 50,000 feet, we have increasing difficulty and require supplemental oxygen. It also gets very cold. At above 50,000 feet we need to live in a sealed suit or cabin. Most people live below 500 feet elevation, but in Tibet, some people live at altitudes of more than 15,000 feet. Olympic athletes like to train at high altitude because this improves their gas exchange performance in competition at lower levels. Some mountain climbers pride themselves in their ability to go high without supplemental oxygen.

Dalton's law explains why people have breathing difficulties at higher altitudes. The partial pressure of oxygen is reduced - we suffer from hypoxia. Air traffic populates the 20,OOO-40,000-foot region because airplanes at that altitude get better gas mileage and are less affected by winds. Typically, commercial airlines pressurize their cabins at the equivalent of 10,000 feet. The International Space Station orbits at about 240 miles, at 17,500 miles per hour!

Our bodies are full of trapped gas - lungs, ears, sinuses, stomach, intestines, and sometimes beneath the fillings in our teeth. At sea level all is well, but as we move upward, the external pressure drops and the trapped gas tries to expand - Boyle's law. Trapped gas in the middle ear can be particularly painful, especially among children and people with upper respiratory infections. Performing the Valsalva procedure can relieve the pressure pain: Take a breath, close your mouth, pinch your nose, and blow hard. This clears the Eustachian tube (between your middle ear and throat).

Some of you may actually listen when flight attendants explain what to do with those little yellow masks that drop in front of your face from a trap door above your heads. You're flying at 37,000 feet, and suddenly you say to yourself: "I feel funny, I must be getting a dose of *Dalton's law -* or could it be *the law of gaseous diffusion?"* You go on to explain to all who are listening: "I believe that I have an impairment of mental and bodily functions as a result of hypoxia." You explain to the less well informed that the symptoms are tingling in the fingers, blurred vision, dizziness, euphoria, belligerence, headache, apprehension, blurred vision, hot and cold flashes, and gasping for air.

Meanwhile, the people around you have already got their masks on and are sucking down the oxygen while you are pontificating. What happened? The cabin lost pressure - or, rather, the air inside assumed the same pressure as the air outside. This led to a reduction in the partial pressure of oxygen to a level below that in the blood flowing through your lungs. You could even have oxygen going the wrong way! Golf enthusiasts may recall that Payne Stewart died after a spectacular uncontrolled flight across the country, probably caused by hypoxia.

Henry's law is insidious - ask any SCUBA diver. If we move rapidly from a high-pressure to a low-pressure environment, the dissolved nitrogen comes out of solution and forms bubbles in the joints (the bends, or Caisson disease), under the skin (skin bends), in the lungs (the chokes), and in the central nervous system. All these illnesses are characterized by pain, sometimes extreme, and loss of function. When the condition affects the lungs, it is sometimes mistaken for a heart attack, which increases the level of apprehension. Central nervous system effects include headache, delirium, and a loss of the senses, similar to being drunk. The simple answer is prevention: Avoid rapid decompression - denitrogenate slowly. Another approach is to prebreathe 100% oxygen to flush nitrogen from the body. The treatment is recompression in a

hyperbaric (high-pressure) chamber.

Astronauts performing extravehicular activity (EVA) wear a suit that is pressurized at about 4.3 psi. The suit has to provide thermal, radiation, and mechanical protection; provide oxygen; and allow the astronaut to move sufficiently well to do his or her job. The atmosphere inside the space shuttle or International Space Station is at 14.7 psi, with 21% oxygen - the same as sea level. One reason the pressure in the suit is low is that with zero pressure outside, a high-differential pressure would prevent movement of the joints. Problems are associated with the transitions between the low- and highpressure environments.

The simple approach to preparation for EVA is to prebreathe 100% oxygen for three or four hours prior to the reduction of pressure in the suit. Unfortunately, this long preparation time has unwanted operational difficulties, so various protocols have been developed, which include a reduction in the cabin pressure to 10.2 psi and an increase in the partial pressure of oxygen for 12 hours prior to the EVA. Understanding the gas laws is key to the survival of humans in space.

**Energetic Gas**

One of my more spectacular jobs was with a gas and oil well instrumentation company developing the computer displays to be used on the drilling rigs. That was in the early 1980s, when computer graphics was in its infancy. However, there was sufficient capability to improve considerably on the simple black and white lists of important drilling variables and their values.

There are various modes of the drilling operation, the main ones being drilling, tripping (pulling out the pipe to add another section), and blowout control. If the drill penetrates a region of gas that is at high pressure, or if the tripping process is carried out too quickly, the result is something like putting a down-turned cup full of air below the water in the bathtub and then tilting it a little. The difference between the bathtub experiment and the drilling rig is simply one of scale. The whole drilling pipe can come flying out. If the gas ignites, the surviving roughnecks will have put the emergency escape facilities to good use. Good (predictive) display design and knowledge of the gas laws can help (Peacock, Schlegel, & Dorman, 1983).

Before the invention of the microwave oven, we had pressure cookers. If you fastened the lid on tight and then heated up the cooker, the gas inside could not expand until the valve let it out. The result was more heat and quicker cooking and if we can believe the advertisements – better flavor and essential nutrients. Stevenson and Watt, who were train enthusiasts, also discovered this gas behavior. They took a cylinder of water, lit a fire under it, and used the escaping steam to drive iron wheels on an iron track.

Much earlier in our history, a certain Guy Fawkes was dissatisfied with the behavior of the British Government, so he tried to apply the relationship between gunpowder and rapidly expanding gas to blow up the houses of Parliament. His failed attempt is still celebrated on November 5. Old Guy learned his tricks from the Chinese, who thousands of years ago intuitively understood the concept of explosive energy.

Gas behavior can be useful, but you've got to know how to manage it. It is the job of human factors engineers to design the controls and displays that facilitate this process. Sometimes the engineers who design these gas-powered machines don't trust us to deliver interfaces that will keep the operator out of trouble. They believe in automation with governors and valves and feedback circuits. They even use computers to do the calculations in case we forgot the stuff we learned in Calculus III and Thermodynamics. But we need not despair; the more automation, the more information the controllers need and the more we are needed to develop these increasingly complex interfaces.

It is left to the inventive human factors engineer to apply his or her knowledge to the control of gas; if we don't do a good job, it will just happen. Here is a take-home quiz:

*Present theoretical and hypothetical arguments regarding the relative effects of cars, cows, and computers on global warming; devise an empirical investigation to evaluate your hypotheses.*

**References**

Astrand, P., & Rodahl, K. (1977). Factors affecting performance. In *Textbook of work physiology* (pp. 628-638). New York: McGraw-Hill.

Nelkon, M., & Parker, P. (1962). *Advanced level physics. L*ondon: Heinemann.

Nicogossian, A. E., & Parker, J. F. (1982). *Space medicine and physiology* (NASA, Scientific and Technical Information Branch). Washington DC: U.S. Government Printing Office.

Peacock, B., Schlegel, R., & Dorman, R. (1983). Dynamic drilling displays. In *Proceedings of*

*the Human Factors Society 27th Annual Meeting* (PP' 445-448). Santa Monica, CA: Human Factors and Ergonomics Society.

**Chapter 5**

**Murphy's Law: If It Can Happen, It Will**

**(Winter 2002)**

This is an article about probability and system reliability, with particular reference to the homeland security issue. The title of this article actually describes Finagle's law of dynamic negatives. The implication is that if there is a possibility of something's happening, however

improbable, then it will happen. Finagle's focus is on rare events, serendipity, and multivariate coincidence.

In contrast, Murphy's law states, "If there are two or more ways to do something

and one of these will result in a catastrophe, then someone will do it." Thus Murphy was more concerned with outcomes than probability. Murphy was also of the opinion that some person will

precipitate the event in question, but he did not indicate whether this human action was attributable to a mistake or to malice.

In recent months we have heard about two catastrophic incidents involving airplane crashes, one clearly arising from malicious intent and the other apparently from error somewhere along the design, manufacturing, maintenance, and operation chain. It appears that Murphy had a hand in both incidents. Readers of this article are directed to the World Wide Web for wide-ranging discussion of Murphy's law and its many extensions.

**After the Event**

In his excellent book *Set Phasers on Stun,* Steven Casey (1992) described a

series of cases in which some design or operational error led to a catastrophic outcome.

But Steve is a Monday morning quarterback: He explains how the situation arose and how the application of appropriate ergonomics interventions in design or operation could have prevented

the unfortunate turn of events. A patient was zapped with a high dose of radiation, a pilot took a plane beyond its operationallimits, a cosmonaut was not strong enough to close a valve, a ship's captain ignored the evidence, and a worker lacked situation awareness while hunting for a rabbit. All the incidents in this book had the dual characteristics of rarity and serious outcomes. The incidents also involved complex combinations of events, unfortunate contexts, human error, and

the lack of a safety net.

Henry Petroski wrote *To Engineer Is Human - The Role of Failure in Successful Design* (1992). Like Casey, he described a series of notable failures, mainly in civil engineering. He argued in his first chapter that "falling down is part of growing up" - failure is an essential part of design success. But where we don't have sufficient safety nets in place, system failure can have unacceptable human and societal costs. Petroski's message is that we must learn from failure; the recent terrorist attacks have heightened our sensitivity to our vulnerabilities. Casey, a human factors specialist, overtly identified the need to apply our knowledge of the shortcomings of humans when we design complex systems. Petroski, an engineer, acknowledged the role of Finagle and Murphy as inevitable but stressed the importance of learning. Winston Churchill was more optimistic; he is quoted as saying, "The further you look back, the further you will be able to look forward." Our current challenge in homeland security is to learn from experience, apply our excellent analysis tools, and design safe and secure systems.

**System Analysis and Design**

Systems engineering classes address the theories of serial and parallel arrangements and the value of redundancy. The electrical examples of this theory tell us that the shortest (or fastest) distance between two points is not necessarily a straight line; rather, it is the path of least resistance. Stafford Beer, an icon in operations research history, posed the question of how to prevent 11 players from moving a football across a goal line. His solution was arrived at not so much through detailed system analysis and design as through placing another 11 intelligent and versatile players on defense. Soccer is a game that adheres to Beer's advice: It relies on flexibility and mobility for both defense and offense to provide the appropriate redundancy. American football, on the other hand, is much more analytic and applies the theories of parallel and serial systems to win the game.

But what about the "end around" and "Hail Mary"? The former strategy was invented by Murphy and is defended by having versatile "safeties" who have enough time to inspect the developing scenario and position themselves to prevent a catastrophic outcome. Conversely, the Hail Mary, a conception of Finagle's, says, "If you throw a ball high enough and far enough, one of these days it will come down in the right place." Soccer, football, and terrorist strategists have one thing in common with Murphy - they exploit the mismatches between offenders and defenders.

**Reliability Theory**

Mechanical systems are known to fail because of either overload or wear. Complex mechanical systems are like humans - they exhibit variability in their response to sudden or cumulative stress. Reliability theorists demonstrate the use of the Weibull distribution in describing burn-in, steady-state, and wear-out failures: the so-called bathtub curve. The evidence from such an analysis is used to drive preventive maintenance schemes. Such schemes are deliberately designed to predict and thwart the surprises offered by Finagle and Murphy. But even the best-laid plans are not immune to unusual circumstances. The Weibull distribution describes the probabilistic nature of failures commonly encountered in such things as vehicle fleets, but it has widespread applicability in other complex systems such as security. Oliver Wendell Homes, whose poem, "The Deacon's Masterpiece," is reprinted in Petroski's book (1992), parodied reliability theory. The deacon built a one-horse shay in which all the components failed at the same time, precisely as designed. The World Trade Center collapse was an eerie structural parallel.

The Weibull distribution can also be used to describe human failures. At the gross level, the bathtub curve describes the trial-and-error learning behaviors of children and the increasing frequency of failures among older people as one system after another deteriorates. It may also explain the human behavior of inspectors - early failures caused by lack of experience, steady-state performance, and, finally, inattention from increasing complacency. The empirical testing of prototype security strategies since September 11 clearly demonstrates a learning curve. We are now becoming accustomed to the process, and the lines are moving more swiftly. But we all know that the absence of calibration signals can lead to complacency, which, if not addressed,

will lead to system decay. Both Finagle and Murphy were students of reliability - they understood the confluence of rare events and human fallibility.

**Signal Detection Theory**

If you are going to stop terrorists, first you have to detect them. Signal detection theory is a powerful concept describing the behavior of people and automated systems in their efforts to

detect anomalies in the context of system «noise." Further, it explains the effect of alternative strategies that optimize the relative probabilities of false positives and false negatives. Signal detection theory says that you can find the needle in the haystack if you are prepared to

pay the cost of taking the stack apart straw by straw. Strategic system design comprehends the opportunities described by signal detection theory through the incorporation of more sensitive and selective equipment and processes. Metal detectors, X-ray equipment, chemical residue sensors, biological binoculars, and perceptive pets enhance the capabilities and versatilities of increasingly highly trained inspectors and sophisticated strategies.

The human inspector's selection and training are also aimed at enhancing the ability of the human inspection process to discriminate between signal and noise. Human factors/ergonomics (HF/E) experts and managers are quick to point out the fallibility of human inspectors. We know much about human sensory and perceptual variability, attention, vigilance, motivation, fatigue, and complacency. Just the other day a harassed airline check-in clerk said to me, "It doesn't matter about your drivers license - they will check it at security." On another occasion in my earlier career, when managers of a paint shop suggested that an eyesight test would be appropriate for paint inspectors, a union representative responded that choice of job was a matter of seniority, not eyesight. But the human security inspector's job is not easy - just look at the many strange but benign electronics gadgets that passengers like to put in their carry-on bags.

We also know that human signal detection performance can be greatly improved by eliminating a lot of the obvious noise, thus reducing the complexity of the task and potentially increasing the signal rate and the hit rate. Thus racial profiling is a great temptation. After all, the overt issue in the current situation is indeed racial, as it was with the Japanese in World War II. But such efforts are crude; they increase the political fallout of false positives (the Fourteenth Amendment is a powerful concept); and, worse, they ignore Murphy.

Paul Revere's certain detection of the British army led to a successful response, but nowadays terrorists don't wear bright red coats. We need to apply much better methods, such as contemporary datamining techniques, if we are going to find that needle. Homeland security will also depend on the improvement of automatic signal detection methods coupled with the use of HF/E knowledge applied to the selection, training, and maintenance of the perceptual processes of inspectors.

**The Human Factors Challenge**

Unfortunately, a lot of traditional HF/E approaches also ignore Murphy. We revere the normal distribution and the 5th percentile. But when it comes to system reliability, our greatest challenges are

* Einstein, who wants to beat the system
* Rambo, who is impatient
* Granddad, who is growing old
* Timothy, who doesn't like the status quo
* and all the other people who are incompetent, inattentive, inebriated, incapacitated. or simply tired or fatigued.

*Most system failures don't occur to normal people under normal circumstances. They occur to people who are permanently or temporarily susceptible to unusual challenges. Murphy knows that.*

Over the years, human factors/ergonomics has made great strides using systematically designed experiments and drawing conclusions regarding averages, first- and second-order interactions, and statistical significance. We eliminate outliers, for good cause, in our pursuit of science. The constraints of budgets, dissertations, and time often encourage us to ignore high-order interactions as being either unexplainable or spurious. Contemporary HF/E researchers, however, use a different approach: data mining. Often the data are opportunistic, collected for some other or nonspecific reason. The strategy is to use the enormous sort and search power of computers to detect anomalies. Such techniques, if coupled with the intuitive capabilities of people to interpolate and extrapolate, will, one hopes, lead to better signal detection than classical experimentation. Contemporary graphic display technologies can be applied to enhance the differences that occur in large data sets.

**The Normal and Other Distributions**

The normal distribution is an enormously powerful concept that can be brought to bear not only on human and biological variability but also on mechanical and operational systems - and even, as Gauss showed us, the behavior of the heavenly bodies. The tails of the bell curve go off to infinity, so we often focus on the usual rather than the normal.

Fortunately, there are many other statistical distributions that describe more extreme system behavior. One very powerful one is the Poisson distribution that Moroney (1956) used to describe the occurrence of rare events - goals, floods, and the death of soldiers caused by horsekicks. Another is the Gumbel or extreme value distribution, which addresses the probability of records. There is also the Non-Central Chi-Squared, which can be brought to bear on "collateral damage" (Evans, Hastings, & Peacock, 2001).

Terrorism is commonly about rare events that exploit a system's defensive weaknesses. Murphy has showed us that these weaknesses commonly involve aberrant human behavior and performance. The human factors/ergonomics profession should stop looking merely at averages and start paying more attention to the extremes and high-order interactions. Otherwise Steve Casey will have material for another book and Henry Petroski will continue to teach us how we can learn from failure.

**Certification**

In his excellent book, *The Human Factors in Aircraft Accidents (1969),* David Beaty addressed many human factors issues that are widely discussed today, including perception, design, and training. In his appendices are many accidents that were caused by either design or operator failure. A recent book, *Human Factors in Certification* (Wise and Hopkin, 2000), also addresses the challenges of aviation safety. The book does perpetuate the unfortunate theme that human factors are often ignored in system design and poses a fundamental question: "Can we certify systems in which humans are involved?" Unfortunately, we are now faced with the double problem of mistakes and malice. Finagle and Murphy have provided us with the incentive, and our HF/E tools and techniques are relevant, but we need to earn our place at the table with success stories. In the words of NASA flight director Gene Kranz, "Failure is not an option."

**References**

Beaty,D. (1969). *The human factor in aircraft accidents.* NewYork:TowerPublications.

Casey,S. (1992). *Set phasers on stun.* Santa Barbara, CA:Aegean.

Evans,S.,Hastings,N.,& Peacock,B.(2001). *Statistical distributions.* NewYork:Wiley

Kranz,G. (2000). *Failure is not an option.* New York:Berkeley.

Moroney,M. J. (1962). Facts from figures.[n *Goals, floods and horse kicks - The Poisson*

*distribution* (chap.8). London:Pelican.

Petroski,H. (1992). *To engineer is human.* New York:Vintage.

Wise,J.,& Hopkin,V.D. (2000). *Human factors in certification.* Mahwah,NJ:Erlbaum

**Chapter 6**

**Wrong Number: They Didn’t Listen to Miller**

**(with Georgina Peacock Goebel, Spring 2002))**

In1956, G. A. Miller offered a general statement about operational memory. He said that on average, people could remember seven plus or minus two chunks of information over the short term. This guideline, like Fitts' law from the same era (see Laws & Rules, Winter *2001 Ergonomics* in *Design),* has proved to be a remarkably robust statement about human capabilities, with important implications for design.

**The Evidence**

Like most parents, I have helped my children with their science fair projects and have steered them toward things that interested me - human factors. Some 15 years ago, when my daughter, Georgina, was in eighth grade the cashier at the food store had to punch in all sorts of numbers, such as driver's license number, telephone number, item code, and price. The lines were long and mistakes were made. So Ginny and I decided to study which characteristics of alphanumeric codes lead to errors in information processing. She wrote down lots of examples of codes on cards and presented them in random order to everyone she could persuade to spend about 15 minutes reading or listening to number, letter, or alphanumeric groups and then to write them down.

Miller was right! The more letters and/or numbers in the group, the more mistakes people made. When long sequences were put into groups of two, three, or four characters, recall was better than for longer groups. We also confirmed the primacy, recency, and Von Restorff (1933) effects - characters at the beginning and end of a code were recalled more accurately than were those in the middle; when alphabetic and numeric characters were mixed, the character following a change was often associated with errors.

We told the retail industry about our findings, and the rest is history: Bar coding and credit/debit card readers were introduced. Now my only problem in the supermarket is when the specials have not been put into the computer or have been entered incorrectly - you've still got to keep your eyes open, despite automation.

Unfortunately, none of our children entered the human factors profession. Ginny became a pediatrician. A couple of months ago, she asked the nurse for a baby's weight and was told "17." So she calculated the dose and fortunately, at the last moment, remembered that doses are calculated in metric, not English units. She was not unique in this respect - even NASA scientists mixed up their feet and meters on a Mars mission, and an automotive company ergonomist mixed up pounds and kilograms while using the NIOSH lifting equation in Europe. So we have two chances to err with numbers - we may forget the number or what the number means.

Automation may have helped to reduce the cognitive workload of checkout attendants, but it certainly hasn't solved the problem of wrong numbers. In fact, nowadays we can hardly turn around without coming face to face with an alphanumeric code:

"Your flight reservation confirmation number is M for Mary, P for Peter, 5 as in five golden rings, o as in nothing, and D for Dog."

"Please write the 26 digit identification number on the back of your check when paying your utilities bill."

"When making a calling card call enter the 9-digit access number, followed by the 16-digit card number, followed by the country code, and the area code and number of the person you are calling. If you make a mistake, then start again"

''I'm sorry that you have forgotten your password. Please choose a new one with between 6 and 12 characters, at least half of which must be numeric with no repetitions, telephone numbers, or pet's names."

The following examples are from a vehicle title application form:

Vehicle Identification Number: 2GCD P28T 422266518

Transaction ID: 10167746889353323

Document Number: 10167746889353323

Sticker Number: 15877821WB

Plate Number: 4TFD87

Employee ID: NAWPSAD

So what did Miller tell us in 1956? People can remember about seven independent chunks of information, a chunk being a natural grouping of one or more letters or digits. Some people can remember only five chunks, but others sometimes can remember nine. This rule is pretty useful and has stood the test of time - at least in the classroom. But working, operational, or short-term memory is a little more complex. It all depends on what a chunk is, what you know before being faced with the memory task, how long you have to remember the information, and what goes on while you're trying to remember what you were supposed to remember (Peacock & Davis, 1972).

An interesting thing about what affects how well we remember is that it creates endless opportunities for class projects. Any self-respecting human factors ergonomics engineer can program a Pentium™ to present all sorts of challenges to a cranium. The adventurous ones do not need to limit themselves to digits and letters; they can manipulate shapes, sizes, colors, music, faces, pictures, mazes, operational procedures, travel directions, assembly instructions, accident scenarios, English, Spanish, or Chinese. They can even require the user to do things with the information before he or she recalls it. Champion students of working memory puzzles mix modalities, provide pacing and payoff, observe the young and the old and the inebriated, and even don't let their subjects sleep for days. But they usually find that Miller's magical number 7 ± 2 is in the right ballpark (Baddeley, 1990).

Sometimes wrong numbers are useful. A number of years ago, a wrong Swiss bank number alerted investigators to a lead in the Iran-Contra affair. But mostly wrong numbers are a nuisance.

**The Design of Codes**

So why don't the designers of this information age get it right? People forget, and forgetting leads to all sorts of problems. Therefore, our job as human factors/ ergonomics professionals is first to design memory tasks so that people are not likely to forget, and if they do forget, we must design in a mechanism for recovery. We should start with a brief review of combinatorics. If we have three characters (e.g., 1,2,3), there are 6 possible orders (permutations or selection without replacement) using all three digits. But if we allow repetition of one or more of the digits (selection with replacement), we have 27 unique combinations. Finally, if we allow the choice of I, 2, or 3 characters, we have effectively added one or two leading "blank" characters and can stretch the number of unique numbers to 39.

Now, if we allow the use of all 10 digits and select groups of not more than three digits, we have 999 unique combinations. I encourage readers to calculate how many unique combinations of 26 characters we can have if we include both digits 10) and letters (26). But, of course, bureaucrats must allow for growth in their system, such as the number of vehicles that will be registered in the state over the next 10,000 years. So we have Rule #1: "Use only as many characters as necessary." In this regard I would like to mention that many years ago, I learned to communicate at 25 words per minute using only combinations of up to four dots and dashes. Mr. Morse was an efficiency expert.

If you don't like Rule #1, then go to Rule #2: "Make use of the richness of natural (English, Chinese, etc.) language and its redundancies, with due regard to easyto- remember abbreviations, such as MA, ME, MI, MO, and MN (Rule #2a)." I have had three major employers - in academia (aU), industry (GM), and government . (NASA) - and in every case I have had to learn new abbreviations and acronyms: eduspeak, carspeak, and spacespeak. Unfortunately, the use of abbreviations and acronyms guarantees that many of the users will make errors much of the time, in part because of their inability to link a chunk of to- be-remembered information with its reference system. Unfortunately, although Rule #2 is usually very effective, it is not very efficient, and Rule #2a may be efficient but not always effective.

So we have to rely on Rule #3: "Do a human factors/ergonomics evaluation of your system before you foist it on an unsuspecting public." If you select your participants well, you can get away with anything. Even if you select only those with fifth percentile and above brains (memories) and usage conditions, you will satisfy most of the users most of the time. But what about Granddad, or the user who is new to the system or in a hurry or under some other kind of stress? Now we have the real challenge for HF/E: We must not design for average Joe in average conditions. We must evaluate our system with extreme users in extreme conditions, especially when the implication of error is serious - like when my pension check goes to Rochester MN instead of Rochester MI.

Finally, if all else fails, we should use Rule #4: "Pay attention to Miller's magical number 7 ± 2, and you will accommodate most of the people most of the time without the need for a costly and perhaps unreliable usability study." You might add Rule #4a: Break Miller into two or three subgroups - like CA TDO GCOWPIG.

I would be remiss if! did not finish this article where I started - at the science fair.

A couple of years ago, I had the good fortune to be a judge at an international high school science fair competition held in Detroit. The projects were outstanding; many students didn't get help from mere parents and teachers but worked with professors at local universities. But they were on their own for the interview with the judges. Among the judges were at least six Nobel prize winners. Unfortunately, the only part of the whole event that came near to HF/E was when a tall Nobellaureate in physics made a comment about the headroom in his car. The message is

clear: Shepherd your children, grandchildren, and their friends toward human factors/ergonomics science fair projects, and perhaps one day they will be awarded a Nobel prize for resolving the problem of trade-offs between styling and anthropometric accommodation or preventing

errors in information processing.

**References**

Baddeley, A (1990). *Human memory.* Needham Heights, MA: Allyn & Bacon.

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits of our capacity

for processing information. *Psychological Review* 63, 81-97.

Peacock, J. B., & Davies, P. (J 972). A method for the investigation of interpolated information

and time effects in short-term retention. *Ergonomics,* 15, 701-704.

Von Restorff, H. (1933). Uber die Wirkung von bereichspildungen im Spurenfeld. *Psychologie*

*Forschung* 18, 299-342.

Wickens, C. D. (] 987). Information processing, decision making and cognition. In G. Salvendy

(Ed.), *Handbook of human factors* (chapter 2.2). New York: Wiley.

**Chapter 7**

**JNDs, SDs, HSDs, and DNDs:**

**The Weber-Fechner Law**

**(with Barrett Caldwell, Summer 2002)**

This article is about differences, people's perceptions of them, human factors techniques of analyzing them, and designing jobs to accommodate them.

**Differences**

More than 100 years ago, research psychologists articulated a general rule about the capabilities of the human perceptual system to distinguish differences between physical quantities (Fechner, 1860). They focused on how small a difference could be detected - *a just noticeable difference* (JND for short.) Around the same period, statisticians started to worry about *significant differences* and adopted a risk-based approach in which they admitted that they would be wrong 1%, 5%, or 10% of the time due to chance. Clinical, psychological, and human factors researchers have agonized over these problems for decades (e.g., Jacobson & Traux, 1991; Jacobs, 2000). Some people call the 1% risk level an *honestly significant difference* (HSD). Those who opt for a 10% significance level are talking about a *perhaps significant difference* (PSD). People in the real world often require even greater confidence by using the *definitely noticeable difference* (DND). A convenient DND rule is to require a quantum tenfold difference, although some people will accept smaller exponents.

Other differences include *practical difference* (.01 seconds in a race), *small differences* (peas in a pod), *big differences* (chalk and cheese), *irrelevant differences* (the number of points scored in a losing game), and *meaningless significant differences* (differences on dimensions irrelevant to the problem at hand). There is also the challenge of *actual differences* and *perceived differences.* Readers will no doubt offer even more classes of difference, some of which may be fuzzy and some semantic.

Two outstanding discussions of the perception of differences are to be found in James (1890) and Ladd and Woodworth (1911). Names discussed the conditions of discrimination, first stating the obvious ("the things must be different") and then, "the sensations excited by differing objects should not come simultaneously but fall in immediate succession" (James, 1890, p. 494). Ladd and Trumbull discussed "The Quantity of Sensation," paying particular attention to the robustness and conditions of Weber's (see Stevens, 1951) and Fechner's (1860) development of psychophysics. Contemporary psychologists and instructors in research methods have addressed the differences between statistical and clinical significance (Jacobson & Traux, 1991) and "When is an apparent difference a significant difference?" (Jacobs, 2000). Luce and Krumhansl (1988) provided a comprehensive overview of psychophysical techniques.

**Psychophysical Methods**

To detect differences, one must first establish measurement procedures. It is necessary to establish a standard for the purpose of anchoring the comparison. For example, someone who works 60 hours a week may be said to "work hard," but this is relevant only in comparison to a standard 40-hour week. Larger and smaller imply a comparison, and large, largest, small, and smallest are also relative. This is where Weber and Fechner made their contributions: They were interested in how people judged the differences between physical quantities such as sound, light, temperature, size, distance, weight, vibration, and all the things that human factors engineers manipulate to make the world a better place.

The *method of paired comparisons* requires one to establish a range of values of the thing one is studying (think of packets of French fries marketed by fast food restaurants) and present them to the participants in pairs. Their job is to say which of each pair is the bigger. The method of limits (sometimes called the *method of serial exploration)* requires one to establish a fixed standard stimulus and a set of comparative stimuli that differ from the standard only on the dimension of interest. The participants are presented with the standard, then with the comparative stimuli in series of ascending and descending trials, and are asked to say whether each comparative stimulus is bigger, smaller, or indistinguishable from the standard.

In the *method of adjustment,* the participant or experimenter adjusts the comparative

stimulus in an ascending or descending direction until it is perceived (or not perceived). This is the approach commonly used in audiology.

Finally, there is the *method of constant stimuli.* The experimental approach in this case is to present the comparative stimuli in random order and have the participant judge whether they are larger, smaller, or indistinguishable from the standard. Think of a set of equally sized parcels at a post office distribution center, each parcel having different content and weight. An ophthalmologist uses a combination of these approaches (method of limits and method of paired comparisons) to select the right kind of eyeglasses for people's aging or genetically deficient eyes. Kirk et al. (1967) conducted a study of the judgment of seat height by blindfolded participants and detected a JND of about a quarter of an inch.

The purpose of these experimental approaches is to find out how perceptive people are - that is, how good are they at differentiating between the standard and the comparisons. This is where the question of probability rears its ugly head. People are different, and even the same person is not always consistent. Some can become very good at discrimination when they are well trained. So we need to define the concept of a just noticeable difference, which is the difference value that is detected on 50% of trials.

**Statistics**

This 50% idea is all right in the laboratory, but in practice one may wish to know how small a difference can be detected most (or all) of the time by most (or all) of the people. Think of the French fries. It would be no use advertising different-sized packets of fries if only 50% of customers could detect the difference. So when one applies these psychophysical methods in the practice of ergonomics, one must use different detection criteria - say, 95% or 99% - depending on the importance placed on the detection. One needs *honestly significant differences* or *definite differences.*

An important purpose of statistics is to prevent researchers from hasty conclusions regarding differences with insufficient evidence. The familiar Student's T test may be used to detect the difference between two averages, providing you also obey certain rules regarding the amount and type of data. This is called a *statistically significant difference,* which may not always be a real difference because you may be wrong in your decision some of the time due to chance. To complicate things, a statistical analysis that tries to compare lots of pairs at once can distort whether or not the 95% rule for detecting differences still holds. Tukey (1953) developed a method for making multiple comparisons and detecting which pairs or groups were significantly different. Much has been written on the subject of how to use statistics to detect differences and associations.

**Drift**

People are not consistent. Their memory of the standard drifts as they are exposed to the comparative stimuli. Picture an inspector whose job it is to detect changes in the color of cloth or

anodized metal as it comes out of a manufacturing process. An easy way of demonstrating drift is to have a blindfolded participant close his or her eyes, draw a line on a piece of paper, and then

draw a series of other lines the same length. You will observe that there is autocorrelation between the lengths of successive lines and that the influence of the standard diminishes as a function of the number of intervening stimuli. Similar effects may be observed with weights, sandpaper grades, and wine tasting.

These observations lead to studies of interference effects on short-term memory (the basis of difference detection) for stimuli on all sorts of dimensions (visual, auditory, spatial, tactile, etc.). What about judging in ice-skating and gymnastics, and grading students' essay-type answers?

The practical importance of the drift problem relates to the calibration of inspectors in industrial and other processes (think of sheet metal inspection and airline security). If you calibrate the inspector every time, you double the amount of inspection time. If you don't calibrate the inspectors, very soon they will "drift" and become biased in their judgments. In practice, inspectors are usually required to look for many different differences, and the drift on one dimension may be complicated by attention and importance weighting factors.

**Work Design Considerations**

The laboratory methods of investigation of one-dimensional variables can be applied in the outside world. Over the past few decades, psychophysical methods have been applied to various aspects of manual materials handling and similar physical work (Snook, 1978). The purpose has generally been to find a maximum acceptable weight of lift, given various conditions, or the maximum acceptable work rate given a fixed weight and various conditions. Such methods are combined with epidemiological, biomechanical, and physiological methods to establish work standards. It's no good talking about a 50% difference detection rate when the challenge is to protect the majority of workers or, conversely, to talk about 99% "protection" when the challenge is productivity. So when one applies psychophysical methods, the choice of definition of a difference, acceptability of lift weight, or detectability of a flaw must take into account the policy implications of the research or job design processes.

Both Piaget and Inhelder (1966) and Genaidy, Karwowski, Christensen, Vogiatzis, and Prins (1998) had opinions about this problem of calibration in psychophysical methods. Pia get recognized that small people generally perceived things to be bigger than did large people. Genaidy et a1. were troubled by the calibration of participants in psychophysical manual materials handling experiments by virtue of the instructions. If you ask for the maximum acceptable weight of lift under an incentive scheme, you will get very different answers than if you stress such criteria as safety or comfort.

**The Design of Detection Processes**

So what is a *useful difference?* One guideline is that it is a physical difference of between 1% and 5% if the issue is detection. Thus, the design of inspection processes should be such that the human inspector can have a good chance of sensing and perceiving the difference most of the time.

Training has a big effect on the sensitivity of the inspector, but human perceptual ability can go only so far. When the level of human ability or reliability is not sufficient, one must use job aids or automation. In airports, one finds fixed sized boxes to judge whether a carry-on bag will fit in the overhead bin, scales and luggage tags to identify heavy items, or X-ray and metal detectors to supplement the human sensory systems in security screening. When the difference can be detected automatically but not perceptually, the designer must resort to labeling - as in the price per unit in the grocery store. When amplification of differences is not possible, it is appropriate to design redundancy into the detection processes. In industry it is common to have serial inspection stations so that the eventual customer is assured of a flaw free product.

System design must comprehend human variability in detection capability through appropriate allocation of function, selection, training, difference amplification, mechanization, automation, and redundant processes. These designs must also be context sensitive. We all may be able to detect the difference between a dime, a nickel, and a quarter, but what if it is dark, we are wearing gloves, or we are in a hurry or

are floating about in microgravity?

**References**

Fechner, G. T. *(1860/1966). Elements of psychophysics.* D. H. Howes & E. E. Boring (Eds.), H.

E. Alder (trans.). New York: Holt, Rinehart & Winston.

Genaidy, A., Karwowski, W., Christensen, D., Vogiatzis, C, & Prins, A. (1998). What is heavy? *Ergonomics,* 41, 320-432. Jacobs, D. F. (2000). Juvenile gambling in North America: An analysis of long-term trends and future prospects. *Journal of Gambling Studies, 16(2/3), 119-152.*

Jacobson, N. 5., & Traux, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology,* 59, 12-19.

James, W. (1890). *The principles of psychology.* New York: Holt. Kirk, N. S., Ward, J. 5., Asprey, E., Baker, E., & Peacock, J. B. (1967). The discrimination of chair seat heights. *Ergonomics,* 12,403-414.

Ladd, G. T. & Woodworth, R. S. (1911). Elements of physiological psychology. In *The quantity of sensation* (chapter III). New York: Scribner's.

Luce, R. D., & Krumhansl, C L. (1988). Measurement, scaling and psychophysics.

In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D.Luce (Eds.) *Stevens' handbook of experimentalpsychology.* New York: Wiley.

Parasuraman R., Masalonis, A. J., & Hancock, P. A. (2000). Fuzzy signal detection theory: Basic postulates and formulae for analyzing human and machine performance. *Human Factors, 42,* 636-659.

Piaget, J., & Inhelder, B. (1966). *L'image mentale chez I'enfant* (The mental images of children) Paris: Presses Universitaires de france.

Snook, S. H. (1978). The design of manual handling tasks. *Ergonomics,* 21, 963-986.

Stevens, S. S. (1951). Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1-49). New York: Wiley.

Tukey, J. W. (1953) *The problem of multiple comparisons.* Unpublished, Princeton University, NJ.

**Chapter 8**

**Bias in Human Judgment:**

**Is Your Halo Slipping?**

**(Fall 2002)**

This article is about bias in human decisions and some steps that can be taken to reduce the effects. It is a very important issue because subjectivity is what really drives the worlds of designers, consumers, managers, politicians, and teachers, with a little bit of objectivity here and there.

**Buying a Truck and the Weighting of Evidence**

Cooper (1981) described the *halo effect* as being the most serious and pervasive of all rating errors. All human judgments and decisions are biased in one way or another. For example, the decision may be the choice of which car to buy. The evidence is contained in the advertisement, the inspection of the shiny toy in the showroom, the test drive, and the persuasive patter of the salesman. The astute car buyer will also delve into consumer reports, car magazine reviews, and ratings offered by powerful pundits. The more strategic decision maker will also ask the advice of friends and family because, if things go wrong after the event, he or she can spread the blame. In the end, I bought a Chevrolet pickup because I worked for General Motors and got a discount.

The halo effect traditionally refers to human judgments about other humans, but the mechanisms are similar whether one is talking about cars, pets, cell phone companies, food, or government spending priorities. Much in sky (1990) described halo errors as "good or bad evaluations based on the rater's general feelings about an employee." In general, the rater has a strong feeling or attitude on one criterion, and this feeling spreads over all the other criteria.

College students are faced with the responsibility of course and professor evaluations at the end of the semester. Because this information has important implications for tenure, promotion, and salary, the cunning but transparent professor will distribute pizza along with the evaluation forms. The dutiful students will place a check mark in the right-hand box for all questions. But the administration is one step ahead: It reverses the polarity of the rating scales to catch those students who don't read the questions.

**Halos Come in Different Forms**

Bartlett (1983) distinguished between two kinds of halo. With *invalid halo,* the evaluator fails to discriminate on the criteria and gives uniformly high or low ratings. In the case of *valid halo,* the person or object is indeed good or bad over all criteria. The invalid halo is a true rater error. The valid halo may be a process error in that the criteria from which to choose are not exclusive and exhaustive in terms of the decision to be made, or they are not worded or communicated clearly to the judge. On the other hand, some people and products are indeed superior or inferior on all relevant criteria.

The halo effect is similar to two other forms of bias in human decision making. *Positive or negative leniency errors* refer to those errors in which a judge generally gives high or low ratings. This so-called systematic error may be addressed by having multiple raters, with the expectancy that, on average, the cohort of raters is not biased. Of course, when there is only one rater - the professor - the ratees search out the courses taught by those with positive leniency. Clever marketers exploit this kind of leniency by emphasizing criteria that will be generally viewed as positive – they appeal to the consumer's interests.

Another kind of error is one of conservatism: Some raters are reluctant to give either high or low ratings. The lenient raters have a high average error, whereas the conservative ones have a low variance in their judgments. Some raters have high variance, perhaps beyond the evidence, and a correspondingly higher influence on the final decision. Statisticians draw our attention to the more insidious problems of skewness. Astute judges may show a general level of negative

leniency but slip in a very high rating for their favorite product or person.

**Some Beautiful Minds (and Other Examples)**

Many statisticians and operations researchers have addressed the problems of decision making under conditions of uncertainty and competition. Baird (1989) analyzed these issues in the context of managerial decision making. He described a wide variety of quantitative

techniques that are aimed at reducing bias from decisions by addressing both the probabilistic evidence and the utility of the different outcomes.

Recently, John Nash "hit a triple" with his development of game theory, after earlier work by von Neumann and Morgenstern (1947): He was awarded the Nobel Prize, he had a great movie made about him, and he drew attention to the pervasive challenge of mental illness. Many years earlier, the Reverend Thomas Bayes addressed the numerical nature of convergence given sequentially presented data. Most human judgments are based on sequentially presented information, from either external or memory sources, and provide numerous opportunities for bias.

In 1972, I presented my Ph.D. thesis, "Information Retention in Manual Contro!." It was initially meant to address human behavior in control tasks with sequentially presented data using a Bayesian mode!. It became clear to me early on that people are generally sub- Baysian in their behavior - that is, they do not make the most of available information. Ward Edwards and his successors in the decision theory world have written extensively on this topic (for example, Edwards & Tversky, 1967).

My two years of postdoctoral experience took me into the area of clinical decision making. With the Royal College of General Practitioners, I devised a training exercise in which groups of doctors were given sequences of symptoms and signs and updated their diagnoses and treatment selections as more information became available. Doctors have halos. Some physicians were conservative and were reluctant to offer "extreme" diagnoses until they had conclusive evidence; others refused to budge from their initial decisions despite the evidence. When all the information had been presented, the assembled group discussed the evidence and attempted to persuade their colleagues of the correctness of their diagnosis and treatment selection. We dubbed this the *slipping halo exercise.* The exercise was very instructive to all involved.

I learned about inconsistent weighting of evidence, the physicians learned that they are human, and patients learned to get a second opinion.

In the mid-1970s, I was faced simultaneously with two other problems involving people judging other people. This was my first appointment as a lecturer at the University of Hong Kong, and I had to grade students' work. I also had a consulting appointment with the Royal Hong Kong Police Force related to its promotion system. Many professors - even human factors professors – believe that they are unbiased in their ability to make absolute judgments about their students. They are even confident in their assertion that "this year's bunch is not as good as last year's bunch." I beg to differ. Even professors have halos. I also learned that senior police officers have halos when making promotion decisions, but those in Hong Kong recognized this problem, and we agreed to work together to improve the accuracy and reliability of the promotion process.

The approach in both cases, student grading and the promotion of police officers, involved the development of a statistically based, computer-aided decision process with prescribed criterion weightings, and with manual override where appropriate.

**Steps to Avoid Bias**

The first step in any decision process is to identify a group of experts who have the task of agreeing on the relevant criteria and their weightings. Ideally, the criteria should be exclusive and exhaustive. The criteria should be as measurable, objective, and quantitative as possible. However, some measures may be ordinal (ranked) or even categorical (pass/fail).

Next, each individual is assessed on each criterion independently. Ideally, there should be replication. In the case of the assessment of students, the criteria will probably include midterm and final exams, individual and joint projects, pop quizzes, attendance, participation, and a host of other more or less valid devices. Some professors may not give partial credit in calculation questions based on the thesis that "if the bridge or circuit fails, then it doesn't matter if you understand the theory." Again, ideally, each criterion should be assessed independently and every effort made to avoid the drift effect by randomization. Standard answer templates help to reduce bias.

For those evaluators who are really interested in the removal of bias, there is the familiar method of paired comparisons, but in practice, if the class size is large, such methods are time-prohibitive.

In the case of the assessment of police officers, it was harder to come up with criteria that were not conducive to bias. There was also the problem of differing sample sizes. One could use the number of traffic tickets or arrests, but then some customers might object. One could use the size or difficulty of the crime that was solved, but this might invite the halo effect for those who hit home runs instead of singles. The role of the human factors specialist in such cases is to ask questions regarding the criteria - the knowledge of domain experts is very important, but the rigor of human factors analysis is also important.

A similar problem came with the comparison between departments for university

prizes at the University of Hong Kong. The "civils" thought they were intrinsically better than the "electricals," who in turn were biased in their opinion about the "mechanicals," "industrials,"

and "chemicals." (When I arrived at the University of Oklahoma, the biases changed, but the problem still existed.) Justification for these halos was based on high school grades and standardized test performance at entry level, scores on common courses (like statistics), and

salaries after graduation. Of course, professors and heads of university departments have halos because they have a vested interest in the outcome of the decision, and they compete in zero-sum

games.

The rational approach that I used in all these cases was to standardize the scores on each criterion and apply predetermined weightings to each criterion in the amalgamation process. Letter grade decisions were made based on natural groupings of the overall total and the performance profile on individual criteria. There are various ways in which this profile analysis can be conducted, such as requiring a passing score on each of the criteria. Despite my bias toward this rational approach to avoiding bias, many of my academic colleagues got their halos stuck based on the premise that they were unbiased in their judgment of other people. May their halos never slip, like those of the judges of gymnastics, figure skating, movies, grant proposals, and art.

**Naturalistic Decision Making**

The controversy surrounding rating is not new (Meister, 1985). Contemporary researchers looking into the topic of bias in human judgments have articulated theories of naturalistic decision making (Klein, Orasanu, Calderwood, & Zsambock, 1993). Their collective thesis follows the reservations of Edwards and Tversky (1967) that human decision making is not always rational - we all have halos.

**Postscript**

Don't forget primacy, recency and Von Restorff.

**References**

Baird, B. F. (1989). *Decision making under uncertainty.* New York: Wiley.

Bartlett, C. J. (1983). What's the difference between valid and invalid halo? Forced choice measurement without forcing a choice. *Journal of Applied Psychology,* 68, 218-226.

Bayes, T. (1763). *Essay towards solving a problem in the doctrine of chance* (posthumous). London: Royal Society of London.

Cooper, W. H. (1981). Ubiquitous halo. *Psychological Bulletin, 90, 244-281.*

Edwards, W., & Tversky, A. (1967). *Decision making.* London: Penguin.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambock, C. (1993). *Decision making* in *action.* Norwood, NJ: Ablex.

Meister, D. (1985). *Behavioral analysis and measurement methods* (ch. 9). New York: Wiley.

Muchinsky, P. M. (1990). *Psychology applied to work.* Belmont, CA: Wadsworth.

Peacock, J. B. (1972). *Information retention in manual control.* Unpublished Ph.D. thesis, University of Birmingham.

Von Neumann, D. & Morgenstern, O. (1947). *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

**Chapter 9**

**What Kind of Shape Are You In?**

**Anthropometry and Appearances**

**(Winter 2003)**

This article is about body size and shape and our attempts to measure, analyze, and accommodate them by design. It addresses various rules of measurement, design policies, and some opportunities with the recent availability of detailed full body scan data and digital human

models. Anthropometry and its applications are probably the most frequently used and abused areas of ergonomics.

"You look f't, mon." On a beautiful Sunday afternoon in 1960, I was playing a cricket match in a cup final against a team made up of West Indians. Our captain, who was a somewhat overweight but otherwise excellent player, walked out to greet the opposing captain to perform the ritual coin toss. The greeting took our leader by surprise until he translated the Jamaican "a" into an "i." Appearances matter. In modern-day America, we are bombarded with messages of how to change the rotund "a" into a slender "i."

A few years later, while attending my first lectures in anthropometry at Loughborough University, I learned a lot about bony landmarks, segment lengths, joint centers, girths, and allowances for soft tissue and sometimes for clothes. The inexactness of this science was emphasized by things like clearance "allowances" and inter- and intraobserver differences. I also found the robust normal distribution to be very forgiving. The need for good definitions and observer training became very clear. Exactly where on the acromion should I make my mark?

Recently, our nation's shortest astronaut went on a jaunt to repair the Hubble space telescope. Shortly after arriving in the microgravity environment, she participated in a test similar to the one we all do with our children: standing upright against the door frame to measure our height. Astronauts grow (their spines lengthen) in space, just as we do when we go to bed at night. But when astronauts grow, they might no longer fit into their equipment, and this could cause serious difficulties.

**Analysis of Size and Shape**

There is more to anthropometry than segment lengths - shape is also important. Back in the middle decades of the last century, Sheldon and coworkers developed the system of somatotyping. This system classifies people according to their degree of fit to three body shape dimensions: ectomorphy (tall and thin), endomorphy (round), and mesomorphy (broad-shouldered and muscular). Each dimension is rated on a sevenpoint scale, and the so-called average man (my anthropometry professor) would be described as a 4,4,4. Later, this theory was extended to suggest that mesomorphs were more likely to have criminal tendencies. This sounds like a good opportunity for a discussion of the differences between causality and correlation!

Drillis and Contini also pursued the correlation issue (see Chaffin & Andersson, 1991; Kroemer, Kroemer, & Kroemer- Elbert, 1994). They articulated a series of equations that allowed individual segment dimensions to be predicted from measures of stature. Tall people have long legs and arms and backs and broad shoulders - or do they? The NASNWebb (1988) publication showed the relatively low correlation coefficients among 20 dimensions. Readers are encouraged to measure various segment lengths of a group of people to check out the reliability of this approach.

Here is an easy classroom example. Ask the class members to stand in a row in order of their height, and then have them take one step forward for each 10 pounds of body weight (or other measure). Appropriate use of tape and string for the axes and regression lines adds to the fun and is a way of showing the random nature of residuals in regression analysis. Certainly, big people usually have big segments, but the correlations between measures are often very weak.

Clothing manufacturers have the practicalities of measurement down to a fine art. They aspire to accommodate a large proportion of the clothes-buying public with the widely recognized symbols XS, S, M, L, XL, and XXL for men and women separately by reference to key segment lengths and girths: chest (bust), waist, inseam, sleeve, and hips. They prescribe so-called proper procedure for measurement, such as "Stand with your feet together, measure around the fullest part of your lower body" or "Use a pair of pants that fit you and measure the inside seam." Those who size shoes, gloves, and rings have also developed simple and usually sufficient devices to ensure the fit between their products and their target populations.

The Western love-hate relationship with ectomorph ism is not totally unreasonable - too much fat is unhealthy, but how much is ideal? Body shape and size have five determinants, four of which are under the control of the individual. First there is inheritance: We can't choose our parents. Next come the big "E's" – Eating and Exercise. Although it is difficult to undo what years of neglect have created, it is possible to change our circumferences. The next opportunity, that is partly under our control with a little help from the apparel industry, is to buy accentuating for forgiving clothes. If all else fails, there is the plastic surgeon's knife. We may also look forward to genetic engineering.

**The Science of Shape**

The medical and physiology communities are also concerned with shape and size. Their concern is that measurable fat on the outside reflects the presence of more fat on the inside, especially in our arteries. They go on to suggest a strong correlation between shape and health, with clearly described causal links. Like the clothing industry, the medical community seeks to describe the complexities of shape by reference to simple indices. The most attractive index, because of its ease of measurement, is the ratio of height to weight. Both components of the index are amenable to sufficiently accurate measurement by the general population.

The more adventurous growth anthropometrists describe variations of the Body Mass Index and the Ponderal Index. Unfortunately, when we dig a little deeper, we find that different body constituents have different densities – fat is not as heavy as muscle. This provides the incentive for the measurers to develop new measures. These measures include skin calipers, underwater weighing and electrical impedance.

The white-coated skin caliper brigade uses these plier-things to attack selected body parts, such as the tissues over our biceps and triceps, and just above our iliac crest. They then amalgamate the individual observations to produce an index of fattiness. This is more physically painful than the cursory "You look ft, mon" and may be more reliable, but it contains a similar message. These sadistic scientists get even more aggressive when they try to drown you in a tank of water while they watch the scale display. But the contrast between open air weighing and underwater weighing does provide a good indication of the ratio between body fat and lean body mass.

If you want to lose weight, become an astronaut - you don't weigh anything in space. Unfortunately, your body fluids also rebel when they become gravity deprived, and you change shape. The best estimates of body shape and size in space are based on a very elaborate set of linear and girth measures of large samples of astronaut candidates under earth gravity conditions. This lack of data is a pretty important challenge because the cost of space hardware is very high and the risks of a mismatch are important: You can't fit into the space lifeboat if you are too big!

Back on earth we are still faced with the challenge of measuring body size and shape for the multiple purposes of health checks, clothing fit, and equipment and workplace design. Fortunately, a dedicated band of anthropometrists and ergonomists over the years has developed well-described measures and produced oodles of population data. Roebuck and colleagues (Roebuck, Kroemer, & Thompson, 1975) wrote an outstanding book, *Engineering Anthropometry Methods,* and the late Stephen Pheasant wrote *Bodyspace* (1986) on anthropometry, ergonomics, and design. The World Health Organization produced a broad survey of anthropometry related to selected regions around the world. These sources - together with many others, such as NASNWebb (1988) - contain great detail not only of population anthropometric methods and data but also of the translation of this evidence into a form useful for the design of cars, workplaces, tools, equipment, and clothing.

**Full-Body Scanning and Digital Modeling**

Recently, the high-tech world has moved into the size and shape arena from two directions: laser scanning and digital human modeling. The CAESAR project (Robinette, 2000) is a collaboration among international anthropometrists to use laser scanning to describe in great detail and with great accuracy the shape and size of people. These data are not merely related to whole bodies; they also focus on particular segments such as hands, feet, and heads for the design of more precisely fitted clothing and protective equipment.

One problem with laser scanning is that it provides a lot of data that are hard to reduce to meaningful and useful parameters. Indices like maximum circumference vary from person to person, and it is not always possible to do a direct translation of an individual's data into the relatively simpler criteria for equipment design. But great strides are being made with the help of multivariate statistics and nonparametric methods.

Digital human modeling has really taken off in recent years (Badler, Phillips, & Webber, 1993; Chaffin, 2001). The Society of Automotive Engineers sponsors an annual international conference and has formed a broadly based committee (SAE G13) to facilitate communication between scientists and engineers interested in measurement, modeling, and design. These anthropomorphic model makers and the users of models are faced with the age-old challenge of correlation between measures of shape and size. Simplistic use involves the selection of a 95th or 5th percentile person based on stature with the hope that the other measures are a reasonable fit. Monte Carlo methods allow the accommodation of prescribed percentages of multiple dimensions. Laser-scanned data enable the designer to comprehend exactly how many of a known population will be accommodated by a particular design and where a design is most restrictive. Most of these models are static, but when the joints are allowed to move, all sorts of new inaccuracies creep in. In recent years, the University of Michigan HUMOSIM project (Chaffin, 2001) has attempted to add naturalistic movement to improve the usefulness of these models.

These major advances in measurement and modeling science will be useful and provide much-needed improvement in validity. But the techniques have a long way to go to match the universal simplicity and usually sufficient utility of the clothing manufacturers' size charts and the casual observer's "You look f't, mon."

**References**

Badler, N. I., Phillips, C. B., & Webber, B. L. (] 993). *Simulating humans: Computer graphics animation and control.* New York: Oxford University Press.

Chaffin, D. B. (200]). *Digital human modeling for vehicle and workplace design.* Warrendale, PA: Society of Automotive Engineers.

Chaffin, D. B., & Andersson G. B. j. (1991). Anthropometry in occupational biomechanics. In *Occupational biomechanics* (chapter 3). New York: Wiley.

Kroemer, K., Kroemer, H., & Kroemer-Elbert, K. (1994). Anthropology and anthropometry. In *Ergonomics: How* to *design for ease and efficiency* (chapter 3). Englewood Cliffs, NJ: Prentice Hall.

NASA/Webb Associates. (1978). Anthropometric sourcebook (NASA 1024). Houston, TX: Author.

Pheasant, S. (1986). *Bodyspace.* London: Taylor & Francis.

Robinette, K. M. (2000, Summer). CAESAR measures up. *Ergonomics in Design.* 8(3), 17-23.

Roebuck, J. A., Kroemer, K. H. E., & Thompson, W. G. (1975). *Engineering antlrropometry methods. New* York: *Wiley.*

Roebuck, J. A. (1995). *Anthropometric methods: Designing* to *fit the human body.* Santa Monica, CA: Human Factors and Ergonomics Society.

**Chapter 10**

**Stand (Sit) Up Straight: The Functional Anatomy of Posture**

**(Spring 2003)**

Posture is a very elusive concept because it can be used to mean many different things. This article attempts to describe the mechanisms involved in posture and resolve some of the difficulties associated with the idea of "good posture."

**The Biomechanical Basis**

Body posture commonly refers to an instantaneous set of relationships (angles) between various body segments (bones) in response to external forces (gravity, loads, and external supports) and the muscular and nervous responses to those forces to create a static equilibrium. The standing and sitting postures are familiar examples. In reality, posture is a dynamic issue with ongoing minor and major modifications in response to changing task demands. The dynamic nature of posture is largely affected by the need to maintain balance over the posture base and the physiological demands of the muscles to stimulate circulation. A particular posture may also be affected by body size and shape, short-term fatigue and long-term habits. *Optimal postural variety is the ideal aim of design and operational interventions.* A truism quoted among ergonomists and therapists is, "The next posture is the best posture."

Posture is managed by the activation, reciprocal inhibition, and cocontraction of muscles around joints, commonly through the medium of proprioception and spinal reflexes, but sometimes with cerebellar, subcortical, and cortical inputs. Most postural muscles span two joints, which leads to complex interactions in the biomechanical chain in response to external forces and internal automated or volitional control activity. The low-level muscle activity needed to maintain, for example, an upright posture is generally referred to as *postural tone.*

Humans vary in their size and shape. The science and technology of anthropometry is applied to the description of this human variability (Kroemer, Kraemer, & Kroemer-Elbert, 1994; Pheasant, 1986; Roebuck, 1995.) One key finding is that the variability exists at the segment level so that predicting an individual segment characteristic may not be made with certainty from knowledge of another measure such as stature. The same is also true of posture and movement, although trained observers may be able to extract certain common features of posture and movement. Such patterns enable one to recognize a familiar person under very degraded lighting conditions or recognize a gait typical of a particular neurological dysfunction. Segment characteristics, postures, and movements under Earth's gravity conditions vary considerably among individuals and may conveniently be modeled by the (multi)normal distribution, functional regression (Faraway, Zhang, & Chaffin, 1999), and associated correlation matrices.

**The Educational and Therapeutic Basis**

"Good" posture is the result of inheritance, habit, and changes in motor, biomechanical, and external factors (Gardiner, 1957). Growth and physical activity create positive changes in the biomechanical functions that facilitate a healthy (dynamic) posture. This is clearly demonstrated by observing children learning to walk in their first two years (Kottke, Stillwell, & Lehmann, 1982). Injury and pain result in changes to the mechanisms and hence the postures. Physical education and therapeutic interventions are employed to restore functionality. In general, such interventions employ gravitational and other resistive forces to (re)develop the biomechanical (and motor) capabilities necessary for good posture.

Modification of the gravitational component by the use of water or other suspension devices may be an important element of the rehabilitation process. Such interventions lead to the natural adoption of postures that vary from the normal, upright, gravity-resistant one. Similar joint relationships may be observed during the "side lying" posture commonly adopted for sleeping.

Inherited characteristics, long-term behavioral habits, and disease may give rise to asymmetries of the body. The classic scoliosis (sideways bending of the spine) can, in extreme cases, have serious effects on both biomechanical and internal organ function. Variations in lumbar lordosis (increased concavity of the lower back) may be accentuated by obesity and pregnancy. Thoracic kyphosis (increased convexity of the middle region of the spine) is often noted as a result of aging, both through osteoporosis and intervertebral disc degeneration. One behavior-induced condition is termed "shop girl's hip," a posture in which a person who stands all day carries most of his or her weight on one leg with the other knee bent. The result is a dropping of the hip to the nonweight- bearing side and a compensatory lateral curvature (scoliosis) of the spine.

Some health advocates focus on balance and symmetry and the restoration of these characteristics through manipulation, exercise, and self-awareness. Their diagnostic methods may include calibrated instrumentation or a trained eye for deviations from the horizontal of the pectoral and pelvic girdles.

**Empirical Evidence**

The measurement of body size and shape - anthropometry - is a mature science, although recent technology using laser scanning is opening up many more opportunities for the analysis and communication of human spatial complexity. Movement analysis has been the province of sports analysts (gymnastics judges, baseball pitching coaches, etc.) and rehabilitation professions for many years. These approaches have usually been subjective and qualitative, sometimes with quasi-quantitative overlays. In recent years, the University of Michigan HUMOSIM project has gathered statistical data on common naturalistic motions, such as reaching and grasping (Faraway et aI., 1999). Gait analysts and industrial ergonomists have used goniometers with

various degrees of sophistication. As in anthropometry, the reduction of this naturalistic motion data involves complex statistical methods.

Posture has received attention from seating, materials handling, and workspace analysts. The right –angled posture is a classic artifact of discussions of seating, but even casual observation shows that naturalistic seating postures vary considerably from this ideal. Materials handling analysts hover between straight backs and bent knees and the intermediate naturalistic segment relationships. Workspace analysts add consideration of upper limbs and necks to discussions of the lower limb and lower back. As with anthropometry and motion analysis, the description of posture is categorical and statistical. And posture exhibits wide within- and between individual variability.

**Joints, Muscles, and Ligaments**

Adjacent joint surfaces are shaped to permit various degrees of freedom; we have hinge, saddle, ball and socket, and sliding (synovial) joints. These convenient descriptions define the directions and, to some extent, the ranges of motion. It must be noted that the surfaces, although smooth, do not exactly fit formal geometric shapes, with the result that the centers of rotation associated with particular joints are not fixed. Muscles, or their associated tendons, tend to cross the joints in the plane of their major motions. Thus, the knee has the quadriceps in front and the hamstrings behind, whereas the more mobile shoulder has muscles crossing all planes. Static posture is maintained by co-contraction of opposing muscle groups.

All synovial joints are surrounded by fibrous capsules and ligaments, which are generally thickened in the regions where there is least movement, as on the lateral and medial aspects of the knee. In relaxed postures (minimal co-contraction of muscles), ligaments may have major roles in joint stability. Such postures occur when external perturbing forces are minimal and the line of gravity passes close to the joint centers. People may be described as "hanging in their ligaments" when adopting such postures.

Another situation occurs in the posterior ligaments (including the elastic ligamentum flavum) of the back. Electromyography measures demonstrate minimal electrical activity of the muscles when the lumbar spine is fully flexed, both in (stooped) standing and (slumped) sitting (Floyd & Silver, 1951). Members of primitive populations are noted for their ability to squat, with knees and hips fully flexed, for long periods. Such postures are rare among those populations accustomed to using chairs. Similarly, members of traditional agricultural communities (rice and vegetable planting and harvesting) have the ability to work for many hours in stooped postures. Relaxed sitting as observed at work, at home, and often in the car also may involve postures with extreme lumbar spine flexion.

Joint surface shape, muscle and tendon length, ligaments, obesity, clothing, and other restrictive orthoses or splints affect joint mobility and range of motion. Some of these characteristics are inherited; some can be modified by mobility exercises and some by design of clothing and work equipment. As noted earlier, most postural muscles pass over two joints, so the position of one joint affects that of the adjacent one. Lying on one's back and attempting to touch one's nose with one's knee, with and without straight knees, demonstrates this clearly. The main culprits in this instance are the hamstrings.

Casual observance of yoga classes indicates that women are generally more flexible than men, although some leveling can be achieved by dedicated practice; this typically involves the gentle stretching and holding of outer-range positions. Similar stretching activity is observed among athletes of all ages at the behest of their coaches and trainers.

Restrictive protective clothing is very important in various hazardous occupations. The shoulder-pelvic straps and/or carried loads of construction workers, firefighters, skydivers, foot soldiers, and mountaineers may restrict shoulder girdle motion. Such restriction - especially if associated with resistance to lateral rotation at the shoulder joint - will restrict upward reach. This restriction may also be observed among football players with shoulder pads and extravehicular astronauts with pressurized suits, hard torso units, and bulky helmet connections. Such restrictive clothing constrains not only posture and range of movement but also speed and force of movement (Gonzalez, Maida, Miles, Rajulu, & Pandya, 2002). Furthermore, the compensation required in a joint (e.g., the shoulder) may be excessive and even damaging if adjacent joints (e.g., shoulder girdle) are compromised. Similar observations may be made with regard to downhill ski boots, which greatly restrict ankle movement and thus require compensatory knee movement.

**The Role of Pelvic Tilt in Standing and Sitting**

Relaxed standing may involve relaxation of the abdominal muscles and an increased lumbar lordosis, causing the pelvis to rotate forward on the hip joints and an outward and downward shift of the abdominal contents. The instruction to "stand up straight" by the parent, gym teacher, or drill sergeant is followed by contraction of the abdominal and hip extensor muscles, retraction of the shoulder girdle, straightening of the neck, and backward rotation of the pelvis.

Pelvic tilt in sitting is also important. In car driving there is a slumping tendency because of the accelerations, vibration, seat orientation, and material. The pelvis rotates backward with an associated flexion of the lumbar spine and, because the foot is fixed, knee flexion. These movements are also associated with a desire to relieve pressure on abdominal contents by using the seat recline function. Similar tendencies may also be observed in the seated postures of computer operators and those in other seated occupations. The draft standard, *Human Factors Engineering of Computer Workstations (2002), d*efines three seated reference postures - reclined sitting, upright sitting, and declined *sitting.* These *classifications* are at best an oversimplification of the wide *variety* of *sitting* postures.

Microgravity, as the name suggests, removes the *primary* external posture provoking force. The consequence *is,* as *with suspension* in water, that the segment relationships change from those associated with earth gravity. However, motor and *biomechanical* memories are such that, over the short term, people exposed to microgravity may adopt some composite posture. Over *time,* however, adaptation to the absence of gravity will lead to the adoption of a neutral or microgravity posture. As in earth gravity conditions, *this* posture will be subject to variability, both within and among *individuals.*

The *microgravity* neutral posture described in *Manned Systems Integration Standard* 3000 (NASA, 1995) shows average *joint* angles with some *indication* of variability. For example, the so-called neutral knee angle is 1330 with a standard *deviation* of 8°. The *hips* tend to be flexed and the ankles plantar flexed. The shoulder girdle is elevated and the shoulders abducted. A study by Mount, Whitmore, and Stealey (2003) indicates that there may be considerable variability in such postures, probably caused by volitional and motor memory overlays.

In space operations, *it* is necessary to hold fixed postures for extended durations during robot arm control and glove box operations. These postures are *achieved* by the use of fixed (toe loops, grab rails) and adjustable (foot and leg bars) restraints and simple wedging against convenient surfaces. As expected, crew reports indicate a need to easily change postures from time to time. It should also be noted that crewmembers report calluses on the tops of their feet, *which* are not a common locus of prolonged pressure on earth.

**Modeling**

In recent years, there has been an explosion of anthropometric models. Some are based on population data for segment properties, and others allow the *input* of individual characteristics, sometimes based on photographic evidence (Chaffin, 2001). *Static* strength models have been developed to add to the anthropometric content and postures derived from observational data. Dynamic models benefit from recent increases *in* comput*ing* power. Digital human models such as JackTM, Safework™, and *Ramsis™* are widely used representations of human segment *characteristics.* For accurate posture modeling, it will be necessary to *obtain* earth gravity, partial *gravity* (e.g., watersupported) and *microgravity* data to articulate the interactions between anthropometric, force context, and behavioral factors and their effects on posture. Posture *modeling* will also have to account for the *interactions* between adjacent joints, as affected by the length of two joint muscle-tendon systems.

**Design Opportunities**

Conventional, but perhaps flawed, *wisdom* recommends the liberal use of postural support through orthotic devices. Such devices include orthopedic *inter*ventions for the severely disabled or *injured* and simpler *interventions* such as the medial arch orthotics recommended by podiatrists. In addition, splints and clothing of various materials may be used to constrain posture and movement, as in backbelts. Finally, the many variations of the chair (in the space context,

restraints) serve to constrain (support) certain postures with the overt intention of reducing muscle fatigue.

In practice, the static postures encouraged by such devices may, if overused,

cause greater fatigue and adverse circulatory effects. Consequently, the design and use of posture-constraining devices must comprehend the dynamic needs for optimal postural variety and not usurp the *responsibilities* of the musculoskeletal system. For this reason, the use

of posture-constraining devices should always address the *time* dimension.

**References**

Chaffin, D. B. (2001). *Digital human modeling for vehicle and workplace design.* Warrendale, PA: Society of Automotive Engineers.

Faraway, J. J., Zhang, *X.,* & Chaffin, D. B. (1999). Rectifying postures reconstructed from joint angles to meet constraints. *Journal of Biomechanics,* 32, 733-736.

Floyd, W. F., & Silver, P. H. S. (1951, January). Function of erector spinae in flexion of the trunk. *The Lancet,* pp. 133-138.

Gardiner, M. D. (1957). Posture. In *The principles of exercise therapy* (chapter 14). London: G. Bell and Sons.

Gonzalez, 1. ]., Maida,]. c., Miles, E. H., Rajulu, S. 1., & Pandya, A. K. (2002). Work and fatigue characteristics of unsuited and suited humans during isolated, isokinetic joint motions. *Ergonomics* 45, 484-500.

Human Factors and Ergonomics Society. (2002). Reference postures. In *Human factors engineering of computer workstations* (Section 5.2.1). Santa Monica, CA: Author.

Kottke, F. J., Stillwell, G. K., & Lehmann, J. F. (1982). Therapeutic exercise to develop neuromuscular coordination. In F. ]. Kottke (Ed.), *Krusen's handbook of physical medicine and rehabilitation* (chapter 19). Philadelphia: W. B. Saunders.

Kroemer, K., Kroemer, H., & Kroemer-Elbert, K. (1994). Anthropology and anthropometry. In *Ergonomics:How to designfor ease and efficiency* (Chapter 3). Englewood Cliffs, NJ: Prentice Hall.

Mount, F. E., Whitmore, M., & Stealey, S. 1. *(2003). Evaluation of neutral body posture 0'1 Shuttle Mission STS* 57 *(SPACEHAB-l)* (NASA TM-2003-104805). Houston: NASA.

NASA. (1995). *Manned systems integration standard NASA-STD-3000.* Washington, DC; Author.

Pheasant, S. (1986). Posture. *In Bodyspace* (chapter 8). London: Taylor & Francis.

Roebuck, J. A. (1995) *Anthropometric methods: Designing to fit the human body.* Santa Monica, CA: Human Factors and Ergonomics Society.

**Chapter 11**

**Pay Attention**

**(Fall 2003)**

This article is about the theories of attention and vigilance, some of the practicalities of human behavior and performance when faced with complex search challenges, and some organizational analysis, decision, and design opportunities to ensure effectiveness and efficiency of the search process.

**Attention Research**

Attention is a bit like posture: It keeps shifting. It's difficult to do more than one thing at a time. We have multiple joints and senses that work sometimes together and sometimes independently. In addition, we acquire bad habits. Attention is also about sensation and perception, which in turn recruit memory, learning, and a whole array of other cognitive and motivational functions. Annunciation, or large changes in one's sensory environment, may direct attention to a particular source, but the instruction to "pay attention" implies that we should perceive the content of the message or wait diligently for a low-probability signal. Vigilance or sustained attention fluctuates and generally drops off over time because of the confluence of many cognitive and motivational changes.

Students of attention like to distinguish different kinds of behavior (Sanders & McCormick, 1993). Selective attention means searching for a particular kind of anomaly (signal) in a sea of noise. Focused attention implies that the source has been identified and that other sources of sensory input are ignored. Divided attention implies that one shifts among multiple sources of information that are pertinent to parallel tasks.

N. H. Mackworth developed the clock test to study deterioration in human attention. Every now and then, the hand made a double jump (Mackworth, 1950). Needless to say, human detection performance deteriorates even in this exciting task. The generally used description of this deterioration is the *vigilance decrement.*

Since those early days of human factors/ergonomics research, attention has focused on the characteristics of the task at hand, the task context, and the variable patterns of deterioration of human performance (Drury & Fox, 1975). One paradox commonly observed in most human activity is that people improve with practice (they learn), but their perceptual performance deteriorates with repetition (they exhibit a vigilance decrement). Another apparent paradox is described by the Yerkes-Dodson law, which implies that (vigilance) performance improves with inspector motivation but deteriorates if one tries too hard.

Contemporary research addresses the broader idea of *situation awareness* (Endsley, 1995; Wickens, 2001), the ability of a human operator to comprehend the current and likely future state of affairs; and ramifications of *signal detection theory* (Parasuraman, Masalonis, & Hancock, 2000), the ability to detect important signals in a sea of noise. A comprehensive effort to identify the multiple human, task, and context factors associated with human performance deterioration over time is contained in *Stress, Workload and Fatigue* (Hancock & Desmond, 2001). Cognitive workload is a major factor in attention and situation awareness. Fatigue and boredom also affect performance in vigilance tasks. Scerbo (2001) offered the heretical solution that people should be allowed to stop doing a task when they become bored!

**Paying Attention at Work**

The foregoing research was carried out in the laboratory, in simulators, and in a wide variety of field situations, including industrial inspection, airport security, air traffic control, process control, car driving, and aircraft piloting. In fact, almost all work and play have elements of inspection, signal detection, and vigilance. And mistakes are everywhere. False positives usually have less serious implications than do false negatives, unless you are the innocent guy in a police identification lineup. False negatives may upset the customer who receives a faulty product.

Recently, I had the opportunity to walk through the woods looking for debris from the STS 107 space shuttle. The inspectors walked in (somewhat) orderly lines through the woods to look for signals of many forms. We were instructed to be very conservative in our judgments - false positives were not a problem because secondary teams of specialists checked out any flagged item.

I got first-hand experience of being an inspector in a critical task. It rained heavily and it was cold; the woods were full of undergrowth with the usual mix of forest debris such as leaves, rotting branches, and empty beer cans. The terrain and flora were very irregular - hills, streams, thickets, and a mixture of conifers and deciduous trees and shrubs. It is amazing how many shapes, sizes, and shades of brown, green, yellow, red, and gray are to be found in the winter woods.

We were all highly motivated and looked for anything that was different. But I got to thinking about my reliability as an inspector. The task was visually and perceptually very difficult. I thought about my scan pattern - up and down, side to side, behind obstacles and sometimes up into the trees. What about signal detection theory? What about contrast? What

about training? What about my old eyes? What about physical discomfort? What about fatigue and the vigilance decrement? There certainly were no boredom or motivation issues. How good were we at inspection? Who knows? But our task was easier than that of the police divers, who had near zero visibility in cold, murky lakes. In the end, the results of our efforts were declared a success.

**Some Different Inspection Situations**

Many years ago, I was asked to advise on an inspection task related to the color of anodized metal window frames for a high-rise glass building. The frames were dipped into a tank of liquid that varied somewhat in its temperature and chemical makeup; an inspector lifting the frames out and looking at them determined the duration of the dipping. It was dark inside the large factory, and the architect and customer were upset that the building was going to have frames of different colors. Here was a job that people were not good at, however experienced. The ultimate solution was to invest in process control of incoming materials, bath temp. erature . and chemistry, and duration of immersion.

Color difference is also important in the textile industry, especially where blue jeans are concerned. In a psychophysical experiment, the strategically presented comparative stimuli varied in shade from much darker to much lighter than the standard, which was presented every time, every tenth time, or every hundredth time. I was able to detect drift in the apparent standard as a function of the number of items and their shades after the presentation of the actual standard. Of course, there is a trade-off; if you recalibrate" the inspector every time, you greatly reduce his or her inspection productivity in the interests of quality. However, if productivity is the game, you get drifting quality.

A very important inspection activity occurs in the automobile industry, where the inspector's task is to search for surface blemishes, dings, and scratches. Great efforts are made in the lighting environment, including directional lighting and relative movement of the sheet metal surfaces to make the signal stand out. But inspectors are people; they are usually trained and experienced, and there is self-selection for the job through a seniority agreement. What is more important in sheet metal inspection - eyesight or experience? There is more to inspection than meets the eye.

I saw a somewhat similar task in an automobile component-manufacturing factory. The component was a couple of inches long with attachments at each end and various appendages. The inspector was checking the components at a rate of about 10 per minute, and he told me that he was expected to identify 35 different flaws. He was self-paced and had an incentive to hurry but was confident in his ability to perform his task flawlessly. The evidence proved otherwise.

On another occasion, I observed a worker standing, all day, in front of a continuous, yard-wide roll of cloth as it moved, with back lighting, between two rollers. His job was to mark flaws with a pen – if he caught them in time. The cloth moved very quickly. The supervisor of the task apparently believed that performance would improve if the inspector "stood to attention."

One dark evening I was on guard duty at my Royal Air Force station. A car drove up and stopped. I glanced at the occupants and their passes before stepping back to attention. But the car didn't move. The passenger said to me: "Do you know who I am?" I had to admit that I had not read the name on the pass. "Airman," said the passenger. "I am your commanding officer, I suggest that you pay more attention to your duty."

**Airport Security**

Recently we have become accustomed to the fact that airport check-in activities sometimes take longer than the flight itself. The processes, training, and equipment have become much more sophisticated since the immediate response to 9/11*.* The Good Guys become impatient with the general hassles of juggling the laptop, the shoes, coats, knitting needles, three carry-on bags, and the cell phone while being pushed from behind. The essential conservatism of this preliminary screening leads to many false positives and the associated humiliation of detailed inspections. The "free safeties" stand in relaxed postures with their eyes, sniffing dogs, and weapons primed. Meanwhile, the Bad Guys learn how the system apparently works and plan accordingly. The uniformed inspectors do their jobs to the best of their ability with existing equipment, uniforms, signal enhancement, specific training, and rotation. But they are human and, unlike machines, have a vigilance decrement and perceptual performance imperfections.

Certainly, people perform better if they have learned what to look for. In fact, experts will often detect signals that a novice would not even recognize (Drury & Sheehan, 1969). Also, the insertion of dummy signals where real signals are very infrequent serves the dual purpose of motivating the inspector and recalibrating him or her in terms of recognizing what really is a signal. A third intervention is to exaggerate the contrast between the signal and the background or even to use automation or other job aids to annunciate the presence of a signal. On the

human end, it is appropriate to select (train and assign) inspectors who have both good experience (knowledge) and good eyesight! Work design interventions include improving the environment by attention to lighting, temperature, and distractions. Organizational interventions

revolve around work - rest schedules and job rotation.

**What Can Ergonomists Contribute to the Challenges of Attention?**

Think of the special challenges of teenage driving. The experienced parent reader will no doubt add more specific rules:

1) Signal Characteristics

a) Clearly describe what has to be attended to and why.

i) road conditions, road signs, other vehicles, pedestrians

ii) in front, to the side and behind, blind spots and way out in front and way back

b) Amplify the signal and increase its contrast with the background.

i) lighting, size, location, color, motion

c) Automate the detection process and rely on the operator for interpretation.

i) RODS and NODS (rear obstacle detection systems, near obstacle/lane change detection systems)

ii) road, traffic, weather, vehicle, and driver condition advisories

d) Identify the likely source (where) and time (when) of signals.

i) standard signal locations, forms, and situations (where, what, when)

e) Provide alerting information.

i) warning signs

ii) alerting devices

2) Environment Characteristics

a) Remove or reduce specific distractions.

i) passengers, entertainment, cell phones, etc.

b) Remove or reduce general distractions.

i) thermal, auditory, lighting sources

ii) exterior information sources

iii) traffic density (especially during training)

3) Operator Characteristics

a) Select/assign operators with the basic sensory, perceptual, and motor faculties.

i) driver education

ii) parental choice

b) Train operators in anticipation, signal detection, signal interpretation, response selection, and response implementation.

i) practice, practice, practice - 1000 miles with a parent in passive and active modes

ii) graduated workload, part training

c) Impose processes for calibration, maintenance, quality control, and continuous improvement.

i) probationary period, with feedback

ii) formal testing

d) Impose processes that address short-term performance degradation.

i) drugs, alcohol, fatigue, distractions

ii) no driving late at night

e) Educate operators regarding the implications of correct and false (positive and negative)

detection and response.

i) statistics, case studies, demonstrations

4) Response Selection

a) Describe response options and preferred choices.

i) go, go slower, go faster, stop, turn, signal

b) Describe implications of correct and incorrect responses.

i) physics - mass, velocity, acceleration, momentum, inertia, force, friction, center of mass

ii) time - reaction time, decision time, system response time

iii) pain and suffering - you, the other guys, your friends and relatives

iv) dollars - repairs to people, vehicles, and the environment; insurance

5) Operation Characteristics

a) Reduce the duration of continuous monitoring.

i) journey duration - stay close to home

b) Increase the time available for signal detection.

i) slow down, back off

c) Provide calibration, alerting signals.

i) vehicle operation information - steering, accelerating, braking

d) Provide feedback.

i) identify correct and false detections

e) Provide incentive.

i) listen to Skinner on reinforcement (Annett, 1969)

ii) listen to others on the value of negative feedback and sanctions

Attention is a very common human function, and, like all human functions, it exhibits variability - mostly in the negative direction! However, an understanding of the nature of attention and its temporal partner, vigilance, can lead to successful engineering, operational, and training countermeasures.

**References**

Annett, J. A. (1969). *Feedback and human behaviour.* Harmondsworth, England: Penguin.

Drury, C. G., & Fox, J. G. (1975). *Human reliability in quality control.* Basingstoke, England: Taylor & Francis.

Drury, C. G., & Sheehan, J. J. (1969). Ergonomics and economic factors in an industrial inspection task. *International Journal of Production Research,* 7, 333-340.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors,* 37, 65-84.

Hancock, P. A., & Desmond, P. A. (2001). *Stress, workload and fatigue.* London: Erlbaum.

Mackworth, N. H. (1950). Researches on the measurement of human performance. *Medical Research Council Special Report* (No. 268). London: Medical Research Council.

Parasuraman, R., Masalonis, A. J., & Hancock, P. A. (2000). Fuzzy signal detection theory: Basic postulates and formulas for analyzing human and machine performance. *Human* *Factors,* 42, 636-659.

Sanders, M. S., & McCormick, E. J. (1993). Information input and processing. In *Human factors in engineering and design* (Chapter 3). New York: McGraw Hill.

Scerbo, M. W. (2001). Stress, workload and boredom in vigilance:A problem and an answer.

In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload and fatigue* (p. 276). London: Erlbaum.

Wickens, C. D. (2001). Workload and situation awareness.

In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload and fatigue* (p. 443). London: Erlbaum.

**Chapter 12**

**Newtonian Moments**

**(Winter 2004)**

The purpose of this article is to bring back happy memories to human factors/ergonomics

practitioners of their high school physics classes, and how some fundamental mechanical principles are very important in the way we analyze and design jobs. *Newton's Laws (Hibbler, 1983a)*

1. A particle originally at rest or moving in a straight line with constant velocity will remain in this state, provided the particle is not subjected to an unbalanced force.

2. A particle acted on by an unbalanced force, *F,* experiences an acceleration, *a,* that has the same direction as the force and a magnitude that is directly proportional to the force. If *F* is applied to a particle of mass *m,* this law may be expressed mathematically as *F=ma.*

3. For every force acting on a particle, the particle exerts an equal, opposite, and collinear reactive force.

**Rugby and Needles**

In rugby union football, the number 9 player (the scrum half) is usually small and the number 8 player (the loose forward) is generally big. One cold day about 40 years ago, a number 9 put the ball into the serum and ran around to the back to pick it up. He pretended to throw the ball to his colleague (number 10) but then quickly turned to sneak around the "blind side" of the serum. As the serum had rotated, the number 8 of the defending side was hidden from the eagle eye of the referee and was able to break a little too quickly from the serum. The small number 9 ran nimbly, but the tall number 8 stuck out his stiff left arm and caught the small number 9 around the neck. Now the small number *9's* legs kept running but his upper body did not, so he fell on his back with a hard thud. This small particle, mass ml' originally moving at a constant velocity VI (or perhaps accelerating a little), was met by a large particle, mass *m2,* and history is blurred by posttraumatic amnesia. The pertinent law here is that of the conservation of linear momentum, .

The small particle was a slow learner and often lacked situational awareness.

A short time later, while playing touch rugby (that's like flag football), he madea feint to the right, side stepped back to the left, and was free and clear until he ran into the goal post. This equal and opposite force led to an even more traumatic demonstration of Newton's laws.

After stitching up his split lip, the ER nurse instructed him to drop his trousers for a tetanus injection. If any of you know how to insert a sharp object through skin, subcutaneous tissue, and a couple of inches of gluteus maxim us, you will realize that it is important to impose a high velocity on the small needle in order to minimize the pain. But the nurse, who was a professional colleague of this captive patient, deliberately ignored her physics lessons.

**Friction**

Another mechanical principle that was demonstrated by the needle incident is that of friction, which is helpful in stopping things. The amount of frictional resistance to motion depends on two things. First there are the characteristics of the two surfaces. The coefficient of friction, µ takes on a value between 0 and 1, or a little more if we introduce adhesion into the discussion. Also there are the normal forces (N) applied to the objects in question; thus the limiting static frictional force, *F* = µN. Ice has a pretty low value - the coefficient of friction between metal ice skates and ice is about 0.04, whereas that between two pieces of wood is about 0.5. If we rub two pieces of aluminum together, we may see coefficients as high as 1.5. I suspect that the frictional resistance to a needle penetrating soft tissues is somewhat higher!

Slips and falls, because of low-friction surfaces, cause many accidents. Trips and falls, on the other hand, are caused by large moments. Runners are very sensitive to these issues; when a heel going fast hits a piece of ice, the shear force may be greater than the frictional force, so the trick is to use short steps on slippery surfaces. Efficient long-distance runners shuffle along and don't lift their feet very high off the ground. Potholes have been known to act as a temporary fulcrum, resulting in a large moment arm and grazed knees and elbows, which lead to more experience with the friction of tetanus shots in the ER.

**Forklift Trucks**

How long does it take to stop a forklift truck and how far will the truck have traveled before it stops? First of all, the driver has to be able to recognize that it would be wise to stop. This decision is hampered somewhat by high loads or broad "masts" while going forward and the reliance on peripheral vision while backing up. This visual part of the story is sometimes amplified by poor factory layout and lighting and the fact that most forklift truck drivers have high seniority (age) and often deteriorating eyesight.

The next part of the reaction process is moving a heavy boot to the right pedal (unless a rocker pedal is installed). Then Newton and friction and a few other mechanical and fluid principles take over. If the floor is composed of oil-soaked wooden blocks, there is a coefficient of friction between 0.1 and 0.3 between this floor and the solid rubber wheel. But forklift trucks, especially loaded ones, have a big N(the normal force that is a sum of the weight of the vehicle and load). Now we come to the important issue: the brakes. When the driver puts his foot on the brake pedal, brake fluid is compressed in a cylinder, which in turn pushes a high-friction pad or shoe onto a shiny metal brake drum directly connected to the wheels. The effectiveness of this truck-stopping system is related to the surface area of the contact between the brake pads and the drum and the amount of normal force that can be generated by the hydraulic braking system. The matter of rolling resistance complicates the discussion somewhat.

There are three more factors that compound this problem of stopping a forklift truck. First are the mass and speed of the truck. Of course, there are speed limits and governors, but those are for the other guys. We now have to move into part 2 of Hibbler's book *Engineering Mechanics -* dynamics (l983b). As this big vehicle is trucking along at 8 to 10 miles per hour, it encounters linear momentum, which is defined as the mass *(m)* of the vehicle times its velocity *(v).* We also have to consider impulse, impact, deformation, elasticity, and restitution - mechanical and financial! Now the brakes really have to work (assuming the vehicle doesn't skid on the slippery floor); they have to overcome this combined effect of mass and velocity. The law of conservation of momentum suggests that the small mass and velocity of the unwary pedestrian who just stepped into the aisle in front of the truck is not going to help much; it's back to number 9 versus number 8 on the rugby field, only more so. The second factor is when the floor slopes and this mess of mass is moving with increasing momentum (mass x velocity) downhill. The third factor is more insidious and controversial. Suppose instead of carrying the load on the forks, the truck is towing a train of trolleys loaded with heavy components. So the normal force over the forklift truck's wheels is reduced, thus reducing the frictional effect, while the mass of the material in the trolleys is increasing the momentum. The brakes of the truck were not designed to deal with this double whammy, so the truck will take a little longer to stop and travel a little farther in the process, which is just fine, providing it does not meet Newton standing in the way with a small mass.

**Moments**

Newton's findings were important, but as these examples show, there are complications that have to be dealt with by other mechanical principles. For the tall number 8 to put the impudent number 9 on his back, he had to impart a turning moment to the small particle by aiming the stiff arm some distance from the small particle's center of mass. Many sports injuries are caused by a combination of Newton's laws and the mechanical advantages offered by moment arms. Some, like skiing into a tree or heading a soccer ball, are direct examples of Newton's principles. Games played with racquets, bats, feet, and hands make great use of moments. How can a golfer hit a drive over 300 yards?

Long before Newton - a few hundred years BC to be exact - Archimedes was relaxing in his bathtub, with water splashing over the side, thinking about how to move the Earth. Being an empiricist, he suggested getting a very long piece of wood, anchoring it under the Earth and over the Moon, and then getting his buddy Phidippedes to run along the lever until his weight was sufficient to move the Earth. Archimedes discovered that a moment is calculated by multiplying the force by the length of the lever arm. But would Phidippedes have been weightless when he got to the end of the lever?

If I lift a 20-pound box close to my body, the moment will be on the order of20 pounds x 10 inches = 200 inch pounds. If I have to reach to the center of a pallet, then the moment might be on the order of 20 pounds x 30 inches = 600 inch pounds. Meanwhile) my back muscles, which are about 2 inches behind the fulcrum (a lumbar disc), will need to pull with a force of *600/2* = 300 pounds, and sometimes they may resent this imposition.

**Newton the Astronaut**

Sir Isaac Newton, who lived between 1642 and 1727, was an astronaut. He must have been, because his laws work when there is not much gravity and not much friction due to the drag of the atmosphere. About 240 miles above the Earth's surface, the velocities and the masses are big. The International Space Station is zooming along at about 17,500 miles per hour, and the driver of the space shuttle or *Soyuz* is trying to catch up and dock at the station. This is all right if the driver can see where he is going and has practiced in the simulators back at the Johnson Space Center. But when *Progress* comes along, the cosmonaut in charge of docking -like the forklift truck driver - may not have a very good view, and this may result in another case study on momentum.

Another space job is to stick out a remote-controlled, multijointed arm and

catch a satellite that weighs a few tons on Earth and will do funny things if it gets an off-center Newtonian nudge. Remember the number 8 who caught the number 9 some distance from his center of mass? Once one of these satellites is spinning, it takes a well-trained rodeo star to stop it (that may be why they put the Johnson Space Center in Texas).

Recently, we at the center were asked to advise on the location of handles on large pieces of equipment that had to be manipulated in and around the International Space Station. Just imagine a Newtonian nudge with a good lever arm setting a large rack careering around the confined interior of an ISS module. These astronauts are quick learners, and it is a joy to see how they maneuver large objects with a little push of a foot here and a little pull of a spare hand there. The trick is to get all the handles and foot and hand restraints in the right places to allow maximum accessibility, stability, and mobility- this is where biomechanical analyses and models come into play.

**Newton and NIOSH**

Back down on Earth, we have manual materials handling, bad backs, NIOSH, OSHA, insurance, and politics. In the middle of this mess we have the industrial ergonomist. The consumer wants to send big parcels through the mail or courier service. Household appliances and furniture are, for functional reasons, big and heavy; buildings are built of big components; batches of produce on the farms and in the food and drinks industry are somewhat heavy. And speed *(v)* is always a priority because time *(t)* costs money ($).

The first job of Newton the ergonomist is to reduce the mass of objects whenever possible, and the next job is to reduce the lifting moment arm. When things get moving, velocity (acceleration and jerk) is often bad, and more rather than less friction can be useful, especially between feet and floors. When one puts all these Newtonian nuggets together, one realizes that the delicate junctions between tendons and ligaments and bones and muscles line up to see whose turn it is to fail when the foot of a manual materials handler slips and a jerk runs through a mass of moments to be stopped by failure of the weakest link in the biomechanical chain. That is when stress and strain lead to an increase in tension.

**Conclusions**

This mechanics stuff that Newton started is fun. Engineers use it to design machines. Biomechanists use it to develop analytic models. Ergonomists can apply it to games, syringes or vehicles, manual materials handling, assembly operations, and space operations. But for ergonomists to be able to win the confidence of their engineering customers, it is important that they learn something of the language and principles of statics and dynamics.

**References**

Chaffin, D., Anderson, G., & Martin, B. (1999). *Occupational biomechanics* (3rd ed.) New York: Wiley.

Hibbeler, R. C. (1983a) *Engineering mechanics:Statics* (3rd ed.). New York: Macmillan.

Hibbeler, R. C. (1983b) *Engineering mecl1allics: Dynamics* (3rd cd.). New York: Macmillan.

**Chapter 13**

**Expectancy and Compatibility**

**(with Robert Schlegel, Spring 2004)**

This article is about how people behave when faced with a system interface based on their expectancies. It also addresses the compatible design of interfaces to facilitate desirable behavior and performance by both expected users and possible misusers.

**Introduction**

When we drive along the right-hand side of the road (in the United States), we expect our fellow drivers to do the same. When we walk into a dark room, we flip the light switch up to turn on the lights. When we want hot water in our hotel room, we turn the left-hand tap counter clockwise. When we want to move between modules of the International Space Station, we glide forward, do a half roll combined with a yaw to the right and a backward pitch and then move forward again, only to find that the "floor" has moved relative to our expectation. Some expectancies are stronger than others.

Expectancy is the result of peoples' lifelong learning about how the natural

and built worlds respond to our actions. Compatibility, in the broadest sense, describes characteristics of the design of the input-output relationships of the natural and built worlds that is the basis of our expectancies and behaviors. Compatibility comes in many dimensions, starting with spatial and temporal association, followed by logical and familiar associations. Finally, special associations are acquired by repeated practice (Kantowitz & Sorkin, 1983; Sanders & McCormick, 1993).

**Mistakes**

Why do people make mistakes, and what happens when they do? Sometimes we make mistakes because the natural and built worlds don't always respond consistently to our inputs - for instance, when we travel to England, where switches go down for on and cars are driven on the left. On other occasions, our perception of the world around us is flawed because we didn't make the most of the information available to us, as in reading the instruction manual.

The first result of our mistake is that things go wrong - sometimes very wrong depending on the amount of energy or information released by our actions. The second result of our error is that we may modify our expectancies for future occasions; in other words, we may learn from our mistakes (Annett, 1972; Johnson, 2004). However, sometimes we are stubborn and persistent - we repeat our mistakes because our expectancies are so strong. As the saying goes, "Fool me once, shame on you; fool me twice, shame on me!" As people get older, their expectancies and response sets become so ingrained that they may ignore the available evidence and feedback and continue to make the same mistakes over and over again.

The third result of our error, if we're lucky, is that we recognize what went wrong and make a hasty correction. If we're unlucky, in the stress of the moment, our "correction" may simply compound the original error and create some novel learning opportunities.

**Designing around Expectancies**

The first step in the design process is to study how the natural and built worlds behave. The second step is to get the knobs and dials in the right places. The third step is to have those knobs and dials move in the right direction. The fourth step is to have them move the right amount. The fifth step is to check the design with a bunch of strangers and get ready to call an ambulance or the Institutional Review Board when things fall apart. The sixth and final step is to write an article for *Ergonomics* in *Design* about our experiences.

A popular description of human factors/ergonomics is "knobs and dials engineering." This cryptic definition may be taken in various ways or may be considered too simplistic. But if one is liberal in the definition of knobs and dials by taking it to reflect all the means and conditions under which people receive and communicate information from and to the built world, then it is an acceptable description. For example, a dial or display is the device that conveys information to the user about the status of the system of interest - a car windshield, telescope, or computer screen is a display. Similarly, a gear lever, handle, harness, or keyboard is a knob (or control).

The cognitive interface is a bit more difficult to comprehend, but it is of great importance. It is the fundamental relationship between the knobs and dials as they reflect the status of the system of interest, the links between the user's eyes and fingers and the likely effects of their inputs. Think of the *and* as this cognitive interface, and one can put these semantic wanderings to rest.

Some aspects of behavior of the physical world were described by Newton. Things fall down; big things are hard to push; the harder one pushes, the harder the thing pushes back. These relationships lead to strong expectancies. A distant relative of Newton - "CogNewton" – developed the following laws of mental inertia:

1. A brain originally at rest will remain in this state providing the brain is not acted on by an unbalanced idea.

2. A brain acted on by an unbalanced idea experiences a change that has the same direction as the unbalanced idea and a magnitude that is directly proportional to the unbalanced idea.

3. For every unbalancing idea acting on a [boss'] brain, the brain exerts an equal and opposite idea.

Other aspects of our perception of the world are that differences are relative and our perceptions are nonlinear. Weber got us going on that one (Peacock, 2002). The observed behavior of the built world is more problematic because human designers are not always students of ergonomics.

One reason for apparent design errors is the sheer complexity of contemporary systems. It may be easy to design a cook top or car instrument panel, but try designing a nuclear power plant control room, space shuttle, VCR, or PDA. The laws of compatibility and expectancy appear

to be less robust and, like the tax laws, difficult to implement.

Fortunately, however, design principles that are derived from natural-world experiences and lessons learned from built world failures can be brought to bear on future designs. Journals and textbooks in the human factors/ergonomics field (e.g., Kantowitz & Sorkin, 1983; Sanders & McCormick, 1993) are full of these lessons, which include the following:

Rule 1. Put the knob next to the dial.

Rule 2. On similar systems, always put the knobs and dials in the same relative locations.

Rule 3. Arrange the knobs in the same pattern as the associated dials.

Rule 4. Group together knobs and dials with similar functions.

Rule 5. Put important and frequently used knobs and dials in the middle. This may lead to a conflict, which can be resolved by asking the question, "How important or how frequent?"

Rule 6. Put knobs to be used by the right hand close to where the right hand is comfortable.

Rule 7. If the activity demands great precision and is critical (i.e., the implications of error are serious), accommodate the needs of left handers by control location adjustability.

Rule 8. Allow for adjustability in your design if you expect your user population to vary on other physical dimensions such as size or strength.

Rule 9. If your expected users are likely to be limited on sensory and cognitive dimensions, make sure that the knobs and dials are few in number and in the best possible place for users to see, and obey all the previous rules.

Rule 10. Test your design on a bunch of strangers.

**Turning the Knob in the Right Direction**

Generally, up is for on, right for increase, push for forward, and pull for backward. But remember that electrical and fluid conventions are contradictory. A number of rules describe the expected relationships between the direction of motion of a control and the expected effect. Sometimes these human behaviors are called *direction of motion stereotypes* (Wierwille & McFarlane, 1993). Unfortunately, when the function, location, and orientation of the controls vary, for whatever reason, these stereotypes sometimes become weaker. This leads to a need for

more detailed qualification of the so-called stereotype.

A strong stereotype is one that results in the correct first-time response on more than 90% of occasions, whereas a weak one results in errors on more than 50% of occasions. This talk of stereotypes raises some important issues. Is it reasonable simply to take the first response to an encounter with a new system as being reflective of the goodness of design, or should one consider the rate of learning over a sequence of encounters? On many occasions, people adapt well to apparently less-than-satisfactory designs, with the result that the human factors input regarding the so-called correct design is seen as unimportant. The ergonomist may say, "Ab, but how will the user behave in an emergency or while under stress?" to which the designer replies, "Don't worry about that task analysis trivia; the sensible user will learn quickly and the others will confirm Darwin's theory of evolution."

The second problem facing the purist is that the variation in function, location, and orientation plays havoc with the simple theory, leading to weaker stereotypes. If the plane of motion is rotated in relation to the operator's reference (anatomic) plane, then first-time users will be less likely to choose the correct response on the first trial. It is sometimes easier to handle 90 degrees than 45, 135, or 180 degrees. Did you ever try undoing a screw on the back side of a panel while standing in front of the panel? The good news is that if you put the control in the right place, you're less likely to have direction -of- motion difficulties.

Other ramifications of direction-of motion stereotypes are the fundamental problems of individual differences in experience and therefore expectancy, the effects of the conditions under which the transaction takes place, and the seriousness of the possible outcomes. The theoretical definition of a strong stereotype acknowledges that mistakes may be made by some of the users some of the time. We may accept a mistake 1 in 10, 1 in 100, or 1 in 1000 times. But our acceptability of the performance of a design will depend on the severity of the outcome and the frequency with which the transaction is made. The more severe the outcome, the greater the amount of time stress, and the greater the frequency of exposure, the more important it is to get the design right. Weak stereotypes generally fail with inexperienced users under physical, cognitive, or temporal stress.

**Time and Recovery**

Studies of human reaction time from Donders in 1869 and onward suggest that it takes about a quarter of a second to see a light and respond by pressing a button. Temporal expectancy created by appropriate training, warning signals, and payoffs can reduce this time to zero or even less. But, as the Hyman-Hicks law tells us, added complexity - the number and difficulty (probability) of choices - causes an increase in the observed times (Sanders & McCormick, 1993, p. 50). When there is no external time limit on the response, users will generally impose their internal time limit when confident that they have used all the available information and weighed all the consequences of success or failure.

One danger of taking too much time to respond is that is provides the opportunity for false logic or erroneous calculations to creep into the decision. Unnecessary complexity can be added by designs that don’t obey the rules of compatibility and thus reduce the strength of expectancy.

The design philosophy of organizations that involve substantial amounts of energy and information with potentially catastrophic outcomes is that we need to design in a safety net. In fact, triples redundancy is a requirement of some design processes.

One way of creating pseudo-redundancy in human-system interactions is to use feedback and time to provide that opportunity for recovery. Another way is to use two or three pairs of eyes and their associated fingers. Contemporary computer interfaces often ask, “Are you sure?” and provide and “undo last” facility.

Control gain can be very important. The greater the gain – the size of the effect caused by the amount of input – the faster will be a response. But if the fain is too great, the response may be inaccurate, with control being achieved – if we are lucky – only after a series of overshoots. If the fain is too small, we may try to be more accurate, but it may take all day to achieve the desired effect. When time is of the essence, all day might be a problem.

**Checking out the Design with a Bunch of Strangers**

The danger with usability trials, especially with small samples of self-selected volunteer users, is that you may get the wrong answer. If possible, your design should be based on first principles, with the bunch of critical strangers being brought in for fine-tuning. There are good and bad things about strangers in usability trials. If selected according to representativeness of the user population, they will produce useful evidence most of the time. Friends and family, however, will be your greatest critics. Also, usability trials should involve the devil’s advocate and extreme individuals (Granddad, or Joe after he’s had a few beers). Under these circumstances, you may get evidence that will help you avoid problems down the road. Remember, we should design for expected use and foreseeable misuse.

A facilitator is a temporary device – such as a label or instruction – that allows the new or infrequent user to make the right choice. Instructions are not always used, often because of spatial, informational, and temporal inconvenience. The car driver usually bases his or her responses on expectancy rather than reading the owner’s manual. The average VCR operator may be more likely to read the book first; the PDA user improves his or her expectancy by working through the tutorial.

Labels act like fast instruction, but because of spatial constraints, they are often so cryptic that they become worse than useless – even counterproductive. Contemporary labels use acronyms, abbreviations, alphanumeric and color codes, and wonderful icons. In a 1980s parking lot study of common car icons, some were understood less than 10% of the time!

**Conclusions**

Compatibility in design is the basis of expectancy and performance. The design rules are clear for many situations, but when complexity increases of the context imposes constraints on design, it is usually wise to conduct usability trials to predict expected use and possible misuse.

**References**

Annett, J. (1973). *Feedback and human behavio*r. London: Penguin.

Johnson, A. (2004). We learn from our mistakes – don’t we? *Ergonomics in Design,* 12(2), 24-27.

Kantowitz, B., & Sorkin, R. (1983). *Human factors*. New York: Wiley.

Peacock, B. (2002). JNDs, SDs, HSDs, and DNDs: The Weber-Fechner Law. *Ergonomics in Design*, 10(3), 4-5, 30.

Sanders, M., & McCormick, E. (1993). *Human factors in engineering and design* (7th ed., pp 302-314). New York: McGraw-Hill.

Wierwille, W., & MacFarlene, J. (1993). Role of expectancy and supplementary cues for control operation. In B. Peacock and W. Karwowski (Eds.), *Automotive ergonomics* (pp. 2990320). Longon: Taylor & Francis.

**Chapter 14**

**Hotter than Houston: Body Temperature Control**

**(Summer 2004)**

This article is about physical work, how the body reacts to the thermal environment, and how clothing, engineering design, personal protection, and exposure management can mitigate effects on comfort, performance, and safety.

**Units!**

Just this morning, I learned from the radio that if! drank a gallon of iced water, I would burn 100 calories in energy to heat it up to body temperature. So I tried to do the analysis in my head. If, on a hot day, the iced water is 39°F and the body's core temperature is 99°F, it is necessary to warm the gallon of water up through 60°F (about 33°C). I also learned that a calorie is the amount of energy needed to raise one gram of water from *14SC* to ISSC - the range of maximum density.

Nowadays we should be talking about joules, like the rest of the world, and recognize

that 1000 calories or 1 kilocalorie or 1 Calorie (in food terms) is really 4184 joules. So 4.184 joules (1 calorie) is the amount of energy needed to raise the temperature of one gram of water one degree Celsius. The weight of one gallon of water is 8.33 pounds, which equals 8.33/2.2 =

3.786 kilograms = 3786 grams. Therefore, to raise 3786 grams of water through 33°C, we will need about 3786 \* 33 = 124.938 Calories, which is only a handful of French fries more than the number I heard on the radio.

You may have realized by now that some of the difficulties of this subject area are caused by the variety of units of measurement (Astrand & Rodahl, 1986; Best,

1990; Johnson, 1991). The medical world seeks to resolve some of the complications by dealing in *mets,* which address energy consumption per unit of body weight.

**Work and Energy**

When we do physical work, we convert the energy stored in the French fries into mechanical and heat energy; some leftover energy is stored as fat. There are many ways of measuring this process, each of which has advantages and disadvantages and offers various levels of accuracy. It is possible to measure the mechanical work done or the amount of contraction in muscles. It is more convenient to measure heart rate or oxygen consumption to convert into energy utilization. However, for the purposes of this article, it is more relevant to measure body temperature.

Our body core temperature is kept relatively constant at 98.6°F (about 36° C) by balancing heat input from the external environment and the heat produced as a byproduct of mechanical work with the heat lost by radiation, conduction, convection, and .evaporation. This balance can be modified somewhat by our choice of clothing materials and by modifying

the effects of the environment by the use of fans, air conditioning, parasols, and limited exposure duration.

Our muscles are only about 20% efficient in the production of mechanical work, depending on the kind of work we are doing. That is, for every 100 Calories we burn by muscular contraction, about 80 Calories are converted into heat energy. When we run at 7 or 8 miles an hour, we use about 10 kilocalories per minute, of which 8 are in the form of heat. The 2004 Boston Marathon was run with an environmental temperature of 85°F. The median

finishing time was about 10% slower than usual, and the demand on the course medical services was greatly increased, because the heat regulatory system of many runners could not keep up with the heat production. Therefore, core body temperatures increased.

Put another way, if someone ran 26.2 miles in just under 4 hours - that is about 9 minutes per mile, or 6.66 miles per hour - he would be burning about 10 Calories per minute or 2400 Calories for the whole race, of which about 2000 Calories were in the form of heat. Long distance runners generally prefer, and perform better in, cool conditions – about 50°F. However, it should be noted that the winner's time in the hot 2004 Boston Marathon was the eighth fastest in more than 100 Bostons.

**Heat Dissipation**

Given that it was a hot day in Boston, there would be very little heat lost by conduction and radiation, some lost by convection (depending on how fast one ran and the wind direction), some by breathing out hot, moist air, and the rest by evaporation. Some of the runners may have lost a few pounds because their water intake did not keep up with that lost through sweating. For example, it is possible to have a net weight loss of 61bs (96 oz, or 2727 gm), despite an intake of 4 oz at every mile or a total of 100 oz or 2840 gm.

An additional problem with evaporative heat loss is that sweat also contains salts, which need replacing by salt tablets or sports drinks; otherwise hyponatremia (a severe drop in sodium in blood plasma) may occur. Running (or working) in Houston presents the additional problem of high humidity, which interferes with the essential evaporative heat loss. It should be noted that if your body temperature increases from its optimal level of 98.6°F (37°C) to 107°F-llloF (42°C-44°C), you may die (Centers for Disease Control, 1994).

Work physiologists have developed an index, called P4SR (predicted 4-hour sweat rate), to reliably quantify the physiological effects of work in hot environments. The National Institute of Occupational Safety and Health, Occupational Safety and Health Administration, and the Canadian National Occupational Health and Safety Resource produce many charts that describe the acceptable and hazardous thermal conditions of employment.

**Control of Body Temperature**

The control of body temperature involves the production, transfer, and dissipation of heat. By far the greatest source of heat production is through muscle contraction. Heat transfer within the body is to some extent by conduction, but mainly by the active control of blood flow from

the deep areas (digestive system and muscles) to the surface. The distribution of blood among these areas is controlled by the diameter of local capillary networks, depending on whether you are digesting French fries, lifting weights, or cooling down.

Once at the surface, the heat dissipates by radiation, conduction, and convection. When the heat load becomes too high, the sweat glands are activated to enable evaporative heat loss. Under extreme conditions of dehydration, the sweat glands may cease to function efficiently, which leads to an increased risk of hyperthermia. After Mile 20 in the 2004 Boston Marathon, I noted that my singlet, which had been soaking, was dry despite frequent water stops.

Barrow and Clark (1998) presented a contemporary review of heat-related illnesses. The etiology of these illnesses includes a wide range of predisposing illnesses as well as age (young and old) and obesity because of the ratio of volume to surface area. Heat-related illnesses range from minor discomfort through cramping and dizziness to damage to various tissues and organs described as heat stroke, which is characterized by a core body temperature of greater than 104.9°F. The general solution is to apply rapid external and/or internal cooling to bring the core body temperature down below 102°P and to replace fluids and other nutrients.

Observation of the human responses to work in hot environments -like most observations of human preference, behavior, and performance - is fraught with variability. By virtue of their body composition and shape, individuals may be more or less efficient at heat loss. In addition, people can become acclimatized to some extent to hot environments by behavioral and physiological adjustments, although older people are less adept at such adaptation. In some work contexts, such as mining, systematic acclimatization regimes have been shown to greatly reduce the prevalence of heat -related disorders (Wyndham, Rogers, Senay, & Mitchell, 1976).

A unique challenge is encountered by astronauts who do space walks. They are encased in a multilayered suit and exposed to very high radiative heat loads while carrying out physically demanding work. Consequently, their suits are equipped with an outer reflective layer and an inner water-cooling system, arranged to make maximum use of those

body areas that are most efficient at heat transfer from the source (the muscles) to the surface (Kosheyev, Leon, Coca, & List, 2004). Chauhan (1988) presented a comprehensive review of the literature on cooling garments.

**Cold**

In space, on earth, or on or under water, people may be vulnerable to the effects of too little heat they may get cold despite their metabolic heat production. Wind chill may be a major culprit. In these cases, the solution is to use clothing as a buffer to reduce heat loss. SCUBA divers wear either wet suits or dry suits. Wet suits cover various amounts of the body surface depending on the temperature of the water. These suits trap a thin layer of water, which remains close to body surface temperature and thus insulates the body from the effects of the surrounding water. Dry suits are used for colder conditions; they cover most of the body and entrap a thin layer of air as the insulating medium.

Whereas these underwater suits are relatively thin and do not interfere unduly with mobility, the clothing worn for work in cold climates may cause considerable interference with performance. Cold interferes particularly with manual dexterity. One morning some 40 years ago, I saw a man, who was not wearing gloves on a very cold morning, fall backward off a ladder to his death. Running suits made of synthetic materials allow good protection and comfort in very cold conditions, whereas natural materials - such as animal skins, fur, and wool - despite their good thermal properties, may be too heavy, cumbersome, and restrictive. The international ergonomics community, through the International Organization for Standards, is working on establishing metrics for the insulating properties of a wide range of materials used in clothing.

**Measurement**

Ergonomists and physiologists have measured the thermal context of work for many years. Key parameters include air (dry bulb) temperature, wet bulb temperature (measured by a thermometer encased in a wet sheath), radiant temperature (measured by a thermometer encased by a black globe) and air movement. These basic parameters are combined with various weightings to develop indices such as relative humidity, dew point, heat index, wind chill, and wet bulb globe temperature. These derived indices are more indicative of the combined effects of the basic parameters noted earlier on comfort and performance.

More ambitious indices (effective temperature) also include the effects of clothing and metabolic heat load. They are often displayed as charts and nomograms with wide-ranging levels of complexity. Weather forecasters have addressed the challenge with enthusiasm and color. Whereas some of their attempts are very meaningful to a wide variety of viewers - by using umbrellas to indicate the likelihood of rain - they occasionally get carried away with their technology and produce obscure analyses that are of interest only to avid watchers of the Weather Channel. Various recreational organizations, such as SCUBA *(http://www.scubadiving.com/a rtide/ 0,7424,3-40-122-231,OO.html),* paddling

*(http://canoeman. com/SWPaddler/hypothermai.html),* and running *(http://www.teamoregon.com/publications/heat.html)* have also joined the cause of thermal environment analysis and guidance directed at the safety of their members.

**Design**

Engineers have addressed the foregoing challenges of thermal environment management with gusto. Nowadays it is possible to leave a climate-controlled house, climb into a climate-controlled car, and go to work in a climate-controlled office without the various aspects of the thermal environment ever changing, except when one uses the sauna to relax at the end of the day. The only problem comes at the end of the month, when you have to pay the utilities bills.

An example of how engineers sometimes oversimplify the problem of thermal environment management occurred a number of years ago in Hong Kong when I encountered an advertising campaign that touted the magic number 22.5 – the comfort zone. I delved into the background to this claim, particularly because 22,SOC*(nOp)* didn't strike me as much of a zone. I found through a very tortuous search that the participants in a controlled laboratory experiment performed better in and preferred this temperature.

I also experimented by measuring the temperature in a bedroom with an air conditioner of the recommended size for the space when the outside temperature was in the 90s. To my surprise, the drop in room temperature was asymptotic at about 25°P. Although I thought the temperature was comfortable, I was unable to achieve the so-called comfort zone even though I left the air conditioner on all day. I then noted that the advertiser of the comfort zone was none other than the Hong Kong Electric Company, which sold both air conditioners and electricity.

During the same period, I was serving as the ergonomics consultant to the Hong Kong Mass Transit Railway and had the opportunity to address the issue of the thermal environment in trains and underground stations. With summer temperatures in the high 90s, the cost of achieving the comfort zone would have been very high. So I developed an algorithm whereby the train temperature tracked external temperature, but without the wide fluctuations. I believe that the range was set between 700P and 80oP. The rationale for this cost-saving approach was a combination of human acclimatization over the seasons and the fact that people's perception of temperature is relative, not absolute.

Another alternative to the engineering approach of modulating the environmental context is by rationing the exposure of individuals to sub-extreme conditions. Given the opportunity, many of us like to alternate the outside heat with indoor air conditioning. Some industrial processes require the maintenance of hot conditions to cure paint or of cold conditions to preserve meat. The people who work in these industries may follow a systematic exposure process that permits some discomfort but maintains a safety margin.

The OSHA heat exposure chart below provides simple guidelines for exposure based on temperature and level of metabolic activity. It appears from this chart that an appropriate strategy for the 2004 Boston Marathon would have been 50% work and 50% rest each hour, although it may be expected that endurance athletes are better able to endure thermal stress than, say, the 5th percentile industrial worker. OSHA also provides guidance for work in cold environments: *http://www.osha.gov/plsloshaweb/owadisp.show\_document?p \_table=NEWS\_RELEASES&p\_id=82*

**Conclusions**

The human response to the thermal environment is well understood, and the mitigation approaches - including both engineering and administrative (exposure) controls - are relatively straightforward. However, heat and cold stress remains a prevalent problem because of the increased susceptibility of some individuals, the behavioral choices of others, and the

financial burden of large-scale thermal management. 

**References**

Astrand, *P.-G.,* & Rodah!, K. (1986). *Textbook of work physiology: Physiological basis of exercise* (3rd ed.). London: McGraw-Hill.

Barrow, M. W., & Clark, K. A. (1998, September 1). Heat-related illnesses. *American Family Physician.* Online at <http://www.aafp.org/afp/980901ap/barrow.htm>l.

Best, r. B., (1990) *The physiological basis of medical practice* (12th ed.). Philadelphia: Lippincott Williams & Wilkins.

Chauhan, D. T. (1986). *Liquid cooling garments.* Unpublished master's thesis, University of London. Centers for Disease Control. (1994). Heat-related deaths - Philadelphia and the United States. *Morbidity and Mortality Weekly Report,* 43( 45), 453-455.

Johnson. A. T. (1991). *Biomechanics and exercise physiology.* New York: Wiley.

Koscheyev, V. S., Leon, G., Coca, A., & List, N. *(2004). Comparison of different cooling regimes within a shortened liquid cooling/warming garment on physiological and psychological comfort during exercise.* Presented at Habitation 2004, Orlando, FL.

Rogers, G., Goodman, c., & Rosen, C. (1997). Water budget during ultra endurance exercise. *Medical Science and Sports Exercise,* 29, 1477-1481.

Wyndham, C. H., Rogers, G. G., Senay, L. C., & Mitchell, D. (1976). Acclimatization in a hot, humid environment: Cardiovascular adjustments. *Journal of Applied Physiology, 40, 779 785.*

**Chapter 15**

**Rule-Based Ergonomics**

**(Fall 2004)**

Initially this article may not be universally popular among many of my academic colleagues and some professional practitioners, but my mission is to confront them with some realities of practice in a world where the customer is not prepared to wait or pay for an in-depth investigation. Customers are also confused by "ergobabble." Those words that often threaten our profession - *common sense -* can be found to be useful, especially when the sense is indeed

common.

Now I'll backpedal a little on those fighting words and admit that sometimes more sophisticated approaches, requiring in-depth knowledge and technique, may be useful; a test of our judgment and integrity is when to employ these costly methods. The measure of our profession is that we must be effective (relevant/valid, useful, and sufficiently accurate), efficient

(timely and inexpensive), and easy to use (by our own rules and our customers' evaluation) .

**Everyday Rules**

The Ten Commandments and a recently popular sitcom aside, our lives, behaviors, and activities are governed by many simple rules and conventions. Drive on the right and less than 25 mph in a built-up area. Be in by 10 o'clock. Turn the right-hand tap counterclockwise for cold water; attach the red (or is it black?) wire to the "live" terminal; stop at the red light; place the drive-up window level with the car window; make doors a standard size. Put the important stuff in the middle of the display, make road signs big enough to read, put the accelerator pedal on the right.

The analytically minded will already have found some holes in these arguments. First, they will note that 10 o'clock is sometimes negotiable or forgivable; 25 mph may be stretched to 35 if no one is looking. In some minds, the red light means look both ways before stepping on the gas and the yellow one means go for it anyway. Hotel bathrooms don't count in fluids conventions. Does "drive on the right" apply to Grandma, who prefers the middle lane; exactly how far to the right should the gas pedal be? What is the level of a car or SUV window? What if there is a lot of important stuff to put on the screen - how big is big enough? And so on.

**Measurement**

Our first classes in statistics addressed the challenges of measurement, which is the basis of our communications. We learned about qualitative and quantitative, objective and subjective, random and systematic errors, and blunders. We learned to differentiate between nominal,

ordinal, interval, and ratio scales. A little later in our course, we learned that "standards" could refer to performance requirements (which need an interaction between a person and a system) or to design specifications (which focus on the hardware or software characteristics). Our communications with designers can involve any of these concepts, but we must be careful in our consistency and operational definitions.

The design specifications that a warning must be placed on a product or that a control movement should be clockwise are nominal statements. The performance requirements that a car should go from o to 60 in a shorter time than the competition is an ordinal statement, which may require a statistical test to evaluate. Performance requirements can be validated in controlled contexts. The design specification that a key on a personal digital assistant (PDA) should be 4 mm is quantitative and verifiable by anyone with the right tool and procedure.

Our communications are often clouded by variability. For example, the performance requirement that the user should be able to judge the difference between large and supersize fries requires an overlay of psychophysics rules, such as "on 50% of trials. " We are also challenged by our need to define *user* and how to obtain representative samples of users in our performance requirements evaluation trials.

Design specifications also face issues of variability, which often come in the guise of tolerances. In the PDA button size example, do we really mean 4 mm plus or minus 1 mm? It might be better to have a loss function with the absolute minimum being 4 mm with increasing, nonlinear penalties for diameters less than 4 mm.

**Our Customers**

The statisticians among us will note that some of these rules will suit some people better than others. Also, most of us are fairly tolerant of some deviations from the hard rules. Only National Basketball Association centers are worried by door and vehicle headroom heights. Even the Highway Patrol may allow us a bit of a cushion. And the character height problem can sometimes be resolved by two or three glances, and learning.

But the statisticians are not our customers; they are our defenders. Our usual customers are designers, engineers, and users, and they need clear and succinct guidance because they have a lot of other things to worry about. They do not need "Come back next year when I have completed my study, which will cost an arm and a leg:' Their response to this suggestion will be "common sense" or "engineering judgment." But the statisticians are often surrogates for our customers. They can answer the question, "How do I please most of the people most of the time?" Our universal design colleagues offer special challenges to the statisticians; often the maximum is even more elusive than the mean.

The traditional approach to design engineering is to establish a quantitative design specification and then add on a tolerance band. The manufacturer then chooses any point in the tolerance band to suit its other constraints. Then the quality control guy comes along to discuss the problem of tolerance "stack-up" – the summation of within-tolerance deviations

on multiple components - and everybody points to the other guy - *"My* piece is within the tolerance band."

Now we are getting to one of the realities of specifications and design standards:

They are rarely applied in isolation. Most products and processes are complex, and users vary. Only recently I attended a series of meetings on display design for the next-generation space shuttle cockpit, and each meeting addressed a separate dimension - background color, font size,

contrast, item location, symbol format, etc. I don't think future shuttle users will worry about these items independently, although the designer will.

Some time ago I sat on a relatively unsuccessful committee that aimed to standardize vehicle control locations - another integration problem. In 1981, a brave NIOSH committee saw the light when its members integrated the various factors associated with manual materials handling - and they're still dealing with the fallout, both political and operational.

**A Better Way - Common Currencies**

There must be a better way. There is, and I call it *rule-based ergonomics.* It involves a shift in the way we ergonomists give guidance to our customers. First, we do a process or task analysis and identify the desired outcomes and their trade-offs: effectiveness, efficiency, safety, customer appeal, etc. Next, we identify the constraints: mass, volume, cost, schedule, and resource utilization. In the space business, these constraints are measured in *equivalent mass* because getting stuff up there is a major challenge, especially with chemical fuels. I would prefer to use cost as the common constraint currency it's more universally understood.

But what about a common currency for the desired outcomes? How much bang for the buck? Monetary value would be a good idea because then we would have tractable mathematics; we could use either addition or multiplication of similar units. Perhaps we will have to settle for two different units. There are many precedents: miles per gallon, accidents per 200,000 hours, market share by advertising investment, errors per transaction. If we delve a little deeper, we may see that efficiency and safety are really constraints and can be converted to denominator units - dollars. Customer appeal (preference) is really a reflection of effectiveness (performance), although it may be measured by different methods. So the pure effectiveness measure is the probability of a successful transaction.

I realize that my psychophysicist colleagues will wish me to define *successful,* but I will be discrete (and discreet) for the time being. My effectiveness scale can be reduced to a common currency, which I will call green, yellow, red, or one to seven, or whatever practical level of resolution you like. A more detailed description of a universall0-point scale is described by

Peacock et al. (2004).

**Consensus**

The visual thinkers will now recognize that I have the beginnings of a graph, with cost on the abscissa and effectiveness on the ordinate. Most will be asking, "But how do you establish the model that relates the cost to the effectiveness?" Come back in a year when I have done a study. No, I can do better than that, but again, some of you aren't going to like what I have to say. My only defense is that it works - most of the time.

My process for linking abscissa to ordinate values is consensus. My detractors will say that I have thrown empiricism out the window, but I haven't. Because my consensus comprises a group representing all the interested parties (think of a courtroom, with [expert] witnesses, attorneys, a jury, and a judge), some who bring data, analyses, and models that describe their own perception of the problem at hand. For example, a couple of years ago, Houston changed its freeway speed limits from 65 mph to 55 mph, based on two constituencies - safety and the environment. But Joe Public came back with his own analysts and a whole load of "customer appeal/effectiveness" arguments, and the limit went back up again.

Now back to our more familiar domain - take the NIOSH lifting index and character height as examples. A reasonable range of values for the NIOSH lift index is 1 to 5, which reflects cost-related factors of workplace design, load and frequency, outcomes (constraints) of productivity and safety, and the ultimate objective of a successful transaction. For the character height on an automobile instrument panel, we have a range from about 2 mm to 10 mm, again related to the competitive cost of instrument panel "real estate," with a bit of aesthetics thrown in.

In the first example, the consensus group will represent managers, workers, engineers, doctors, some academics, or perhaps some customer representatives - lawyers, union reps, and ergonomists. They will all agree that 5 is bad, although their forefathers achieved a 10 while they were building the pyramids! Most will agree that 1 is probably all right, although some may push for less. Data will show that the curve is not linear, with diminishing returns at both ends. Nobody knows the exact form of the curve, however big his or her sample size. So all may agree that 3 or higher is red, less than 1 is green, and yellow is in between. The engineers in the room will point out that the composite index does not solve their particular problem. The product engineer is in charge of load, the manufacturing engineer designs the workplace, and the industrial engineer is responsible for frequency. So the consensus group will break the index into its constituents and follow the same process to establish cutoff points for the individual factors. The same game will be played out on the character height problem, but with other groups of stakeholders - designers, marketing representatives, engineers, and ergonomists as surrogates for their farsighted grandparents. Policy and Politics One more thing. What does *red* mean? The answer is not ergonomics but policy and politics, which are the province of management, the government, or negotiation. A good general rule is that red means stop or not allowed, green means always allowed, and yellow is usually allowed, unless there are too many "yellows" in the set of relevant measures. A practical rule would be to say that we cannot allow more than 3 or 5 or 10 yellows, given a set of between 10 and 20 criteria. Most decision processes will involve a finite set of criteria - in the case of the NIOSH lift index, there are 6, or a few more in some embellishments.

In display design there may be 20. In the space habitat environment design, there may be many more, including heat, light, noise, vibration, air, water, radiation, and a bunch of spatial variables. Each of these variables may be the province of different engineering groups, but it is up to managers to make the trade-offs. They need to see the forest and decide which trees should get the axe and which get the waiver. The forest can be conveyed by a matrix of reds, yellows, and greens, just like the leaves in fall, particularly if we expand the scale with some oranges and browns.

Engineers will note that the tolerance band has been replaced by a loss function. The green cutoff is that level below which there is unlikely to be any significant cost. The yellow/red cutoff is more elusive. A rule of thumb is that it is that level beyond which the individual constraint will incur a significant cost, independent of any interaction with other constraints. "Unsafe at any speed" got the attention of the carmakers and lawmakers. The yellow region is the region of uncertainty where interactions become important. For example, a low-contrast figure-background relationship will interact with ambient light levels, or the vertical height of a lift will interact with the horizontal reach. Another rule of thumb establishes the red cutoff as the point of inflection of the cost –effectiveness curve - giving it the approximate appearance of an exponential loss function.

There is another complication of rulemaking - human variability. I am safe at 90 mph, but *you* should drive at 50 mph. Perhaps my old eyes are not so good and my reactions are a bit slow, so I may need bigger controls and font sizes (how about 10 mm?). But because you are young and should not be distracted by a radio, you should see me coming and get out of the way. In practice, the consensus approach will address the level of conservatism in rulemaking to accommodate the majority of the population. Perhaps a green should have a 10th percentile limit and the red cutoff should be at the 50th percentile - these are policy matters.

**Stakeholders**

The advantage of the consensus approach is that all the stakeholders "bought into" the mapping model between the constraints and the effectiveness rating. Also, management made a rule that there would be no reds and only so many yellows. The purist will ask about the weighting of the different criteria. The weighting is accounted for in the mapping process - the common currency principle implies that a red is a red is a red! But what about interactions? Where there are known interactions, variables may be combined as in the NIOSH lifting index, but eventually intervention will usually involve separate variables. In practice, counting the number of yellows assesses the likely interactions.

So where does this leave the ergonomist? I argue that the professional ergonomist's job is to facilitate the establishment of consensus-based, context relevant rules, based on whatever data and domain experience are available. In the absence of hard (reliable) evidence, a hypothetical (commonsense) rule or rule set should be established, based on process or task analysis and evaluated by appropriate empirical research or systematic customer feedback.

The next task of the professional ergonomist is to train the stakeholders in the process, including loss functions, and solicit their participation in consensus rulemaking. But first management must be educated in this pragmatic approach to ergonomics design standards in the context of complexity and uncertainty.

**References**

Chen. J. G., Peacock, B., & Schlegel, R. (1989). An observational technique for physical work stress analysis. *International Journal of Industrial Ergonomics,* No.3.

Peacock, B. (2002). Measurement in manufacturing ergonomics. In S. G. Charlton & T. G.

O'Brien (Eds.), *Handbook of human factors testing and evaluation.* Mahwah, NJ: Erlbaum.

Peacock B.• Blume, J.• & Vallance. S. (2002. June). An index of habitability. Presented at the International Conference on Environmental Systems. Society of Automotive Engineers, San Antonio.Texas.

Peacock, B., Blume, J.. & Vallance. S. (2003). Habitability measurement in space vehicles. In N. Stanton (Ed.). *Handbook of human factors and ergonomics methods.* Basingstoke, England: Taylor & Francis.

**Chapter 16**

**Remember Hawthorne**

**(Winter 2005)**

This article is about the Hawthorne effect (the problem of confounding in investigation design and analysis) and some rules of thumb to prevent the ergonomics practitioner from jumping to invalid conclusions. Some parallels are drawn between our profession and the many other professions whose focus is human behavior and performance.

**The Hawthorne Studies**

In 1927, a group of Harvard researchers went to the Hawthorne, Illinois, plant of the Western Electric Company. Their initial purpose was to investigate the effects of lighting on productivity, but they observed that productivity increased whether the lighting was increased or lowered. The conclusion was that worker productivity was influenced to a greater extent by the fact of change and the attention by these prestigious researchers than by the technical workplace changes per se.

The so-called Hawthorn effect describes a positive response to these psychosocial factors followed by a decline to original levels as the workers become accustomed to the attention or the attention is withdrawn (Muchinsky, 1994). These studies became a classic in the development

of industrial psychology (Blum & Naylor, 1986) and the more recent field of macroergonomics (Hendrick & Kleiner, 2001,2002), although later critical reviews of the study methodology and results were less convincing.

The Hawthorne studies also have implications for the practice of human factors/ergonomics in the laboratory and in the field. Are our observations and derived conclusions attributable to the factors in which we are interested, or to some other (contextual) source of variation? In statistical jargon, this is called *confounding.* For example, we may conduct a psychophysical investigation of high-frequency lifting using two different containers or methods and conclude that Container A is better than Container B - until we realize that those using Container A were on average younger and stronger than were those using Container B. A similar comparison of alternative vehicle navigation system interfaces may be invalidated by the different traffic conditions in which each was evaluated. Productivity studies of workers in different countries or with different socioeconomic contexts may lead to very different conclusions.

In the design of investigations (experiments), there are four principal sources of interdependent variation: the factors in which we are interested, other factors that are known to influence behavior and performance and are controlled in the experimental design, important factors that we cannot or do not control, and the ever-present random or residual factors. The Hawthorne studies drew attention to a particularly insidious source of confounding: psychosocial and motivational factors.

**Beyond Hawthorne**

The Hawthorne effect is often put to positive use. For example, athletic coaches may provide some tangible input that changes the individual's behavior and subsequent performance, but they also provide a motivational environment that encourages the athlete to practice harder. Similar observations may be made with sports, industrial, entertainment, and business teams in which the contribution of individuals is enhanced by team cohesion. Similarly, the empirically demonstrable home court advantage in sports is a complex effect of tangible and intangible

factors.

It should be noted, however, that the Hawthorne effect does not necessarily imply a permanent change to individual or team behavior. Individual, company, or team failure, caused by the simple fact of more talented competition, may erode the initial motivational advantage and indicate that it is time for change - a new coach or CEO.

Frederick Herzberg, an industrial psychologist, addressed the matter of job motivation by developing a two-factor theory (Herzberg, Mausner, & Snyderman, 1959). He indicated that workers were positively motivated by factors intrinsic to work, including feedback of success, but were demotivated by dissatisfiers or "hygiene factors." Since that time, a large number of studies have supported many contemporary production and work systems that seek to combine intrinsic and extrinsic factors affecting human behavior and performance. The classic applications have been in the automotive industry, where the quest for high levels of quality and productivity has led to short job cycle times and very mundane work. In the 1980s, Volvo and Saturn developed team structures that included job enlargement and rotation. The Toyota production system also makes considerable use of team structures.

The early paternalistic view of these approaches was sometimes euphemistically termed "job enrichment," but the more objective researchers recognized that this is a personal judgment and not necessarily something that can be imposed by an ergonomist or industrial psychologist. Corlett (personal communication, 1998) pointed out that the role of the ergonomist might be that of activist. In my first career (as a physiotherapist), I quickly learned that my influence went beyond the physical, both in industrial and sports rehabilitation. Even very large and sophisticated implementations of motivational theory have had their problems when the novelty wears off and time for change is indicated in the form of a new guru or system. This is not to say that popular gurus do not offer good advice; rather, that their influence may sometimes go beyond their objective message and decline over time.

**Some Case Studies**

In the 1970s, I was invited to address some perceived productivity and motivational problems in a bookbindery that was attached to the University of Hong Kong library (Peacock, 1979). I implemented many equipment, environmental, and work flow changes, including the development of a team structure. I also noted that the atmosphere of the lesser-paid blue-collar (skilled craftsmen) workers in a white-collar context led to some unreasonable demands for service. The results of the substantial objective changes, which also included level and pay increases, indicated a considerable productivity increase, which persisted for at least six months. On reflection, I may have been a crusader armed with contemporary techniques who

was reluctant to remember Hawthorne. Or maybe not.

In 1980, the Dartmouth, Nova Scotia, police force was considering a change to a 12-hour, 8-day shift system from an 8-hour, 12-day system. I was asked to investigate the effects of this change over a trial period and advise accordingly (Peacock, Miller, Glube, & Clune, 1983). We were able to carry out a broad spectrum of physiological, psychological, and psychosocial tests a couple of months before the change and six months after the change. Because of the nature of police work, it was not possible to evaluate direct performance changes.

The physiological (blood pressure and exercise tolerance) measures showed minor improvements. The psychological (grammatical reasoning and critical flicker fusion frequency) also showed minor improvements. Behavioral measures (sleep duration and quality) showed major improvements. However, all these measures were swamped by the overwhelming support of the police officers for the new system, which gave them longer continuous time off between shift rotations and better compensation time for such things as court appearances. Needless to say, the new system was adopted for psychosocial reasons in the absence of any significant negative physiological or psychological or performance changes.

In another shift work situation, the workers at the Saturn Corporation plant adopted the worst possible shift pattern from the physiological point of view: one week of days alternating with one week of nights. The argument was that this created a "fair" distribution of nights and days.

I got another chance to practice this aspect of my profession in the mid -1980s, this time with the Oklahoma Department of Health and Human Services. The context was a very large open-plan office housing some 200 workers who processed Medicare and Medicaid claims. The workflow was through a series of functional groups with very specialized and routine work tasks. The open-plan environment and endemic mainframe computer glitches compounded problems of the classroom-like layout, where the supervisors sat in front of rows and columns of middle-aged, overweight, and often unhappy "students." There was also pressure from above arising from complaints about errors, productivity, and insufficient funding for these federally mandated

programs.

My silver bullet was to introduce a prototype team structure made up of people from different departments, with each team being responsible for a much greater number of the necessary steps in the process. Again the intervention was a resounding success. Productivity increased and errors were greatly reduced in the pilot team, which soon became both the envy of the whole department and also the "go to" group for problem solving. Plans were under way to extend the team structure throughout the whole department, but in their wisdom the powers

that be decided to outsource the whole program to a private sector company.

In the 1990s, the productivity and quality wars in the automobile industry were joined by massive attention to the epidemic of cumulative trauma disorders, which spread like wildfire through the food-processing industry, nursing homes, the office, and the millions of people whose occupations and recreation were (and still are) tied to a computer and often a telecommunications device. Large manufacturing organizations, the unions, and the Occupational Safety and Health Administration (with the help of the National Institute for Occupational Safety and Health) formed often acrimonious associations in the battle between production and protection (Peacock, 1994a, 1994b, 1994c)

The universal treatment was ergonomics, with engineering controls being favored over administrative controls. The engineering controls consisted of attention to workplaces and tools, particularly to that most personal item, the chair. Administrative controls generally addressed interventions such as job rotation and job enlargement to address the frequency or exposure factor, because the demand for productivity as dictated by line rate and a lean philosophy was untouchable. The unions recognized that reducing exposure to the mechanical (posture and force) stresses by adding labor was an attractive alternative, until the widespread practice of exporting jobs to regions with lower wage rates and younger bones became an even greater threat.

It is very likely that the epidemic of cumulative trauma disorders was partially the result of a sort of inverse Hawthorn effect, which greatly lowered the threshold of reporting. Cumulative trauma became an attractive target for the health and safety community and the unions, and the deluge of educational material perhaps made some workers more sensitive to physical discomfort. This is not to say that cumulative trauma disorders are not real, as some opponents of ergonomics legislation implied. They are very real, and the causes are very clear (posture, movement, force, interface, frequency, and duration). But the level of pain, impairment,

and reporting is very variable, thus opening the doors to the influence of psychosocial factors. Objective observers of this focus of our profession over the past few decades will recognize the escalation through physical and psychosocial influences to the highest level of politics, with ergonomics researchers and practitioners sometimes being unable to dissect the independent influences on behavior, performance, and worker health.

This battle for control of the job cycle was and continues to be fought on the battlegrounds of competition and productivity on the one flank and health and safety on the other. The solution is often "participatory ergonomics intervention programs" (NIOSH, 1994) that mandate worker involvement as a primary component. The premise of such programs has two components. First it is argued, quite reasonably, that the line worker or computer operator does have a very personal interest in and knowledge of the physical aspects of his or her job and what works best or hurts most. Second, the fundamental psychosocial issues identified in the Hawthorn studies of attention to the worker, now formalized through participatory programs, will certainly influence job satisfaction, at least temporarily.

**Other Contexts**

The $64,000 question is the longevity of these programs, given the contentious political context and the undeniable pressure for labor productivity in the competitive marketplace (Imada, 2001). The growth of organizations (and nations) depersonalizes individuals, who sometimes reluctantly delegate their destiny to their elected or imposed leaders. When the result is success, halos grow; when the pay is low or their arm is hurting, then Hawthorn goes into reverse.

Such intervention strategies are also found in other areas of society. Recently I watched an interview with Dr. Phil on the television. He and his team of down-to-earth change makers were going into a small town to address the social problems that are endemic throughout the United States - poverty, crime, drugs, teenage sex, family violence, school dropouts, and so on. He used a participatory approach that involved the town's leadership, including government, law enforcement, business, education, community organizations, and the individual townsfolk - on both sides of the fence. Will this massive Hawthorn study be successful? Will it lead to tangible change beyond the psychosocial hype that accompanies this charismatic guru? Will it last when the attention goes away? Even if it doesn't last very long, it will provide even greater evidence that disenfranchised and discriminated people will rebel and that the intervention must include both attention to the individual and to change in the primary aspects of the individual's education, work, and environment.

Unfortunately, as in many areas of intervention in complex problems, it is practically impossible to employ classical experimental statistics to dissect the precise cause-and-effect relationships that we observe in the laboratory. But this begs the question of whether the professional practitioner (social worker, doctor, ergonomist) should wait for the researcher to use classical formal methods to prove that an intervention or simply attention to a situation is worthwhile.

**Conclusions**

What can be gleaned from the Hawthorne message and the enormous knowledge that has grown over the past 75 years? The first message is that cause and effect may not be simple and that an ergonomist who prescribes a $1000 chair may be addressing more than the operator's posture. Also, ergonomists who evaluate the effectiveness of programs must address the importance of honest investigation designs that avoid confounding. Next, they must recognize the inextricable interactions between primary and psychosocial or motivational factors. The fourth message is that people thrive on attention that comes in many guises and that the ergonomist may often mimic the role of the family practitioner whose task is to treat the person as well as the disease.

The most important message for ergonomics researchers and practitioners is that they must not rely on dogma but, rather, recognize the less tangible influences of customer participation and their own role as activists.

**References**

Blum, M., & Naylor, J. (1986). *Industrial psychology: Its theoretical and social foundation.* New York: Harper & Row.

Hendrick, H. W., & Kleiner, B. M. (2001). *Macroergonomics: An introduction to work system design.* Santa Monica, CA: Human Factors and Ergonomics Society.

Hendrick, H. W., & Kleiner, B. M. (Eds). (2002). *Macroergonomics: Theory, methods, and applications.* Mahwah, NJ: Erlbaum.

Herzberg, F., Mausner, B., & Snyderman, B. B. *(1959). The motivation of work.* New York: Wiley.

Imada, A. (2001). Participatory ergonomics in Chevron. In H. W. Hendrick & B. M. Kleiner (Eds.), *Macroergonomics: Theory, methods, and applications.* Mahwah, N}: Erlbaum.

Muchinsky, P. M. (1994). *Psychology applied to work* (4th ed.) Belmont, CA: Brooks/Cole. National Institute for Occupational Safety and Health. (1994). *Participatory ergonomic intervention in meatpacking plants* (US DHHS 94 - 124). <http://www.cdc.gov/niosh/94> 124.html

Peacock, J. B. (1979). Job redesign in the bindery. In R. G. Sell & P. Shipley (Eds.), *Satisfactions in work design* (pp. 155-160). London: Taylor & Francis.

Peacock, J. B. (1994a). *Users and abusers of ergonomics.* Presented at the Agricultural Machinery Conference, Cedar Rapids. Iowa.

Peacock, J. B. (1994b). *Ergonomics in very high volume manufacturing.* Paper presented at the All Ohio Safety and Health Conference, Cleveland, Ohio.

Peacock, J. B. (l994c). *Production with protection.* Presented at the National Association of Manufacturers Conference on Ergonomics and CTO, Nashville, Tennessee.

Peacock J. B., Glube, R., Miller, M., & Clune, P. (1983). Police officers response to eight and twelve hour shift schedules. *Ergonomics, 26,* 479-493.

**Chapter 17**

**Unwanted Energy – Vibration**

**(with Geoffrey Chase, Robert Fox and Richard Parker, Spring 2005)**

This article is mainly about mechanical vibrations and their effects on human performance and health. It also presents an opportunity to study Hooke's law (1676), chainsaws, fork trucks, and McPherson struts.

**Quality and Health Outcomes**

Once upon a time, a young boy, aged about four, was taken by his uncle on a tractor to plow the stubble in the previous year's cornfields. The ground was rough, and they bounced around a lot on the seat, which consisted of a shaped steel platform mounted on a V-shaped steel spring. The rough field also caused the front wheels to transmit the bumps to the steering wheels. When the little boy was allowed to drive, it soon became clear that the immaculately straight furrows were turning into a zigzagging mess; he was certainly not in line to win the plowing championship at the village gymkhana (a horse show and contests testing agricultural skills). The boy's performance was degraded by his inability to manage the mechanical vibrations inherent in the task and equipment.

Some years later, he returned to work on the farm and was given the much less precise task of towing a set of harrows behind the tractor to further break up the soil before planting the corn. The technology had not improved much - he still sat, day in and day out, on a sprung seat, albeit with an old cushion to provide a modicum of damping, and the steering wheel still harmonized with the dancing front wheels. The short-term costs of these efforts were very tired and sore hands and arms.

After finishing his undergraduate education in ergonomics, he had a choice between going to graduate school and working for the British aircraft industry, which at that time was very concerned about pilot performance and health as affected by vibrating fixed-wing airplanes and helicopters. He chose the former path, but not before he had the opportunity to learn about the substantial fatigue and performance decrements encountered by pilots, who sometimes had difficulty in reading their instruments correctly and in performing target-aiming tasks. To this day, pilots of small aircraft and helicopters are still hampered by a hostile mechanical vibration environment.

**Whole-Body Vibration**

Many years later, this now trained ergonomist was asked to analyze the ride quality experienced by public transport drivers and passengers in Hong Kong (Peacock, 1978). Back in the 1970s, the bus drivers still had primitive seats and spent long hours navigating the extremely busy streets.

At that time, researchers in England (Sandover, 1998) and the United States (Richards & Jacobsen, 1978) were developing methods to measure the mechanical aspects of ride quality and applying their findings to road vehicles, trains, and airplanes (including the Concorde). It became immediately clear that only part of the problem lay with the vehicle - a lot had to do with the road surface. Contemporary vehicle manufacturers have come a long way in addressing the

transmission pathway for induced vibrations by providing sophisticated suspensions that can absorb or even actively counter the vibration. Nowadays it is possible to drive high-end automobiles that are smoother than a flying carpet. But whole-body vibration is still a problem

on many types of heavy mobile equipment.

**Hand-Arm Vibration**

Workers who mend roads with jackhammers and concrete breakers are subject to excessive vibration that is passed through their hands and arms. This socalled segmental vibration can cause serious damage to the musculoskeletal, circulatory, and peripheral nerves of the upper limbs. Cold conditions compound the vibration effects and can result in vibration-induced white finger or Raynaud's syndrome, which in its advanced stages can result in the loss of one's finger (Brammer & Taylor, 1982; Wasserman, 1987). Similar percussion tools such as chipping hammers are used to clean flashings (unwanted pieces of molten metal that seep around the edges of moulds) from castings in the foundry, and these tedious tasks often result in

similar pathologies.

**Exposure**

Contemporary just-in-time manufacturing operations give rise to another vibration problem. Large over-the-road trucks with shock-absorbing seats deliver components to the docks of assembly plants, where floor space is usually very costly. Stand-up fork trucks are used to

move the boxes of components from the docks to temporary staging places (supermarkets),

from which sit-down fork trucks and "tuggers" carry the materials to the assembly line.

Fork trucks usually have solid tires and primitive suspensions because of their weight and the very heavy loads that they have to carry. Consequently, the vibration from uneven factory floors or warped dock plates is transmitted to the driver, who stands directly above the rear axel. Because of the intrinsic design of these stand-up devices, which need to have a small turning circle for their maneuvers, it is difficult to interpose an adequate damping mechanism between the operator and the vehicle and still offer good access to the vehicle controls.

Observation of such operations highlights the most important element of the vibration jigsaw: *exposure.* The traditional work day is eight or more hours, and for many reasons, job enlargement and rotation to reduce exposure duration are not feasible or acceptable, especially in organizations in which job choice is based on seniority and riding a fork truck is more attractive than working on the assembly line. Also, it is a common observation that self-paced drivers with quotas to fill or who have piece-based incentives will hurry to finish a task. This leads them to drive fork trucks too quickly over bumps, which increases the size of the shock, or to put greater force on the hand tool to finish the job more quickly, both of which lead to an increase in vibration stress.

**Measurement and Standards**

Measurement and analysis of mechanical vibrations has progressed enormously over the past few decades. Triple-axis accelerometers can be mounted at strategic sites in the transmission pathways from vehicles and tools to the operator in order to measure the amplitude and frequency spectrum of the vibrations. There are also dosimeters that integrate the mechanical

and temporal components of operations involving exposure to vibrations. But this leads to the biggest challenge of all: how to set and enforce standards that will protect workers from the harmful effects of vibration while attaining an acceptable level of productivity.

The International Standards Organization (ISO 13160) contains some 50 pertinent standards, including:

*• http://www.techstreet.com/cgi-bin/detail?producUd=22795*

• the US Navy identifies many pertinent references *(http://www.safetycenter.navy.mil/acquisition/vibration/resources. htm)*

• the American National Standards Institute *(*[*http://webstore*](http://webstore)*. ansi.org/ansidocstore!find. asp -* search on "vibration" for a list of more than 100 standards)

• the American Conference of Governmental Industrial Hygienists *(http://www.acgih.org/TLV/* describes the organization's latest standards postings with references)

• the National Institute of Occupational Safety and Health *(*[*http://www*](http://www)*. cdc.gov/niosh/handvibra.html* describes current efforts in occupational vibration research and development)

• The University of Tennessee has an Institute for the Study of Human Vibration, under the leadership of Donald and Jack Wasserman *(http://www. engr.utk. edu/ishvl).*

The Occupational Health and Safety Administration (OSHA) as yet has no specific vibration standards, although it does cite vibration as an important cause of work-related musculoskeletal disorders and addresses the issue through its general duty clause. ANSI and the ISO office offer guidelines that limit the exposure of workers to whole-body and upper limb vibration.

**Standards Enforcement**

Although these standards are based on sound physical measurement principles and epidemiological evidence, they face the challenge of most health and safety standards that are developed for application in a context in which speed and productivity, either imposed by the

company or self-imposed, are of importance. They also have to address the everpresent problem of individual differences in operational methods, adaptation, and sensitivity to mechanical stresses. Even though applicable and reliable protective standards do exist, there is an enormous

challenge of monitoring and enforcing them. For a start, measurement and analysis of exposure to vibration requires sophisticated equipment, is technically difficult, requires highly trained engineers and technicians for interpretation, and relies on sampling repeatedly over time and similar operations and operators to be reliable. Also, the adverse health effects of exposure to mechanical vibrations, as with noise and radiation, may take a long time to surface.

The ISO and most national standards are subject to regular and periodic review and update, although disagreements among experts persist in what may require change (Griffin, 1998; Mansfield, 2004). The user is cautioned to review applicable national and international standards carefully and note their differences. Some standards provide information more relevant to the design of products and comparison of alternative designs, whereas others (e.g., ISO 2631-1,1997; ACGIH TLVs® and BEIs®) are more appropriate for the assessment of vibration in existing

conditions.

**How Do We Prevent Vibrations?**

Instructors in mechanics tell us that vibrations might be initiated when we do a bun gee jump, sit on a tractor seat, or hold a hand tool (Hibbeler, 1983). Now it is quite important to check our weight before doing a bungee jump because Hooke's law tells us that the elastic strap will stretch farther for heavier than for lighter people until it reaches its limits of elasticity and we find out about Young's modulus, which reflects the ratio of stress and strain for different materials as they are stretched, or we hit a large mass in the form of the Earth and rediscover Newton.

Inertia is a funny thing; little things going fast will continue to do so until they come across very big things going the other way. Similarly, if we sit a very heavy person on a simple sprung tractor seat, he or she will bottom out uncomfortably. In both the bun gee and tractor cases, if, because of the mass attached, the static stretched length of the spring in question is too long, all ability to isolate vibration at one end of the spring will be lost. If we balance the stiffness of the spring with the mass of whatever is attached to it and start the coupled system moving, it will continue to oscillate.

Springs alone don't prevent vibration; rather, they offer a greater pull or push the farther they are extended or compressed. However, they can simply isolate the oscillation of the mass at one end of the spring while the other is fixed. The down side is that the tractor driver oscillates forever as punishment for his choice of an occupation with obsolete technology. The good news is that the contemporary heavy equipment manufacturers are now making driver accommodations comparable to those in a Cadillac.

If we add a shock absorber or damper in parallel with the spring to settle the motion by removing energy, then the faster we move, the greater will be the resistance to movement. The end result is no movement. But if we continue to abuse the shock absorber, the law of conservation of energy will result in its getting hot. The up side is that the tractor driver can rest a bit now, but the down side is that some vibrations and large forces will come through the damper whenever the tractor hits a bump. The greater the amount of input energy you wish to remove, the bigger the damper capacity must be and the greater the reduction in isolation.

So to prevent vibration being transmitted to the trunk of a seated operator or to the hand and arm of an operator using a hand tool, we must adjust the combinations of mass, elasticity, and damping of the various system components. Passive shock-absorbing systems are created by combining springs and damping materials or fixtures between the user and the offending system. However, because the damper must move to absorb the forces, it is necessary to have sufficient space between the tractor seat and the chassis or between the hand tool motor and the handle. In the car business, the introduction of McPherson struts addressed some of the challenges of compactness of suspension systems.

There are also active vibration and shock-absorbing systems that use sensors and computers to recognize an incoming force and add opposing forces from some other energy source. Such advanced control technology can be used on hand tools, vehicles, vehicle seats, and even skyscrapers. They are most commonly seen in high-performance sports cars that most of us can't afford, and they may take away all the feel and excitement of cornering. Active systems also add significant complexity for not always a great improvement in performance. So, the decision to go active should be made with caution. In other areas, we use active and semiactive (mostly variable orifice dampers and lead-rubber isolation bearings) to control skyscrapers and bridges under earthquake and wind load and, one hopes, given current events, blast load. In these cases, it is almost exactly like the vocational vibration absorption problem. Specifically, we need to manage the energy absorbed by the building. The only main difference is that in these cases, damage occurs over far fewer cycles, even for wind loading. An additional similarity to vibrating hand tools and tractor seats is that there is no exact solution but a series of design trade-offs between improved isolation and energy dissipated.

In all these cases, the goal is to manage energy by moving it from the "subject," where it causes damage, to systems and devices that can absorb it without damage. In each case, there is the trade-off of improved energy absorption with a loss of isolation from the destructive or vibratory input. For example, we can create the highest damping shock absorber in the world, but it would be almost impossible for any reasonable input to deflect it, so it would still have little effect.

**Conclusions**

The best solution to many of these vibration problems is automation; great strides have been made in industry, construction, and forestry to remove the human operator from the direct pathway of the mechanical vibrations that are inherent in the tasks of disassembly. A less

satisfactory solution is to absorb the vibration in the transmission pathway; so-called suspension eats in heavy equipment are becoming more common. Some contemporary hand-held chipping hammers, chainsaws, and concrete breakers are designed with damping material to reduce

the transmission of vibration to the user.

Ironically, some operators have been known to be resistant to such innovations because they interfere with the "feel" for the job and slow operations. An example of this phenomenon involves workers operating pneumatic chisels to cut stone or check weld strength. They insist on holding onto the vibrating chisel with one hand to guide the tool and finesse the job.

The next step in the mitigation hierarchy is the use of personal protective equipment in the form of shock-absorbing gloves that both dissipate the forces and channel them away from sensitive structures such as the median nerve as it exits the carpal tunnel.

The solution of weeding out the weak to lead to a robust survivor population will always be attractive to some employers, and many employees may be happy to leave jobs for which they are not suited, provided there is not too great an incentive (usually economic and social or political and philosophical) to stay too long in the offending job. However, this is a risky

philosophy for all concerned. As mentioned earlier, engineering innovation is the most attractive solution, and appropriate technology already exists in many cases. Trees can be stripped automatically, although the operator must sit all day in a very large, expensive, not very nimble machine. Some castings can be cleaned automatically, and some casting processes can reduce the amount of flashing, thus obviating the need for the use of human operators to perform this unpleasant job. Drivers' seats in trains, buses, planes, and trucks can benefit from vehicle and seat suspensions that make driving a dream, but at a significant cost. Chain saws and hand-held construction and destruction equipment can be fitted with shock absorbing devices.

Where these engineering interventions are deemed to be infeasible, administrative controls, such as job rotation and enlargement, must be used to reduce temporal exposure, and operators who are required to interact with vibrating equipment should be monitored regularly for sensitivity to vibration-induced illness. Where fine motor skills are required, the performance and quality payoff for reducing mechanical vibration will more than outweigh the costs of engineering changes.

**References**

American Conference of Governmental Industrial Hygienists. (ACGIH). (2003). TLVsil!>and BElsil!>. Cincinnati, OH: Author.

Brammer, A. J., & Taylor, W. (1982). *Vibration effects on the hand and arm in industry.* New York: Wiley.

Griffin, M. J. (1998). Review article: Predicting the hazards of whole-body vibration Considerations of a standard. *Industrial Health,* 36, 83-91.

Hibbeler, R. C. (1983). Vibrations. In *Engineeringmechanics: Dynamics* (chap. 22, pp 453-489). New York: Macmillan.

Mansfield, N. 1. (2004). *Human response to vibration.* London: CRC Press.

Peacock, B. (1978). Ergonomics aspects of Hong Kong public transport. *Hong Kong Engineer,* 6(10), 27-34.

Peacock, B. (1998). *Whole body vibration.* Presentation to Michigan Chapter of the Vibration Institute.

Richards, L. G., Jacobson, I. D. (1978). Concorde: Ride quality and passenger reactions. *Space and Environmental Medicine,* 49, 905-913.

Sandover, J. (1998). The fatigue approach to vibration and health: Is it a practical and viable way of predicting the effects on people? *Journal of Sound and Vibration,* 215, 699-721.

Wasserman, D. (1987). *Human aspects of occupational vibration.* New York: Elsevier.

**Chapter 18**

**"WARNING: Do Not Use While Sleeping" - The Role of Facilitators**

**(with Lila Laux, Fall 2005)**

This article is about the role of facilitators: labels, warnings, instructions, procedures,

checklists, tutorials, and augmented reality that are not part of the actual operation of a system but are intended to facilitate appropriate use and to prevent misuse. Often these facilitators are temporary in that, after experience with the system, the user can perform satisfactorily and safely without their help.

A facilitator aims to supplement the user's knowledge or current situation awareness to ensure an accurate and timely transaction and prevent misuse. Unfortunately, the facilitator itself may sometimes be a precipitating cause of failure.

**Previous Work**

Bailey (1982) presented an extensive discussion of facilitators. Designers who heed Bailey's advice and "rules of thumb" will go a long way toward achieving acceptable facilitators, but it is clear that many designers do not follow this advice and instead rely on their own flawed “common sense.”

Driven largely by litigation resulting from personal injury and damage to property, a great deal of scholarly work examining how to effectively warn and instruct has been carried out since 1940. An extensive bibliography of research on warnings and safety instructions is contained in Miller and Lehto (2001). The American National Standards Institute and the International Standards Organization also offer guidance on instructions, warnings, and labeling, both general and in domain-specific contexts. Many other organizations - such as the Chemical Manufacturers Association, Society of Automotive Engineers, and National Fire Prevention Association - also provide guidance in the development of facilitators for specific types of products.

Despite this plethora of guidance and advice, facilitators continue to be a major source of system failure, inefficiencies, and frustration in many domains. Edworthy and Adams (1996) questioned the direction of warnings-related research almost 10 years ago, and although there are examples of good facilitators today, there isn't much evidence that all of the 1200 or so research reports and guidance materials annotated by Miller and Lehto have generally resulted in facilitators that are useful and effective. In this article, we discuss some of the trends and practices that may suggest why we still see warnings like the one in the title.

**Types of Facilitators**

***Labels.***Contemporary airplanes, cars, process operating consoles, and home electronic equipment are typically designed with numerous functions, features, and associated displays and controls. Different equipment brands that perform similar functions may sport different user

interfaces because the designers attempt to use the interface as a mechanism for product differentiation. Standardized items like vehicle foot pedals and the steering wheel do not need labels, but the knobs and dials associated with many other vehicle systems, such as climate control and entertainment, are not standardized in design and location (Laux & Mayer, 1993) and therefore need some type of label. These labels take up valuable real estate.

Competition for real estate between the labels and the controls and displays often results in a one-sided compromise by the labeling department. The initial compromise is to reduce the label to some cryptic word or phrase that captures the essential function of the parent system and sometimes the operational alternatives, such as "- VOLUME +." A further compromise may employ an abbreviation and a smaller font, such as "VOL," with the hope that a clockwise rotation of the associated knob has a strong enough population stereotype. Instead of determining what users actually comprehend, some enterprising facilitator designers may fight back by using obscure acronyms or icons.

***Warnings.***When the inappropriate use of a system may cause harm to the user, system, or environment, it may be necessary to warn unwary users to behave appropriately. Classical guidelines for this class of facilitators require that the warning should attract attention, identify the hazard, spell out the implications of improper use, suggest appropriate behavior, identify foreseeable inappropriate behavior, and provide guidance to the user in the event that things go wrong. There are several options for the signal word, such as "Warning," "Danger," "Caution," or "Be careful!" - which in theory convey different levels of hazard. But some companies are averse to the signal word *danger;* they do not wish to admit that use of their product can be dangerous.

Other conventions add colors such as yellow and red to the signal words to indicate the relative severity of outcomes associated with inappropriate use (Parsons, Seminara, & Wogalter, 1999).

A common but unstated purpose of warnings, by whatever name, is to remind the user about the hazards of inappropriate use, thus passing the responsibility for an accident to the user.

***Instructions.***Users of very complex devices, such as children's toys, are often faced with the dreaded phrase "Some assembly required." Experience tells us that some parents may inadvertently throwaway the instruction material and proceed to spend most of the day and night discovering the correct assembly or operational sequence. When the instructions are good, the user proceeds successfully through the assembly sequence and still has time to play with the toy before bedtime. Instructions generally contain both labels of the components and warnings of incorrect actions. Instructions also contain greater detail of assembly or operational sequences than labels and/or warnings.

Whereas an icon or warning label may consist of pictures that represent somewhat less than 1000 words, the instruction manual for a typical cellular telephone contains many thousands of words and a plethora of lists, pictures, and diagrams. This presents a problem while speeding on the freeway when all the driver wants to know is which button to press to answer the incoming call! But perhaps the Nobel Prize for instruction writing should go to those intrepid designers of the instructions for dealing with a flat tire. Imagine changing a flat in an unfamiliar rental car on the freeway in the rain at night. In reality, maybe the first instruction for flat tire replacement should read: "Use your cell phone to call the Auto Club."

In instructions, there is a tendency to use technical or engineering terms in an effort to be succinct and accurate, but this results in terminology that users simply can't understand. An example of this is the use of "assist starting" in an automobile owner's manual to name what users typically call "jump starting."

***Procedures.*** Procedures generally consist of sets of instructions that each deal with one of many interacting subsystems. For example, the procedure for starting a car requires that the operator begin by arranging his or her workplace, including seats, mirrors, and seat belts; next, the driver must learn how to interact with the ignition system, having first been cautioned to check the status of the fuel, transmission, and braking systems. Finally, he is referred to the troubleshooting section of the owner's manual that deals with dead batteries. The use and utility of procedures depends on the degree and currency of training at the knowledge, rule, or skill level (Rasmussen, 1987).

A challenge for procedure writers is to anticipate *all* possible sequences and failure modes. Their tool of choice is the flow diagram with logical branches. The key to good facilitator design, as with medicine, is first to do no harm. *Harm* in this context includes both inaccurate and untimely advice. However, this raises another common human factors engineering dilemma - the speed-accuracy trade-off. Commonly, instruction and procedure designers err on the side of accuracy by expanding detail and sacrifice speed. Inevitably, operators with little more than basic knowledge may jump around and amalgamate many steps into one.

The American Heart Association procedure for advanced cardiac life support is an example of a facilitator that is complete and accurate but may not always be very effective. Knowledgeable laypeople, emergency medical technicians, and emergency room physicians may use the procedure. The procedure must contend with complex differentiation among choking, stroke, heart attack, and poisoning symptoms; it is also time critical, and the outcome may be catastrophic. Whereas procedures must be correct, sometimes

this obsession with precision may result in temporary overload or impatience among users, with the result that the procedure becomes both inefficient and ineffective.

***Checklists.*** The key to a good checklist is that it should be succinct while ensuring that the user follow each step in sufficient detail. A classic failure of checklists is that an item may not stimulate sufficient analysis by the user before he or she moves on to the next step. If the box is checked, it is assumed that the necessary analysis has been done.

In case the you are skeptical about this discussion, consider the NW255 accident at Detroit Metro Airport in 1987, which killed 150 people: The National Transportation Safety Board determined that the probable cause of the accident was the flight crew's failure to use the taxi checklist to ensure that the flaps and slats were extended for takeoff. Contributing to the accident was the absence of electrical power to the airplane takeoff warning system, which thus did not warn the flight crew that the airplane was not configured properly for takeoff. (<http://www.ntsb.gov/ntsb/brief.asp?evid=20001213X31759&key=l>)

In this instance, the use of a checklist – a facilitator - was complicated by the crewmembers' considerable familiarity with the procedure, a failed warning system (another facilitator), and the unforgiving context.

***Tutorials.*** A first-time user of a digital camera, with all its features, behaves and performs very differently from an experienced photographer who transfers knowledge from his or her previous experiences with traditional cameras. The purpose of tutorials is to bring the inexperienced user up to a level of knowledge sufficient for effective, efficient, and safe operation of the product in question. It is common for the tutorial to consist of basic operation theory, some application examples, and sometimes situations in which inappropriate actions may lead to system failure or damage.

The use of tutorials and walkthroughs has boomed with the growth of software packages and even some complex Web pages. Poorly designed tutorials, like navigation

instructions provided by homeowners familiar with their environs, may address only correct actions, with the result that the unfortunate user, once lost, may be unable to recover.

***Augmented reality.*** Perhaps the most sophisticated of all facilitators, next to an experienced human guide, is the set of technologies called *augmented reality.* A familiar example is the yellow line on the television screen that identifies the target for the next down in football. This line gives sofa-bound spectators an advantage over the referees, who have only old fashioned chains (and a little judgment) to measure forward progress. But for the spectator, this yellow line greatly enhances situation awareness.

New uses of augmented reality include combinations of computers and cameras to view the outside of a device and then show schematic details of the contents, together with diagnostic and maintenance advice. Even more exciting is the use of augmented reality for medical diagnostics and intervention. These facilitators enable the doctor to view the normal underlying anatomy, together with examples of possible pathology and even simulations to answer "what if' questions.

**The Human Factors Approach**

The human factors engineering community approaches this challenge of the bridge between engineering or operational designs and the user in a number of ways, including

a. design of the permanent displays and controls that are normally essential for successfully carrying out a transaction or procedure;

b. completion of a task analysis that addresses the sequences of information and actions that comprise the transaction, including the human and system

contributions;

c. assessment of the conditions and context of the transaction(s);

d. assessment of the knowledge and skills of the intended user (note that users' capabilities change with practice and may fluctuate with the conditions under which a particular transaction takes place);

e. evaluation of expected use and possible misuse of the interface or procedure (this step should involve formal usability testing, in which users of varying degrees of knowledge, skill, and experience perform representative tasks under a variety of conditions and contexts with a candidate set of facilitators); and

f. development of facilitators to ensure appropriate use and prevent misuse.

Facilitators should be designed with due regard to the expected users, possible misuses, and likely conditions and should be subject to formal evaluation.

These analytical activities lead to the design of facilitators to fill user knowledge and skill gaps. Because conditions and users vary, it is necessary to include flexibility in facilitator design while remembering the important rule: First, do no harm. In the context of facilitator (e.g.,

procedure) design, harm includes both inappropriate action as a result of the design, with various possibilities of outcome from minor inconvenience to catastrophe, and unnecessary delays in the transaction caused by the verbatim following of the procedure. It should be noted that errors

in the use of facilitators commonly cost recovery time, which in turn may lead to catastrophe in a time-constrained condition, lack of productivity, or simply frustration.

***Research, measurement, and analysis.***Although there are many so-called rules of thumb for facilitator design, there is no robust theory of the pervasive role that they play in everyday life and technical areas. Such a theory must address the level of complexity, outcome/recovery importance, human and contextual variability (including learning), and the speed accuracy trade-off. Measurement of the use and utility of facilitators must consider the human capabilities to learn and forget. Analysis of facilitators requires tools that can be used by both researchers and designers to evaluate the effectiveness of alternative ways of aiding human performance. Facilitators can become crutches that interfere with knowledge or skill acquisition.

***Requirements and design.***The literature cited earlier presents some considerations, guidelines, and standards for the design of facilitators, but these are rarely supported by clear system performance requirements. Consequently, many facilitators are designed as an afterthought and are refined iteratively following untoward incidents or user comments.

Most facilitators are static in nature and do not address the changing needs of users as they become more familiar with system behavior. There is a need, therefore, to develop facilitator performance requirements and detailed system and facilitator design specifications to ensure safe and productive human performance.

***Implementation and evaluation.***All system design processes should require the integrated design of facilitators along with formal assessment of expected human performance and performance variability. The changing role of facilitators should be evaluated as the system is validated in a simulated or operational context.

***Expected use and foreseeable misuse.***Design should address expected use and foreseeable misuse. For example, a step ladder is used to obtain access to high places, and Murphy's law causes the item of interest to be just out of reach when the handyman gets to the next-to-top step, which is adorned with the caution: "Do not use this step as a step!" In legal jargon, this may be termed an *attractive nuisance.* The PROFORMIS model (Askren, 2005)

goes some way toward the anticipation of product misuse.

***Clear and obvious.***A facilitator is not needed when the assembly, operation, or hazard is "clear and obvious." Some naive lawyers have been known to add " ... to the average user." On many occasions, human factors professionals have pointed out the obvious fact that catering to the average user will by definition not address the needs of many. Other population accommodation strategies insist on eighth-grade or fifth-grade reading levels, which goes some way toward helping the less capable but is difficult when precise but esoteric language is needed for a particular product.

Some time ago, the senior author offered a suggestion that the simple mechanics associated with a seat belt were clear and obvious - when you recline the car seat back, the seat belt won't work. Since that time he has surveyed many intelligent people, including some engineers, who say that they just don't think about physics when they want to take a nap in the passenger seat. What is clear and obvious with hindsight may not be so before the event. What may be clear and obvious to the designer or experienced user may not be clear and obvious to the less experienced user.

***Usability testing.***Facilitators are ideal for human factors student projects, either for test and evaluation of existing facilitators or to try out some theory of attention, perception, or compliance with novel approaches. These projects sometimes address the issue of ecological validity - they test the facilitator with real people under realistic conditions of physical and time stress and distraction. Also it should be noted that a successful usability test does not necessarily predict all possible failure modes. It may be infeasible to address some contexts of use and conditions of users, such as inebriation.

**Conclusions**

Ideally, systems should be designed for effective, efficient, and safe use without the aid of facilitators. This ideal is largely true for simple systems and experienced users. Systems should also be "fail safe" so that misuse by the user is forgiven. But these ideals are usually unrealistic in practice, and so facilitators will always be needed.

As with necessary system interfaces, the utility of facilitators is probabilistic, and it is unlikely that we will ever accommodate all the users all the time. But we can try by formally integrating the development of facilitators into the system design and evaluation processes.

**References**

Askren, W. B. (2005). Predicting and evaluating misuses of products. *Ergonomics in Design.* 13(1), 15-19.

Bailey, R.W. (1982). Printed instructions. In R. W. Bailey (Ed.), *Human performance engineering: A guide for system designers* (chapter 19). Englewood Cliffs, NJ: Prentice Hall.

Edworthy, J., & Adams, A. (1996). *Warning design: A research perspective.* Bristol, PA: Taylor & Francis.

Laux, L. F., & Mayer, D. L. (1993). Informational aspects of vehicle design - A systematic approach to developing facilitators.

Tn B. Peacock & W.Karwowski (Eds.), *Automotive ergonomics* (pp. 401-430). London: Taylor & Francis

Miller, J. M., & Lehto, M. R.(2001). *Warnings and safety instructions annotated and indexed* (4th ed.). Ann Arbor, MI: Fuller Technical Publications.

Parsons, S. 0., Seminara, J. L., & Wogalter, M. S. (1999). A summary of warnings research. *Ergonomics in Design,* 7(1),21-31.

Rasmussen, J. (1987). Cognitive control and human error mechanisms. In J. Rasmussen, K. Duncan, & J. Leplat (Eds.), *New technology and human error* (pp. 53-62). Bury St. Edmunds, England: Edmundsbury Press.

**Chapter 19**

**Whose Agenda Is It Anyway? - Robert's Rules of Order**

**(Winter 2006)**

This article is about committee meetings, their synonyms and euphemisms, and their purposes and processes. All of us have experienced failures (errors) of meetings and committees that were caused, at least in part, by a failure to follow due process. *Robert's Rules of Order* goes back to 1876. The original author, General Henry M. Robert, was an ergonomist at heart; he took on a complex human function (committee meetings) and attempted, with considerable success,

to develop a sophisticated set of procedural guidelines to enhance the effectiveness and efficiency of these activities. He now has an avid following in the form of the "parliamentarian society," as evidenced, for example, by the National Association of Parliamentarians. Online resources and organizations regarding Robert's Rules of Order may be found at http://www.

robertsrules.com/, http://www.parlipro.org/, http://www.robertsrules.org/, and http://www. parliamentarians.org.

**Scientific Management**

Sometime after Robert had published his recognition that the fabric of organizations was contained in the committees, which were parallel to the line management organizations, Frederick Winslow Taylor published his classic text *The Principles of Scientific Management* (Taylor, 1911). Taylor's views were that the line organization ruled and that managers were intrinsically more capable than workers of designing jobs through scientific

analysis and meticulous training. One of Taylor's motives was the elimination of "soldiering": working to rule or deliberately working slowly, but such was his conviction regarding objective

science and harmony that he may well have been talking about world peace.

The human factors/ergonomics *(HF/E)* profession is concerned with scientific work analysis and design. Inthat role, we act as consultants to management, often through the mechanism of committees and meetings, armed with our statistics, analyses, models, and reports. We may also aspire to Taylor's view that "What constitutes a fair day's work will be a question for scientific investigation, instead of a subject to be bargained or haggled over." However, we should be sufficiently humbled to recognize that even after our scientific analysis, much uncertainty and many intangible factors remain to be resolved by debate and negotiation among all concerned, and this is where committees and meetings come in to play. We call this *participation.*

**Committee Meetings**

Committees are not committees unless they meet face to face, remotely, or virtually and communicate. This is where the *agendas* (processes) and *Agendas* (purposes) begin to present opportunities for success or failure. For example, most meetings between management

and unions aim to come to an agreement (decision) on working conditions and rewards,

but the tactics (analyses, weighting of evidence, off-line conversations) would make Henry Robert turn in his grave. The ultimate meeting in which agendas differ is war, and tomes have been written about the resolution of these extreme agenda differences. There are degree courses in conflict management. Even the surrogates for war - competitive team sports, the law courts, and competitive business - are governed by rules of behavior, until the protagonists stretch or bypass the rules or find loopholes. These extreme analogies are pertinent because they highlight many of the reasons for and definitions of success and failure of meetings.

Webster defines a committee as "a body of persons delegated to consider,

investigate, take action on, or report on some matter" or "a self-constituted organization for the promotion of a common object." These definitions suggest that a committee may be delegated or self-appointed, but all agree that there must be some common objective. Unfortunately,

different, incompatible, or unspoken agendas often cast doubts on this common objective. They also suggest processes such as *analyze* (investigate), *discuss* (consider), *decide and implement* (take action), and *communicate* (report.) Common synonyms for *committee* include *group, team, board, commission, working group,* or *agency.* Some might argue that *group* is a misfit because a

group, unlike a team or board, does not necessarily have some predefined purpose, although group members are defined by some common characteristic (cohort) or activity (the passengers on an airplane). Furthermore, this group may morph into a team when the flight is delayed. However, as in some committees and teams, the agendas of the passengers may be personal

rather than for the common good.

In practice, the organization is often complicated by degrees of responsibility and autonomy of line management and committees. The parent committee, "the committee as a whole," often delegates responsibility to subcommittees that are more knowledgeable about the subject at hand. These subcommittees, sometimes consisting of only one or two people, often owe their allegiances to the line organization, and this may cloud their motives in analysis and recommendations. On occasion, these subcommittees may have been delegated considerable autonomy, and the parent committee merely "rubber stamps" their recommendations. Often this hands-off style of management succeeds, until something goes wrong and the auditors and investigative journalists have their day.

**What Did Parkinson, Cleese, Adams, and Alfes Say?**

The topic of meetings has been fair game for gurus and satirists, the latter of whom are sometimes more perceptive. C. Northcote Parkinson (1957) discussed the commonly observed fact that activity increases to fill the time available. He also pointed out that discussion time is inversely proportional to the importance of the topic, whereas Robert liked the five minute rule - implying that no discussion, no matter how important, should ever take more than five minutes. Perhaps Parkinson's most perceptive analysis is of the discussions that take place in the British Parliament, which, although informative, may bear no relation to the outcome of the discussion. Parkinson also pointed out that the most important purpose of meetings was self-aggrandizement.

Meanwhile, comedian John Cleese, in his training video "Meetings, Bloody Meetings" (1990), proposed a similar satirical view of chaotic meetings that mainly reflect different Agendas and agendas. Cleese showed a dysfunctional manager who fails to plan, inform, manage, and record. The cartoonist Scott Adams, in his strip *Dilbert,* often strikes home in his

portrayal of irrational and incompetent managers and their behavior in meetings.

John Alfes, my first boss when I moved into industry, retired and proposed a hiring service for professional meeting goers who would be punctual, not say much, take extensive notes, not go to sleep, always agree with the boss, and bring donuts. On another occasion, two managers argued about the "hardness" of a car seat, despite my (human factors) offer of objective and psychophysical studies. The decision had to be made that day. To be fair, one cannot subject every decision to scientific analysis. Committees will continue to exercise their best judgment in a timely manner.

**Technology**

Contemporary committees and meetings make use of telecommunications and multimedia technology such as the telephone, video conferencing, and computer-aided communication through the Internet. A driving reason for this trend is that more communication can take place with less and less costly physical movement of the communicators. Hundreds of e-mails a day present a challenge of prioritization; the generators of all these messages believe that they are contributing to the success of the meeting, the organization, or their personal Agenda.

In 1973, I was hired to investigate the *uses* of telecommunications in the British Health Service for the years 1980-2000. We explored fax, telephone conferencing, videoconferencing, data transfer, remote diagnosis, and patient management through telemedicine with the purpose of eliminating "the meeting." The explosion

of communication technology since that time has since dwarfed my ambitious predictions, but "the meeting" is still alive and kicking.

**Organizations and Robert**

There are two extreme views of the purpose of organizations. On the one hand, it can be argued that the purpose of all organizations is the employment of people, with the constraint that the organization must survive economically or *so*cially. The more prevalent view is that the purpose of (commercial) organizations is to make a profit for the shareholders, with the constraint that it must address some of the needs of the employees. An even more cynical view of the purpose of organizations is that they exist solely for the promotion of individuals. Recognition of these divergent views can go some way toward understanding the behavior of participants in the meetings that form the mechanism by which the organization operates.

Robert recognized that committees and meetings exist in the greater context of organizations. There is a wealth of literature on organizational behavior, much of which has a background in industrial and social psychology (e.g., Blum & Naylor, 1968). Analysts describe structural and functional organizations and dissect in great detail individual and group motivations, culture, decision-making, power, and politics.

The most familiar organization is hierarchical, with regions, divisions, and departments; in universities, there are colleges, schools, and departments. The military and other government organizations develop colorful names to describe their rigid hierarchies. Hospitals are notorious in their management of specialties and services. Programs, projects, and centers often complicate the structural hierarchy by borrowing people through a matrix arrangement.

These hierarchical organizations are characterized by the structural mechanisms through which people are paid; most people have only one boss, although matrix organizations sometimes muddy the waters. Some contemporary businesses have opted for a shallow, broader structure, with flexible project assignments and leadership.

Most large organizations also have hierarchical interdisciplinary committee structures that run in parallel with the line management structure. These committee hierarchies are composed of members of line management, and in theory, their decisions should influence how the line organizations should behave. However, in practice, there may be power struggles between the committees and line organizations, which leads to conflicts of interest and discrepant behavior between the office and conference room.

**Robert's Rules**

Robert spent considerable time discussing the format and management of the building blocks of meetings: motions - their amendment and their precedence. He was emphatic about the role of the chairperson and who may speak; he also suggested strongly that speeches should

be pertinent and limited in duration. He insisted on decorum among the members and derided behavior that is self-seeking. Voting follows discussion and allows a choice among simple yeas and nays, hand raising, and ballots. Simple and two-thirds majorities are used depending on the class of motion. Although Robert was very detailed in his discussion, he was also pragmatic

and cautioned committees not to put process in the way of effective decision -making.

Robert's rules address at length the appointment of committees and the mechanisms of reporting of committee deliberations, but they do not address the political rift that sometimes exists between committee and line organization hierarchies.

Robert's rules provide the procedures for topic selection, discussion, and decision making so that the assembled committee members, along with the chairperson and recording secretary, work from a common ground. The purpose of the rules is to reduce the influence of prejudice and seniority and to ensure that all points of view have a fair hearing before a democratic decision is made.

**Agendas and Consensus**

The rationale of a committee is based on the truism that two minds are better than one and that more minds are better still, although a committee can become too cumbersome with too many minds. Ideally, each of the minds should be independent and represent a different aspect

of the decision. In practice, the process of debate is such that the minds are rarely independent - the members are "open minded" and amenable to "changing their minds" given the evidence presented by their committee colleagues.

An extreme variation of committee pliability was described by Klein et al. (1993) in the context of naturalistic decision making. Sometimes discussion may result in so-called false consensus, in which the weightings of the evidence that is presented are influenced perhaps through seniority or through some particularly articulate presentation, and, given the uncertainty that exists among the body of decision makers, a conclusion is drawn that is proven to be wrong after the event.

Many years ago, I addressed the challenges of promoting police officers, grading college students, and judging artistic performances by a computer-based method that identified, separated, standardized, and weighted the decision components before presenting the amalgamated evidence, together with a graphic profile of the component scores, to a committee that had to make the final decisions (Peacock & Stewart, 1985). Since that time, I have investigated many mathematical procedures, such as Satay's hierarchy and Bayesian decision making, that are applied to similar problems, including the assessment of grant applications. When numerical values are included in committee votes, there are many subtle opportunities for individuals to have more or less opportunity to influence the outcome. On one occasion, I remember the head of one engineering department was emphatic that his students were intrinsically better than those in the other departments and insisted that they should have their test scores weighted accordingly when it came to the distribution of prizes. On another occasion, I noted that one judge on a panel that was assessing gymnastic performance presented much more variable scores than the other judges and so exerted greater influence on the final outcome.

**Conclusions**

The general purpose of committees and meetings is to reduce uncertainty. Robert's contribution was the establishment of due process, given that the committee members represented the best available evidence. Robert's rules ensure that this evidence is dealt with in a fair way prior to the vote. This principle of independence - one person, one vote following controlled debate – unfortunately only partly addresses the issue of agendas. By definition, most committee members have individual Agendas or subgroup affiliations. Sometimes subordinates have opinions that differ from those of their boss, who may be the committee chair or the chair of a committee higher up the chain. If the subordinate sticks to his guns, he may be rewarded by promotion or demotion, depending on the relative clout of his boss in the line organization. Whereas Robert may influence where an item or motion sits on an agenda and how each topic is dealt with, he will never be able to deal with different Agendas.

**References**

Blum, M. L., & Naylor, J. C. (1968). *Industrial psychology.* New York: Harper and Row.

Cleese, J. (1990). *Meetings, bloody meetings.* London: Xicom Video Arts Film Productions.

Klein, G. A., Orasanu, j., Calderwood, R., & Zsambok, C. E. (1993). *Decision making in action.* Norwood, NJ: Ablex.

McMurran, M. (1998). *If it weren't for people, management would be a science.* Norcross, GA: Engineering and Management Press.

Parkinson, C. N. (1957). *Parkinson's law.* Boston: Houghton Mifflin.

Peacock, B., & Stewart, D. B. (1985). Structured grading. In *Proceedings of the Human Factors Society 29th Annual Meeting* (pp. 142-146). Santa Monica, CA: Human Factors and Ergonomics Society.

Robert, H. M. (1915). *Robert's rules of order.* New York: William Morrow.

Taylor, F. W. (1911/1998). *The principles o fscientific management.* Atlanta: Institute of Industrial Engineers

**Chapter 20**

**Boomer, Sooner, and Donders**

**(with Max Fogleman, Summer 2006)**

This article is about reaction time, information theory, the Hick-Hyman law, and the many practical applications of these theories to human behavior and performance.

President Benjamin Harrison issued a proclamation that the Oklahoma Land Run would commence at noon on April 22, 1889, with the firing of a canon. The "boomers," who had been agitating for the opening of the territory for some time, planned to stake their claims in Guthrie or Kingfisher at that time. However, a group of "sooners" entered the territory before the gun and staked their claims - a negative reaction time, based on anticipation.

In the *1860s,* Dutch psychologist F. C. Donders articulated the concept of three forms of reaction time:

• A simple reaction-time task (a light bulb and a response button) involves perception and motor activity.

• A discrimination reaction-time task involves pressing the button if and only if one of a number of light bulbs lights up; that is, the reaction involves a discrimination stage .

• A choice reaction-time task requires the subject to press one of a number of buttons depending on which light bulb lights up - a discrimination plus a choice stage.

Athletic sprint competitions are won by fractions of a second, so a fast response to the starting gun is very important. In fact, an ideal start would coincide with the gun; this is possible given a constant, or close to constant, warning period (e.g., "On your marks, get set [warning period], GO!"). However, in their wisdom, the athletics regulators believe that the runner should react to the gun, and because a legitimate reaction time of zero is logically impossible, international regulations require that runners be recalled if movement is detected within 100 milliseconds of the gun.

In the world of driving and center high-mounted stop lights, one can clearly show that if drivers traveling at 70 miles per hour with a short headway relied on reaction time, they would all be dead. Because traffic lights are so familiar, we should use the green light as a warning, the yellow light as a signal to stop, and the red light to remind us that we have just been caught on camera and that the ticket is in the mail. Unfortunately, the dynamics of this very simple system, coupled with variability in reaction-time performance, produces a well-documented set of failures. Paradoxically, certain speed and light-change conditions present the driver with an impossible decision.

One simple game to demonstrate reaction time is with a marked ruler that is

dropped through the subject's semi-closed fist. The inch marker level with the lateral Surface of the index finger reflects the reaction time. The experimenter may also use a dollar bill, but in this case, he or she had better be aware of the vagaries of anticipation and predictable warning periods. A practical application using manipulations of beer mats may be used to indicate sobriety. Unfortunately, these measures may not be without their share of false positives and false negatives. The false negatives may have to wait until the traffic light test to measure their true performance level.

**Decision Difficulty**

Anticipation aside (if that is possible), a choice reaction time can be measured by pressing one button for a red light and another for a green light. Have the subject start each trial with a finger placed equally between the two response buttons or one finger of each hand over one or the other of the response buttons. This two-choice reaction time will be slower than an equivalent single-choice time. If one continues to add choices, one will observe that reaction time is affected by the number of choices or decision difficulty.

These observations may be used to demonstrate some very robust theories developed some 50 years ago by our human factors/ergonomics ancestors. Shannon and Weaver (1949) and Edwards (1964) developed the mathematical concepts of information theory, and Welford (1952), Hick (1952), and Hyman (1953) applied this theory to reaction time.

Shannon and Weaver (1949) treated information as a string of binary digits (bits). The more information in the outcome, the greater is the uncertainty in predicting that outcome. The basic quantity of information transmission is average information in the outcome, or H:



Therefore, there is more information in the toss of a fair coin (1 bit) than in the outcome of two events in which the probability of one is 90% (e.g., sun in the desert and the only other possibility is a 10% chance of rain) and there are only 0.47 bits of information.

Hick and Hyman applied this theory to reaction time. The Hick-Hyman law states that reaction time varies proportionately with the base 2 logarithm of the number of choices:



where Hs = log2N, or the bits of information in the stimulus, and N = the number of equally probable choices. More elaborate mathematics can be applied to situations in which the alternatives are not equiprobable.

The effect may be demonstrated by sorting a deck of playing cards. Starting with the cards face down, sort the cards into two piles (face up to standardize movement time and provide baseline data). Next, sort the cards into (in increasing order of information content) picture and number cards, black and red, odd and even, hearts, clubs, diamonds, and spades and then more complex combinations. The times taken for these increasingly difficult choices demonstrate the Hick-Hyman law.

The delay in reaction time that is related to the number of choices is attributable to the so-called thinking time in the reaction process. Thus, a reaction has at least three components: recognition time, thinking (calculation) time. and motor response time.

The Lafayette Instrument Company makes a range of devices that enable one to present subjects with variations of reaction-time tasks and that make data recording very easy. The basic unit has four colored lights, a buzzer, and five buttons. The experimenter moves through simple two-choice and multiple-choice protocols before adding complexities to demonstrate recognition time, movement time, and distraction. Information may be found at *http://www.lafayetteinstrument.com/reaction timing. htm.*

**Components of Reaction Time**

***Movement time.***One way to get ahead of the car next to you at a traffic light is to place your left foot on the brake and your right foot over the accelerator. As soon as the light changes, you simultaneously slide your left foot off the brake and press the accelerator with your right foot. Meanwhile, the young buck in the sports car next to you has his right foot on the brake and his left foot unoccupied. He has to move his right foot some 50 millimeters across and 50 millimeters down to hit the accelerator, and this wastes precious milliseconds. New York taxi drivers habitually use two feet to drive automatic cars. Even normal drivers like you and

I are often reminded by our passengers that we should apply the brakes when in fact we are only halfway through our movement time - the voice is faster than the foot.

***Recognition time.***In order to react to a signal or stimulus, one must recognize that there is a signal. Again, one must use anticipation to be looking in the direction of the signal (think of the traffic lights) or cocking one's ears like a dog or rabbit. The second challenge is to separate the signal from the surrounding noise. If the signal is masked by similar but otherwise irrelevant signals, this discrimination process will add to the overall response time. This can be demonstrated by the card sorting game with the lights turned down low or by surrounding the signal light or sound with similar flashing lights or sounds.

***Interference and distraction****.* The recognition and response phases may be greatly affected by interference and distraction. While you are waiting at the traffic light and dialing your cell phone, you may get an audible reminder from the car behind that you should be paying attention to your primary task. Similarly, if your response resources (e.g., feet) are not where they should be, you will be slow off the mark. If your eyes and feet are in gear but your mind is not, you may delay the traffic behind you. Even when drivers are not at traffic lights, they may adopt a common strategy of slowing down to increase the time available to monitor their primary task (the cell phone) and their secondary task (driving).

***Learning, parallel processing, and multiple resources.***Unfortunately for the simple theory described so far, the human response system has enormous ability to learn and anticipate. With practice and learning, simple tasks can become automated and relegated to a lower level of

neural management. Because people have multiple resources, they can do some tasks in parallel

(Wickens & Yeh, 1986). They will experience minimal interference when different tasks require different resources, but they will experience great difficulty when two tasks compete for the same, scarce capital. We can watch the road and listen to music with minimal interference, although simultaneously watching the road and visually monitoring a televised baseball

game is another matter.

***Physiological and psychological refractoriness.***When an impulse travels along a nerve, it does so by the movement of an *action potential* at a rate of between less than a millisecond and a hundred or more milliseconds, depending on the size and type of nerve. Once a signal has passed along a section of nerve, it takes a finite time before another signal can follow. This delay in nerve sensitivity is called the *physiological refractory period.*

A similar phenomenon can be demonstrated by observing reactions to sequential visual stimuli. If a large-intensity stimulus is processed, the response to a following stimulus will be delayed as a function of the intensity of the first stimulus (Koster & Peacock, 1969). This phenomenon may be observed in a startle response, such as following a loud bang, a bright flash of light, or a particularly attention-demanding or high-decisiondifficulty situation. The victim of such a large stimulus may "freeze" for a short period and be unable to respond effectively to other, perhaps relevant information. This may go some way toward describing the delayed responses in drivers who inadvertently press the accelerator (and continue to do so) in unwanted acceleration situations (Schmidt, 1993).

**Aging**

Because reaction time consists of sensory, cognitive, and motor components,

it is no surprise that older people have slower reactions. Older people with failing eyesight, perhaps in low lighting or glare, may take an undue amount of time to recognize a signal, which is attributable to some extent by their attention-switching processes. In addition, older people's

cognitive abilities, involving operational memory and decision processes, may also slow the reaction process. Furthermore, weak or untrained muscles and stiff joints may compromise the response process.

Collectively, these subsystem delays may compromise reaction time performance in a serious way. Fortunately, learning, experience, and anticipation in particular situations can be brought to bear to mitigate these slower response capabilities. Age and cunning (anticipation) can often beat youth and ability. [In this case, the (baby) boomer will be the "sooner." - Ed.]

**Complexity**

The Hick-Hyman law can be demonstrated in relatively simple situations. However, most day-to-day demands on our reaction time are much more complex. The warning times are not always constant or even clear. The stimuli may be degraded or otherwise buried in a cloud of other stimuli. Some of the pertinent information may have to be obtained from operational or long-term memory. The decisions may not be simple; there may be tradeoffs between alternative choices. Finally, there may be many "response buttons," and we will probably not always

have our index finger poised over the correct one.

Such situations are to be found in the highly over learned skill of car driving, but they become much more complex when we have to deal with a third dimension, such as trying to land an airplane, or the multiple subtasks that are necessary in operating a chemical process plant, controlling a set of unmanned aerial vehicles, making a medical diagnosis, or investing in the stock market. In these cases, there may be a speed-accuracy tradeoff; there certainly will be a reaction time, but simple laws of information theory are inadequate to predict human performance except in gross terms.

**Situation Awareness**

A composite construct has been developed in recent years called *situation awareness* (Endsley, 1993). This concept reflects human performance given myriad relevant and distracting stimuli that may affect a decision or control response. Typically, there may be multiple sources

of information, some of which may have an external source, some embedded in long-term memory through training and some in operational memory as a situation develops. Other information sources may be contained in the response options. "Good" situation awareness stems from "good" anticipation, which leads to "good" reaction times.

Controlled flight into terrain is a situation in which an otherwise "healthy" aircraft flies into the ground or a tall building. Such accidents happen when a pilot flying under visual flight rules encounters instrument meteorological conditions by flying into cloud. Contemporary warning devices (Enhanced Ground Proximity Warning Systems) using synthetic terrain databases and data from the Global Positioning System provide the pilot with about one minute of warning of an impending collision. The technology has advanced to a level that it also provides response selection aiding by calling out, "Pull up, pull up." Even more exciting aviation technology is found in synthetic vision systems, whereby a synthetic map of the terrain ahead can provide the pilot or unmanned aerial vehicle operator with much greater warning times and response guidance.

**Conclusions**

The message to human factors/ergonomics professionals is that we need to recognize the problems and vulnerabilities of reaction time and design systems that give people better and timelier warnings and sound response advice. However, athletes will still have to wait for the gun.

**References**

Edwards, E. (1964). *Information transmission.* London: Chapman and Hall.

Endsley, M. A. (1993). Situation awareness and workload: Flip sides of the same coin. In *Proceedings of the Seventh International Symposium on Aviation Psychology* (pp. 906 9\ ]). Columbus: Ohio State University Department of Aviation.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology,* 4, 11-26. Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental psychology,* 45, 423-432.

Koster, W., & Peacock, B. (1969). The influence of intensity of visual stimuli on the psychological refractory phase. In W. Koster (Ed.), *Attention and performance* (pp. 232 253). Amsterdam: North Holland.

Schmidt, R. (1993). Unintended acceleration: Human performance considerations. In B. Peacock & W. Karwowski (Eds.), *Automotive ergonomics* (pp. 431-452). London: Taylor & Francis.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication.* Urbana: University of Illinois Press.

Welford, A. T. (1952). The psychological refractory period and the timing of high speed performance - A review and a theory. *British Journal of Psychology,* 43, 2-19.

Wickens, C. *D.,* & Yeh, Y. Y. (1986). A multiple resource model of workload prediction and assessment. In *Proceedings of the IEEE Conference on Systems, Man and Cybernetics* (pp. 1044-1048). Los Alamitos, CA: IEEE.

**Chapter 21**

**"You've Got to Attend to Everything"**

**(with Gary Northam, Fall 2006)**

The title of this article is an actual quotation from my flight instructor (S. Woolsey, personal communication, July 2006). The article is about workload from the receiving end. It describes some of the challenges that have to be overcome while learning to fly. Hicks' law, De Jong's law, and Miller's law play a part. A key issue in workload assessment is a comparison among the task demands, stress reactions, and performance outputs as learning progresses.

**Introduction**

Picture this: The brand new student pilot is overwhelmed by a bunch of flight, engine, navigation, collision avoidance, communication, and operations displays, as well as charts, notes, and printouts, not to mention a forest of different-shaped controls. Meanwhile, his flight instructor is talking in a foreign language, and the friendly neighborhood air traffic controller is throwing out coded instructions at a million bits per second. Then there's the weather, with ATIS, TAF, METAR, and a whole gaggle of abbreviations with no rhyme or reason whatsoever. And don't get us going on those aeronautical charts.

Early in the first author's flying career (a few flight hours ago), he was complaining to a colleague about having to learn a handful of checklists that direct the pilot what to do in normal and emergency situations. The colleague, a Vietnam-era F4 fighter pilot, pointed out that if the author had a Mig on his tail, he wouldn't complain about having to master the procedures embodied in the checklists (M. Polay, personal communication, May 2006).

The student pilot brings a lot of baggage from riding a bike and driving a car (negative transfer, proactive interference). In an airplane, one has to learn to steer with one's feet and turn by using a combination of one's hands and feet. Sometimes one steers one way with the feet and the other with the hands to combat the wind and perform controlled slips and skids. The throttle seems to vary in its sensitivity over its range, and the choke (fuel-to-air mixture) is very sensitive - you're supposed to listen for when the engine is "running rough." The throttle is really for altitude control, and the yoke (pitch) is for airspeed. The air brakes (flaps) do funny things to your pitch, especially when you are going too slowly and have decided to go around and try again to get your approach right.

Falling off a bike usually results only in scraped knees, and although car accidents kill 40,000 people a year in the United States, most of these catastrophes are caused by bad behavior, such as speeding and drinking. Many more accidents result only in minor disruptions to metal and insurance premiums. If you fall out of the sky, however, you will have no further need for insurance, although the airplane will generally be treated as a write-off. An early comforting lesson in flying is that you can land safely even if your engine fails, provided you have acquired the appropriate knowledge, procedures, and psychomotor skills.

Following work by Shannon and Weaver (1949) on information theory, Hick (1952) and Hyman (1953) observed that the more information one has to process, the longer it takes to make a decision or response. Miller (1956) described a general rule - Miller's magic number 7 plus

or minus 2 "chunks"- that reflects well the ability of people to retain information over the short term. Some explanations used the concept of a limited-capacity, single information-processing channel. Other theories focused on the response selection process, suggesting that one can respond (attend?) to only one thing at a time. These simple approaches are inadequate to describe the complexities of real world information processing. The real problem is coming up with an adequate description of the ever-changing nature and size of a chunk.

More recent theories suggest multiple channels and multiple limited resources (Wickens, 2001). A comprehensive overview of stress, workload, and fatigue, edited by Hancock and Desmond (2001), addresses both physical and cognitive issues and delves into many domains in which workload is a concern, including driving, flying, and air traffic control.

**Measurement**

In physical ergonomics, the concept of workload is more tangible. There are clear biomechanical and physiological indicators, such as heart rate, oxygen consumption, and the NIOSH Lifting Equation. Industrial engineers' approach to the problem of workload measurement focuses on output and activity at both gross and micro levels. On the production line, predetermined work standards measure and prescribe activities down to the second. These measures usually assume an average experienced worker rate; more perceptive industrial engineers aspire to gauge how hard a person is trying by applying so-called rating methods. Snook (2005) and others developed methods, based on psychophysics, to determine the physical workloads that could be sustained over extended periods.

Borg (2005) developed a method whereby the individual rated his or her own rate of perceived exertion for physical work. Similar subjective reporting techniques for cognitive work are included in the NASA-TLX and other similar rating batteries (Hart & Staveland, 1988). These subjective physical and cognitive workload indicators are characterized by the same problem of individual differences in reference or anchor points. They use verbal anchors such as "quite easy" and "very strenuous" or "low" and "high." A marathon runner's perception of the physical workload while on a morning jog is very different from that of lesser-trained individuals. Similarly, the perceptions of a novice student pilot are very different

frol11those of an experienced one. However, given the calibration of an individual, these subjective methods are reliable in assessing relative differences, especially with large samples of responders.

Physiological measures of cognitive workload include heart rate, heart rate variability (sinus arrhythmia), and various electroencephalographic indices. Perhaps the most convenient, least intrusive, and most practically reliable measure is heart rate. A new generation of neurological measures for cognitive activity may have merit, but the equipment needed to watch people think is prohibitively expensive for widespread workplace application.

**Secondary Task Methods**

When the first author was an ergonomics student, he participated in a test to assess the mental workload associated with different levels of driving difficulty. A car equipped with an externally controlled throttle (replacing the driver controlled gas pedal) was driven over a slalom course with hard and soft tires at different speeds. The subject was asked to tap his now free right foot rhythmically while driving over the course (secondary task). The results indicated that the more difficult task (softer tires and faster speeds) was associated with greater arrhythmia in the foot tapping.

The theory behind this secondary task method was that the subject had a limited

capacity to handle driving and tapping, and as the driving task became more difficult,

performance in the secondary task would deteriorate. As with most empirical investigations of this kind, secondary task performance improved with practice, and after a while the driver could tap regularly, even in the more difficult primary task conditions. This brings us back to

the key matter of chunks and the ever present capability of human beings to learn: What constitutes a single chunk to the expert may represent a sequence of several chunks to the novice.

**Ultimate Automation**

The ultimate form of automation is found in the interactions between the

student pilot and his or her instructor. To the student, the activities of flying can be totally manual or totally "automatic" - the instructor does the flying. Handoff between manual mode and automation is clear: "My controls," "Your controls," "I have the controls." Intermediate levels

of automation are more subtle and vary between probing ("How is your altitude, pitch, and airspeed?"), more direct coaching ("add power, then pitch"), and direct help with the yoke and rudder. At times of high workload - when the student is trying to fly the airplane and a busy air

traffic controller calls out a long list of instructions, some of which are more important than others and must be repeated back - the instructor (intelligent automation) may simply take over the communication task and instruct the student when the airplane is stable and the workload is lower.

A major challenge in all time constrained activities, such as driving and flying, is that of distraction. The distraction afforded by the cell phone during driving is now legendary, although

there is no good solution in sight because of the generally low (but variable) workload

associated with driving and the high desirability of communication. Flying is very similar: There are periods of low workload, especially in highly automated airplanes in which the pilot is charged with the boring and vulnerable task of monitoring while engulfed by all the forces of vigilance decrement. In other flying situations, however, such as instrument flight, the workload can be nearly overwhelming. When a single pilot must fly the less automated airplane in cloud

cover and visibility just enough to land at a particular airport, there is much to do in a short time. Distraction can occur during periods of high workload demand and during periods of low monitoring mode demand, which may turn ugly very quickly.

Flight planning and navigation are key parts of flying. Flight planning involves setting up a series of waypoints and altitudes with due regard to weather and congested areas. It is surprising to find that all mountains and roads look alike when one is lost. Fortunately, modern airplanes have radar-, transponder-, and GPS-based task aids that can be used to confirm a location. If all else fails, the pilot can communicate by two-way radio with air traffic control or a flight service station.

The instructor, to make the point about task priorities, sometimes deliberately distracts the student pilot from his attention to navigation details. It is very humbling to be lost in space, especially in "instrument" conditions. During these moments, there is a great increase in stress and workload - the pilot has to fly the airplane, find his location, re-plan his heading, and, if necessary, communicate with his support services.

Perhaps the most stressful form of flying occurs in darkness and with clouds or fog while flying in a mountainous region. Coast Guard helicopter pilots in Alaska do this all the time (J. Smith, personal communication, July 2006). In these situations, the pilot must rely on instruments and charts to prevent a controlled flight- into-terrain incident. Again, stress and workload come together to affect pilot behavior and performance.

Contemporary synthetic vision systems are based on a terrain database and enable the pilot to see a synthetic picture of the terrain on the primary flight display. This picture is enhanced by distance and altitude cues. In addition, audible and color-coded warnings are displayed based on distance, time, and rate of closure with terrain. The pilot can reduce both workload and stress in these situations simply by obeying the command, "Pull up, pull up." Contemporary automation will pull up automatically if the pilot is too slow in his response and then politely hand control back to the pilot.

**Workload and Stress - Emergencies!**

Much of flight training centers on how to deal with emergencies. The first class of emergency avoidance is rule based. There is a detailed set of three-dimensional airspace regulations: Don't fly too close to the clouds or to congested areas, and give way to sky divers! Give way to the right, and always carry a two-way radio. In the preflight procedures, the student pilot explains what he hopes to do in the case of an engine fire or failure after rotation if no runway remains - a somewhat stressful situation with oodles of workload.

Flight training involves many maneuvers aimed at teaching the pilot how thrust, drag, lift, and weight interact. One learns to fly the airplane very slowly and still maintain control. But suddenly the whole thing starts to shudder, a warning horn blares, and then the nose drops and you are pointed downward. *You* have stalled, and if you're not careful, the stall will turn into a spin. Talk about workload! Meanwhile, the instructor is laughing her socks off, although her feet have moved to the rudder and hands to the yoke just in case you have forgotten your procedures - should you pull out the throttle or push it in? On occasion, the cruel instructor will tell you to close your eyes and then, because your neurovestibular senses don't tell you what is happening, you will be very surprised to find that you're pointing straight down and the air speed indicator is rocketing. Never mind, calm down and just try to remember the rule-based procedures.

Competing with these "unusual attitudes" for emergencies that bring on the biggest workload is when the engine fails while one is flying high. It's as easy as ABC: Aviate (fly the airplane), get the best glide speed, find the Best place to land, and then do your troubleshooting and emergency landing Checklists. Then just land the airplane. No pressure!

**Workload and Learning**

Workload is simply information over time. If information increases and time remains constant, you're in trouble. Similarly, if information remains constant and the time horizon diminishes, you may be in bigger trouble. Common human strategies for dealing with increases in workload are to use less of the available information or to buy time. The former strategy is flawed because it increases the chance or risk of erroneous decisions. The latter strategy is just fine if time can be bought, otherwise the response of failing to respond to the demands may cause a catastrophe.

Fortunately, the old adage "Practice makes perfect" applies to workload management. With practice, the pilot learns to distinguish what is more important in the air traffic controller's message. He also learns to store the less important material for future consumption. Also, the psychomotor demand of flying the airplane becomes more automatic with time. In the early stages of learning, even the apparently different resources of psychomotor control and handling verbal ATC instructions interfere with one another, as does directional control and cell phone use in an automobile. Often in these cases, the challenges of navigation go out the window. Did ATC say "clear for takeoff, runway 21L, no delay" and then "fly runway heading to the campus and then turn right at 6000' after inbound 747 on the downwind at 1 o'clock" all in one mouthful (earful)? Did you just forget the turn into your street while you were sending a text message to your home and then at the last minute trying to make the corner with disastrous results? Experienced pilots learn to deal with this workload with ease; the jury is still out on the driver who attempts text messaging while driving.

**Conclusions**

In this article, we have attempted to explain what workload is all about. One can measure the input, the physiological response, and the output (behavior and performance.) Some measures can be applied conveniently in real time, such as heart rate; others are measured after the event by questionnaires. The rest are guessed at by the accident investigator, who often hangs his hat on situation awareness (Endsley & Bolstad, 1994).

**References**

Borg, G. (2005). Scaling experiences during work: Perceived exertion and difficulty. In N. Stanton, A. Hedge, K. Brookhuis, & E. Salas (Eds.), *Handbook of human factors and ergonomics methods* (chap. ] I). Boca Raton, FL: CRC Press.

DeJong, J. R. (1969). The effects of increasing skill on cycle time and its consequences for time standards. *ErgorlOmics* 7, 51-60.

Endsley, M. A., & Bolstad, C. A. (1994). Individual differences in pilot situational awareness. *International Journal of Aviation Psychology, 4,* 241-264.

Hancock, P. A., & Desmond, P. A. (200]). *Stress, workload and fatigue.* Mahwah, N): Erlbaum.

Hart, S. G., & Staveland, L. E. (1988). Development ofNASA-TLX (task load index): Results of empirical and theoretical research. ]n P. A. Hancock & N. Meshkati (Eds.), *Human mental workload.*Amsterdam: North-Holland.

Hick, W. E. (1952). On the rate of gain of inform ation. *Quarterly Journal of Experimental Psychology, 4, 11-26.*

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology,* 45, 423-432.

Miller, G. A. (1956). The magical number seven, plus or minus two. *Psychological Review, 63,* 81-97.

Peacock,). B. (1972). *Information retention in manual control.* Unpublished Ph.D. thesis, University of Birmingham, UK.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication.* Urbana: University of Illinois Press.

Snook, S. (2005). Psychophysical tables: Lifting, lowering, pushing, pulling and carrying. In N. Stanton, A. Hedge, K. Brookhuis, & E. Salas (Eds.), *Handbook of human factors and erg0l1Omics methods* (chap. 13). Boca Raton, FL: CRC Press.

Walker, G. (2005). Verbal protocol analysis. In N. Stanton, A. Hedge, K. Brookhuis, & E. Salas (Eds.), *Handbook of human factors and ergonomicsmethods.* Boca Raton, FL: CRC Press.

Wickens, C. (2001). Workload and situation awareness. In P. A. Hancock & P. A. Desmond (Eds.), *Stress, workload and fatigue.* Mahwah, N): Erlbaum.

**Chapter 22**

**A Look Back: *Ergo* Is More Difficult Than *Nomos***

**(Spring 2007)**

This article is a personal reflection on applied ergonomics between the mid 1980s and the turn of the Century, mainly from the perspective of ergonomics practice in the automobile industry. It was a period of rapid expansion of electronics and computer technology, especially in cars, trucks offices and personal gadgets, and controversial debates regarding the relationships between production line work and occupational injury and illness. The title of this article - “*Ergo* is More Difficult than *Nomos*” refers to the challenge of bridging the gap between relatively narrow research and measurement (nomos) that often characterizes our profession and the multifaceted problem of designing products and work (ergo), which have multiple, often conflicting requirements – effectiveness, quality, efficiency, productivity, health and safety, and customer satisfaction.

**Car and Truck Design**

“Good morning, Brian,” said John Alfes, director of packaging in Advanced Vehicle Engineering, "Welcome to General Motors.” Why don’t you take off your jacket, you look like one of those university professors. By the way, I hope you’re a good Catholic, because you’ll need 12 promotions to become president of GM and only 4 to become Pope. And, unlike the Church, there will be an organizational change every three years.”

In 1986, General Motors had a substantial share of the world’s car and truck market, but European styling was becoming popular, electronics and computer technology were being introduced apace and the Toyota production system was setting the quality standard. My first job, as manager of the GM Advanced Vehicle Engineering Human Factors Department, was to meet with the Legal Department and find out what kind of human factors problems they had. To my surprise, it wasn’t controls and displays or even passive restraint systems; it was failure to warn. The challenge was on – we had to accommodate the vast majority of the user population on multiple dimensions, address the inevitable learning curves with new technology and the occasional behavioral lapses that produced unwanted outcomes.

At about this time, there was a great increase in the number of failure-to-warn lawsuits. Although there were many trivial suits, the debate raged over what was “clear and obvious to the average user.” One particular series of suits concerned the misfit between 15 ½ -inch wheels and 15-inch tires. This discrepancy was mostly found in older pickups and focused the argument on which warnings were reasonable and feasible on a wheel and tire and which could be placed in the owner’s manual. Other warnings debates and law suits revolved around passengers in pickup beds, hands in the engine compartment, and the effectiveness or otherwise of alternative seat belt designs. (A summary of the theory and practice of “facilitator” design is given in a recent EID article, “Do Not Use While Sleeping” (Peacock & Laux, 2006).

Perhaps the biggest controversy in the late 1980s was the so-called runaway car problem,in which a vehicle suddenly accelerated for no apparent reason, often with catastrophic results.The TV news show *60 Minutes* implied that this was an engine problem. All the automobile companies were faced with similar suits, and I was asked to investigate the behavioral aspects of problem. We conducted a major experiment of some 50 drivers who drove 12 different cars in multiple physical and cognitive loading situations. We demonstrated clearly that the unwanted acceleration problem was largely attributable to (usually) older drivers pressing the accelerator pedal instead of the brake. The brake-transmission interlock resolved a substantial proportion of the problems.

My next stop was at the design department, the organization tasked with anticipating what the future customer would like. At that time, rows of look-like buttons and digital displays were all the rage, so I pointed out politely that if we were to improve the speed and accuracy of interactions between the driver and the vehicle and reduce operational errors, we would have to provide some differentiation mechanisms and perhaps even have common locations for the controls. I learned rapidly that it was unwise to call the designer’s baby “ugly.” I did learn another important lesson -when asked about a problem, the human factors specialist should not say “come back in a year when I have run a designed experiment.” We were often expected to offer an informed opinion there and then at the conference table.

One amusing example of design around the table occurred during a seat design meeting for a popular large car. The chief engineer said that the [prototype] seats were too soft, but the program manager was of the opinion that they were too hard. I explained that I could run a psychophysical experiment and resolve the problem unequivocally, but the decision had to be made there and then. I learned that votes = opinion × salary.

Shortly after my arrival at GM, the company bought EDS and Hughes. EDS was supposed to revolutionize the way GM handled design, manufacturing, personnel, and financial information. Hughes was to teach us systems engineering, with human factors being a major player, and this exercise was successful in streamlining the vehicle development process – from concept through product design and on to manufacturing and production design, with a much shorter life cycle. Perhaps the most successful strategy was the joint venture with Toyota – to learn the Toyota Production System. Although we weren’t enamored of 5S (Japanese for "tidy up your workplace"), this and other simple but effective approaches to quality were a real success. A major component of this strategy was the introduction of team structure and job rotation..

A similar development in the 1980s was the Saturn experiment. This was a new, independent product line in a Greenfield location in Tennessee with a totally innovative and collaborative partnership between labor and management. One shortcoming of this partnership was the one-week shift rotation arrangement ( perhaps the worst possible schedule as employees are forever adapting their circadian rhythms), which was introduced in the name of fairness. Unfortunately unattractive product offerings and union politics eventually resulted in this organization losing some of its autonomy.

The electronics revolution was arriving in the late 1980s and the design department, along with marketing, rushed to add gizmos into cars. First came navigation systems, with a green screen and a whole bunch of buttons and menus. Each transaction (eye movement and dwell time) took a long time which, at 60 miles an hour, was not very wise. Then came head-up displays, and there was pressure to put every function imaginable in front of the driver. Fortunately, the Hughes human factors specialists put their simulator to good use and clearly articulated the problem of cognitive capture. Also the market did not rush to this technology, which had proved to be very successful in the aviation industry.

Next, we got involved in infrared vision for use in the fog and rain. Unfortunately, although it was very good at picking up cats and dogs, it performed less well with cold concrete posts. There were also multiple investigations of “near and rear obstacle detection systems.” One interesting product offered by a supplier was a small mirror next to the rear view mirror onto which movies could be projected. I watched *Top Gun* during a test drive and demonstrated sufficient spare mental capacity, at least on a quiet road. This idea wasn’t going anywhere, although just last year in Bangkok I rode with a driver who watched a World Cup soccer game on a dashboard-mounted TV!

A major safety project was centered on the design of passive restraint systems. The airbag had been introduced, but it was found to be most useful only in frontal collisions. Furthermore, it was observed that some drivers had stopped wearing seat belts under the impression that the airbag would solve all their problems. Consequently, we initiated a university design competition that attracted a couple of dozen proposals. We selected four of these for the first round of funding – Texas Tech, Tennessee Tech, Duke, and Purdue. They all produced very innovative ideas with very large graduate and undergraduate participation, but eventually the Purdue team was selected to design and test a deployable air belt. We conducted a series of crash tests, and the idea produced fairly good results. However, the performance was not adequate for production, and our safety department turned to the development of side air bags.

**Cars for Older People**

A big opportunity arose in 1988, when I became program manager for the ACCESS car – the idea being to design a car for the elderly car buyer, who had a lot of money but failing faculties. The concept embraced physical, information, and social ACCESS – such as entry egress, display design and vehicle ownership. This project also involved numerous university contributions, notably from Rice, USC, the University of Michigan, Virginia Tech, Wayne State University, Michigan State, Louisville, and Oakland. In all, some 220 features were developed by dozens of engineers and suppliers. Two notable innovations were the Emergency Communication System – an ONSTAR prototype, and the Neighborhood electric car. Eventually, after two years of development, this program was discontinued – “you can’t sell a young man an old man’s car.” Also the slogan, “buy this car, it may be your last” was not very sensitive. However, many of the concepts were integrated into new programs, and the four-feet-thick report was full of good human factors applications.

General Motors encouraged contract research, and we approached many outstanding academics to conduct experiments in automotive ergonomics – also the title of a1993 book that I edited with Waldemar Karwowski, which contained many chapters based on GM-sponsored human factors research.

**The General Motors Manufacturing Ergonomics Laboratory**

After the curtailment of the ACCESS car program I got the opportunity to develop the Corporate Manufacturing Ergonomics Laboratory, where I stayed for 10 years. This initiative followed a major thrust by the UAW and the Occupational Safety and Health Administration to address the challenge of work-related musculoskeletal disorders. I established this activity in the industrial engineering area of the organization, based on the premise that ergonomics should work hand in hand with the engineers responsible for workplace and job design.

The challenge was massive, given the size and diversity of the organization, which, in the US alone, had more than a hundred assembly plants (where the cycle time was between 45 and 80 seconds) and many more component plants where the cycle time was as low as 10 seconds. The clear strategy was training of plant ergonomists to do the “reactive work” and program ergonomists to carry out “proactive ergonomics,” - design for manufacturing and assembly. The enabling strategy was to hire laboratory staff from the best stables in the country – the University of Michigan, Ohio State University, Texas Tech, and Virginia Tech and match these young people with hardened plant guys, who already knew the “bad jobs” and how to work with the shop floor. We supported these efforts with a laboratory that was second to none, with all the usual ergonomics measurement equipment plus mockups of cars, component assembly and manual materials handling facilities, and a conveyor line. This became the center for ergonomics training and specific investigation programs.

First, we had to develop the training material - we decided to develop measurement and analysis techniques that were linked to design specifications. Given that we were faced with the usual human variability, we used a consensus approach to develop a sliding scale of stressor measures, together with the ubiquitous NIOSH lifting guidelines, energy expenditure (which was not a real problem), and repetition indices. Armed with these measurement tools, we set about to institute a common global manufacturing ergonomics program, with a major training component

**The National Ergonomics Scene**

The epidemic of “ergonomic injuries” was accelerating in the early 1990s, with OSHA developing an ergonomics standard. I was assigned part-time to the GM office in Washington, D.C. where my task was to offer unbiased ergonomics input to both the Department of Labor and the National Association of Manufacturers. This was indeed a challenge, as ergonomics was the battlefield in the battle for control of the job cycle (Peacock, 1993).

The groundswell of ergonomics was attractive for ergonomists of various levels of expertise and thousands of "snake oil salesmen" who jumped on the ergonomics bandwagon. I also participated heavily in the Institute of Industrial Engineers, HFES, Board of Certification in Professional Ergonomics, Society of Automotive Engineers, and other organizations, with a forest of acronyms ( e.g. NIOSH, NORA, ASTM, ANSI, ISO, NAM, the US Cof C,) that had an iron in the ergonomics fire. A particularly successful venture was a contribution to the design of what is now known as Ergoweb (through a grant from the American Automobile Manufacturers Association). These organizations spawned journals, textbooks, conferences and ergonomics products galore, some with more success than others. The joint venture by the American Association of Automobile Manufacturers and the National Association of Manufacturers in the first half of the 1990s, was superseded by the Applied Ergonomics Conference, which has gone from strength to strength.

The U.S. industrial ergonomics movement was not only focused on the automobile industry. Meat packing, construction, nursing homes, aviation baggage handling and the package mailing industries were all sites of increased incidence of work related musculoskeletal disorders. The proposed ergonomics standard was heavily criticized and some critics even accused us of purveying junk science. We ergonomists recognize that this criticism is based largely on the probabilistic problem of individual differences in capability and susceptibility to injury. (Peacock, 1994, 1995, 1997, 1998).

**Globalization**

In the late 1990s, it became clear that the worldwide competition for vehicle sales was such that streamlining and operational economies had to be made. One principle was that GM (like Ford) could not afford to duplicate design work at multiple centers. Consequently, a global strategy was developed that focused design in a few large centers for cars and trucks to be manufactured and assembled in many countries. A snag was that the manufacturing processes had to be compatible with the laws of all the manufacturing countries. I was dispatched to Germany half a dozen times a year to introduce our “global common process for manufacturing ergonomics.” I was greeted with, “We have been building cars for 100 years!!!” Diplomacy and German beer enabled a consensus on how to implement proactive and reactive common ergonomics processes.

**Conclusions**

It became evident that the major causes of work-related musculoskeletal disorders are not just posture and force – they usually have conceptually simple remedies – rather, the challenges are the temporal issues of repetition, duration, and individual differences in capability and susceptibility, which are less easily dealt with in practice. The 1980s and 1990s were an exciting period for ergonomics which spawned many ergonomics products, processes, guidelines and technical organizations. Unfortunately, although “Good Ergonomics is Good Economics” (Hendrick, 1996) sometimes the short term “bottom line” goals are too attractive.

**References**

Peacock, B., & Karwowski, W. (1993). “*Automotive ergonomics”.* Philadelphia: Taylor & Francis.

Peacock, B. (1994a). “Uses and abusers of ergonomics.” Presentation at the Agricultural

Machinery Cedar Rapids, Iowa Conference,

Peacock B. (1994b). “The Battle for Control of the Job Cycle**,”** Presentation to the American Automobile Manufacturers Association sponsored conference, Nashville TN.

Peacock, B. (1995). “The Philosophy of Ergonomics Standards,” Human Factors and Ergonomics Society Annual Conference, San Diego, October 1995

Hendrick H (1996) “Good Ergonomics is Good Economics” HFES Society lecture.

Peacock, B. (1997). “Ergo is more complicated than nomos” Plenary address, International

Occupational Ergonomics and Safety Conference, Washington DC

Peacock, B. (1998). “The Ergonomics Chessboard: Moving Labor and Management to Cooperate and Communicate” The National Ergonomics Conference and Exposition, Detroit

Peacock, B., & Laux, L. (2006). “Do not use while sleeping: Facilitator Design,” Ergonomics in Design.

**Chapter 23**

**Just a Moment?**

**(Summer 2007)**

This article is about lifting and its unwanted outcome. It addresses the anatomical and biomechanical bases of so-called bad backs, some contemporary analytical tools and the challenges for those designing workplaces, operations, and training to prevent these debilitating injuries.

**Mechanics**

It is very apparent that lifting in the real world rarely occurs in the saggital plane because the starting and ending points are found all around the lifter. Objects to be lifted aren’t all the same size, shape, and weight, nor do they sport convenient handles. And sometimes, as in handling airline baggage and elderly patients, all sorts of people lift all sorts of objects all day long.

A particular anatomical point of note is that the facet joints are oriented differently in different regions of the spinal column, and this orientation has a lot to do with the possible direction of motion in those regions. For example, the thoracic region allows more rotation, whereas the lumbar region is mainly limited to flexion and extension and side flexion. However, in a lifting task, the static and dynamic forces involved must be transmitted all the way up and down the biomechanical chain – from the feet to the fingers. The farther forward one bends, the greater the counter moment has to be in the back tissues, and their moment arm is considerably shorter than the load arm. Furthermore, these moments act in all sorts of directions during an awkward movement. To make matters worse, the dynamic forces (velocities, accelerations, and jerks) greatly increase the stresses. Something has to give, and that something will be the tissue that momentarily has the worst mechanical advantage. It should be noted that both weightlifters and materials handlers often successfully use the dynamic strength of their leg muscles to produce what is called a *jerk* to achieve maximum lifting performance.

An individual movement requires the involvement of many spinal joints. Consideration of the biomechanical chain indicates that although much of the actual spinal rotation occurs in the thoracic region (with help from some hip rotation), some torque (a twisting moment) occurs in the lumbar region, and this, by definition, must stress the limits of the lumbar facet joints and related structures. Also, there must be a transmission of linear forces up and down the spine, as a result of contraction of the spinal muscles, which creates increasingly large compressive forces in the intervertebral discs.

So what is a “bad back”? Contemporary technology, such as MRI, contributes significantly to precise diagnosis and isolation of the damage. Before this technology, X-rays and the clinical evidence of (referred) pain down the leg usually pointed to a popular diagnosis of a prolapsed intervertebral disc pressing on a nerve root. However, as was pointed out earlier, there are so many tissues and interconnections in the region, a precise diagnosis is often impossible.

In practice, this does not always matter. Soft tissue damage may repair in a few days with the help of muscle relaxants, heat, painkillers, and rest, followed by graduated physical therapy. On the other hand, damage may recur, especially if the patient returns to the manual materials handling tasks and behaviors that caused the problem in the first place. Our epidemiology colleagues tell us that the best predictor of an occurrence of a bad back is a history of the same ailment.

**Measurement and Models**

Floyd and Silver (1951) used electromyography to show that in the fully flexed posture, the back extensor muscles “turn off.” This results in stress in the elastic and inelastic ligaments that pass to the rear of the center of rotation – the intervertebral disc. Another assessment technique involved the measurement of intra-abdominal and intra discal pressure Results from studies using these methods, together with static analyses, contributed greatly to the development of lifting guidelines and the University of Michigan Three Dimensional Static Strength Prediction Program (Chaffin, Anderson, & Martin, 1999).

A concerted effort began in the 1980s to address the dynamics of back motion. It was argued that although the static forces were of considerable importance, the addition of time and motion to the analysis painted a much more precise (and bleak) picture. Again, electromyography became the first tool of choice, but later dynamic back goniometry became the preferred method (Marras, 2005) The Lumbar Motion Monitor is composed of transducers, mounted on a corset, that sense the motions of the lower back in all planes. The inclusion of load information – upper body and carried load weights – allows one to develop dynamic equations that can predict the stresses on various back structures and, as with the static models, to predict the probability of structural failure.

Fatigue and its effect on lifting behavior also received attention (Kumar, Fagarasanu, Narayan, & Prasad, 2006). As manual materials handlers become fatigued, the kinematics of their lifting technique changes visibly. Instead of bending and straightening their knees, they tend to keep their knees straight and bend forward at the waist. An explanation of this behavior is that the lifters optimize their physiological cost; less energy is required when the knees are straight because the vertical motion of the upper body weight is reduced (Fogleman and Smith, 1995). Other investigations in looked at the location of the load in relation to the direction in which the lifter was facing and the effect on whether or not the feet were moved. It was observed that the probability of foot movement interacted with subject height and the radial direction of the load origin and destination (Peacock, 1980).

**Variability**

This whole mess is compounded by variability – in the size and physique of the lifter; in the state of training (or selection) of the lifter; in the weight, location, and destination of the load; and in the posture and dynamics of motion of the lifter. It is no wonder that occasionally the weakest link will fail, when all these sources of variability stack up, (Mirka and Marras, 1993). Variability is the both the Achilles heel and Holy Grail of ergonomics – how can we include variability in our analyses, models, predictions, and prevention guidelines? Snook and Ciriello (1991) used a psychophysical approach to this problem of variability in the measurement of human behavior in manual materials handling tasks. The product of these studies predicts the proportions of the (adult) population that can safely and comfortably handle various loads in various spatial and temporal contexts.

Another ramification of variability in anthropometry, posture, kinematics, motor skills, and lifting kinetics is that subtle differences occur not only in gross, observable motions but also in the micro motions that must occur to stabilize the spine during lifting. The semi rigid spine may be subject to numerous small variations, some of which place the delicate structures in vulnerable positions. Such variations may be fatigue induced. The fine motor control of these fixator and synergistic muscles may be subject to considerable situational variability. In addition, the actual speed of lifting is mainly controlled by the lifter with only minimal influence of external pacing mechanisms. Unfortunately, the relatively crude instruments and models that are available are inadequate to measure such complex and delicate variations in the postures, forces, and movements that occur in the real world.

**The NIOSH Lifting Guidelines**

In 1981 and again in 1991, NIOSH produced a composite model of manual materials handling – the widely used NIOSH Lift Equation (Waters, Putz-Anderson, Garg, & Fine, 1993). This model includes load, vertical and horizontal workplace dimensions, interface, frequency, and shift duration components. It was based on a consensus of experts who interpreted the biomechanical, physiological, psychophysical, and epidemiological evidence. By definition, the predictions of the model had to be conservative to protect the majority of the materials handling population. This essential conservatism, together with the model's plausibility and ease of use, has led to both its wide application and its criticisms.

The 1981 model includes two limits to the composite index: a warning limit and an action limit. The 1991 equation produces a lifting index and has led to considerable discussion of exactly what should be an acceptable task design level. A particular criticism is that the model does not address the essential selection and training of people involved in manual materials handling tasks. Another characteristic of the 1991 version is that it is more sensitive to frequency than to load and spatial factors. However, despite these criticisms, the NIOSH lift equation is probably the most robust and widely used practical tool for analyzing manual materials handling. Furthermore, although some experts argue that the tool should be used only within limited lifting conditions, an experienced ergonomist will find the tool useful in a much wider variety of manual materials handling situations because it draws systematic attention to the primary design aspects of materials handling.

A complicating factor is the effect of *encumbrance* on the mechanics and physiology of manual materials handling. Weight lifters, firefighters and extravehicular activity (EVA) astronauts wear restrictive clothing, which opposes the motion of the trunk and the efforts of the trunk muscles. This greatly limits the ranges of motion of the spinal and limb joints. One strategy to reduce the incidence of back disorders is to attempt to restrict the range of motion of the lower spinal joints through the use of a more or less tightly applied back belt. At $20 per belt, this is a very attractive solution to one of industry’s major challenges. Unfortunately, the biomechanical effects of back belts, whether loosely or tightly applied, including their effect on intra-abdominal pressure, are not convincing (NIOSH, 1996).

**Moments**

The problem with moments is that they take place over short and long “moments” and there may be many moments in a day’s work. This focuses attention on the effects of frequency, duration, repetition, and variability of the tasks involved. Are bad backs simply caused by cumulative wear and tear of the tissues, or is the predominant etiology a single damaging moment that is exacerbated by repetition? This dilemma has both analytical and classification implications. The history of an occurrence of back pain often points to a defining moment and thus classifies back pain as an injury. On the other hand, those other so-called ergonomic disorders of the upper limb are more often categorized as overuse, repetitive, or cumulative illnesses. To the patient, this argument is moot – bad backs and carpal tunnel syndrome hurt. But the government, employers, insurance companies, and policymakers use the distinction to argue vehemently for and against an ergonomics standard, with the ergonomics profession being booted about like a political football. There is no doubt within the ergonomics profession that the problem of bad backs is a complex combination of moments, both spatial and temporal.

In conclusion, although bending one’s knees and keeping one’s back straight certainly may have the effect of reducing and resisting the moments associated with lifting, these techniques may not always be practical or practiced. Because of the difficulties faced by analysts in precisely predicting where, when, how, and to whom a structural failure will occur, practicing ergonomists are limited to gross levels of intervention, as suggested by the NIOSH lifting guidelines, which, is probably still the most robust and practical analysis and design tool.

So although lifting is just a lot of moments, which are well defined and analyzed in benign laboratory settings, and despite voluminous research (Marras, 2005), the jury is still out on how to predict precisely when and to whom these moments will be momentous. And bad backs will continue to be a painful and elusive challenge.

**References**

Chaffin, D. B., Anderson, G. B. H., & Martin, B. J. (1999). *Occupational biomechanics* (3rd ed.). New York: Wiley.

Floyd, W. F., & Silver, P. H. S (1951, January 20). The function of erectores spinae in flexion of the trunk.” The Lancet, pp. 133–138

Fogleman, M., & Smith, J. (1995). The uses of biomechanical measures in the investigation of changes in lifting strategies over extended periods. *International Journal of Industrial Ergonomics, 16,* 51–71.

Kumar, S., Fagarasanu, M., Narayan, Y., & Prasad, N. (2006). Measures of localized muscle fatigue. *Ergonomics, 49,* 1092–1110.

Marras, W. S. (2005). The future of research in understanding and controlling work-related low back disorders. *Ergonomics, 48,* 464–477.

National Institute of Occupational Safety and Health (NIOSH). (1996). *Backbelts: Do they prevent injury?* (DHHS/NIOSH Publication No. 94-127). Cincinnati, OH: Department of Health and Human Services.

Mirka, G. A., & Marras, W. S. (1993). A stochastic model of trunk muscle coactivation during trunk bending. *Spine*, *18*, 1396–1409.

Peacock, B. (1980). The physical workload involved in parcel handling. *Ergonomics, 23,*  417–424.

Snook, S. H., & Ciriello, V. M. (1991). The design of manual handling tasks: Revised tables of acceptable weights and forces. *Ergonomics, 34,* 1197–1213.

Waters, T., Putz-Anderson, V., Garg, A., & Fine, L. (1993). Revised NIOSH lifting equation for the design and evaluation of manual lifting tasks. *Ergonomics, 36,* 749–776.

**Chapter 24**

**Ethics and Ergonomics: Customer Satisfaction**

**(with Gary Northam and Erica Diels, Summer 2008)**

This column is the first in a series about ethics and morality in the practice of ergonomics. In this installment, we introduce the ideas and present some of the history of ethics, comparisons with other professions, and measurement challenges. For the purpose of this article, we define *morality* as a set of (human) laws that aspire to ensure harmony among individuals and groups. Ethics, on the other hand, embraces the study of morality and the practical standards that are set to define morality more precisely. In the series, we will address the various articles in the

Human Factors and Ergonomics Society’s Code of Ethics.

**Ethics and Human Groups**

The history of ethics and morality is very long and deep, stemming from the ancient philosophers (Plato, Aristotle, and, more recently, Kant). All religions involve belief in a higher purpose and expectations with regard to behaviors among people in general and those of adherents to a particular religion. Despite differences of so-called commonly accepted practices among religions, ethical concepts generally transcend national and religious boundaries. Maslow’s (1943) highest-level human aspiration – “self actualization” – is a lofty goal; even Mother Theresa expressed doubts about her qualifications for this level. Contemporary ethical practice is more aligned with Maslow’s “esteem” level; hence the title of this article: customer satisfaction.

National and international laws and treaties provide some boundaries to the more general collaborative and competitive relationships among religions, nations, and businesses. However, particularly controversial areas of international disagreement are related to intellectual property

protection and safety standards, which are central to ergonomics research and practice.

The professions of medicine, engineering, law, journalism, and business are explicit regarding their expected standards of practice. These are generally developed pragmatically by attention to the voice of the (external and internal) customer. Unfortunately for those who aspire to establish ethical standards, most organizations and individuals usually have many different customers, whose requirements may not only differ but also may be in conflict. Other general concepts that affect the establishment of ethical standards include relativism and materiality.

**The Human Factors and Ergonomics Society**

One characteristic of a profession is that its members’ behaviors are regulated by ethical standards (V. J. Rice & J. R. Duncan, personal communication, June 22, 2006)**.** The HFES Code of Ethics includes five articles: Professional Qualifications, General Conduct, Publications,

Subject Precautions, and Forensic Practice (HFES, 2005). Each article is further explained by a series of principles. The Professional Qualifications article states that human factors professionals should confine their practice within their area of competence. Furthermore, they have an obligation to “represent their employer’s interests accurately.”

The General Conduct article is similar to that of most professions and suggests a level of decorum and honesty and an absence of bias in professional behavior. Because the field of human factors/ergonomics has a substantial research component, including applied research, the third and fourth articles address the need for subject protection during the conduct of research and intellectual integrity in reporting. The final article deals with forensic practice – a common activity of both professors and consultants. The principles associated with this activity include unbiased testimony and confidentiality.

**Measurement**

Lawrence Kohlberg is credited with being the originator of a measurement process that can be applied to moral development (Kohlberg, 1984). He identified six stages of moral development, which are grouped into three levels, as shown in the table on page 5.

This approach, which is still widely used, presents the subject with a series of scenarios, either in an interview or through a survey. After the subjects have read and, it is hoped, comprehended each scenario, they are presented with a series of related dilemmas and asked how answers to related questions are important to the ethical construct under discussion.

These concepts were formalized by the Defining Issues Test (Rest, Narvaez, Bebeau, & Thoma, 1999), which presents recognizable scenarios followed by a series of questions regarding the subject’s interpretation of the importance of each question as a measure of ethical practice. These scenarios are somewhat generic, and room is left in the instrument for domain specific material using a similar scoring approach. The techniques have been well validated and are used widely for age- and profession-related assessments.



**Moral Development and Ethics Education**

There is a wide body of research related to moral development, and many professional education programs have tailored ethics courses that often include specific case examples. The implication of this research and teaching is that ethics can be taught; it is not necessarily innate or intuitive. However, it is generally agreed that the family, church, and grade school experiences promote a maturation of moral beliefs and ethical behavior.

Ethics education in medical, engineering, business, and law schools addresses domain-specific dilemmas. In medicine, the topics range from the fundamental issues of life and death and right to life to the use of placebos for research purposes. The Hippocratic oath emphasizes that the doctor shall “first do no harm,” although it is well appreciated that some harm may be necessary to achieve a greater good, as with radiation treatment for cancer.

In business, financial management is the focus, particularly as it affects customers, employees, and the taxation authorities. In criminal law, the prime emphasis is on the rights of a prisoner to competent representation, a fair trial, and consideration of evidence on the presumption of innocence. In engineering, the topic is contained in the “Obligation of an Engineer,” which addresses concepts such as *mankind’s benefit*, *practice integrity*, and *utmost effort and fidelity to the profession*. To date, there have been no substantial formal efforts to include ethics education in human factors/ergonomics.

**The Voice of the Customer**

When we look at the life cycle of a product or process, we find that the eventual user or customer is concerned with functionality, form, safety, and cost. An intermediate customer, such as the manufacturing employee, adds cost to the product or process and may be much less concerned with the requirements of external customers.

Other necessary “customers” are company managers and shareholders, whose primary purpose is profit, with the constraints of production costs and the costs of product or process failure in the hands of the eventual user or third party. The ergonomists who put wheels on airline baggage so that passengers could carry more weight were not thinking of the baggage handlers’ backs or the collateral damage to the passenger sitting below the overhead bin. This particular example – the weight of airline luggage – clearly identifies the conflicting requirements of different customers.

The purpose of ergonomics work is usually defined by managers, and if the ergonomist doesn’t agree with the purpose of his or her organization, he or she can usually find employment elsewhere. If these ergonomists stay, they are obliged to rationalize the differences between their personal ethical views and those of their organization. If the purpose is to satisfy the employee by providing expensive safety facilities, then the eventual customer may have to pay a higher price for the product or process. If the cost requirements of the eventual paying customer are emphasized, then there is no doubt that the ergonomist can design a quickly moving production line that may increase the incidence of work related injury.

Of course, there is always the option of moving the manufacturing process to a country where labor costs and safety standards are lower. How should the practicing ergonomist address ethical standards within multinational corporations?

Usability professionals may aspire to universal access but may eventually have to draw a line that excludes the old, the young, the infirm, or the uneducated from product use, because including them may interfere with the effectiveness of the product for the majority of users.

Tradition in HF/E has suggested that we accommodate the 5th percentile, or 95% of the expected user population. This practice is conceptually straightforward when one considers a single anthropometric dimension, but when multiple anthropometric, biomechanical, sensory,

cognitive, and social dimensions are thrown into the mix, these guidelines are simplistic at best and impossible to define and implement at worst.

This dilemma of ergonomics purpose and different customer requirements is particularly pertinent in military applications. Whereas the overt purpose of most military organizations is the defense of their nation’s historical territory and resources, this purpose usually involves conflict with other nations and organizations. Much HF/E research is supported by the military, with customers who vary from field soldier, airplane pilot, and UAV operator to generals and politicians who manage the wars or “military actions.”

It might be argued that government leaders, by definition, have a self-interest in preserving or improving the condition of their citizens. Governments generally aspire to adhere to social norms and internationally agreed-upon treaties and may contribute aid to less fortunate nations. However, the claims that governments aspire to universal ethical principles begs the question of their primary purpose. These ethical dilemmas lead to consideration of theories of relativism (Lewis & Unerman, 1999).

In conclusion, the HF/E profession faces many ethical dilemmas. This series of articles will address some of the different ethical decisions that face researchers and practitioners.

**References**

Human Factors and Ergonomics Society. (2005). Code of ethics. Available at [*http://www.hfes.org/web/AboutHFES/ethics.html*](http://www.hfes.org/web/AboutHFES/ethics.html).

Kohlberg, L. (1984). *Essays on moral development*. New York: Harper and Row.

Lewis, L., & Unerman, J. (1999). Ethical relativism: A reason for differences in corporate social reporting. *Critical Perspectives in Accounting*, *10*, 521–547.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*, 370–396.

Rest, J., Narvaez, D., Bebeau, M., & Thoma, S. (1999). DIT-2: Devising and testing a new instrument of moral judgment. *Journal of Educational Psychology, 91*(4), 644–659.

**Chapter 25**

**Product and Process Evaluation Using the Six Us**

**(with Marc Resnick)**

**Introduction**

This article is about product and process test and evaluation. In order to ensure that evaluators do a thorough job, it often helps to have a mnemonic aid that reminds them to cover the full breadth, depth, and scope that is appropriate for the product’s expected use or misuse. This paper presents an approach called the “6Us” that targets the breadth dimension. It can be built on top of existing breadth and depth mnemonics as described below. The six “Us” are: Utility, Users (Misusers), Usage, Utilization, Usability (Misuse) and User error. It can be used in the evaluation of anything that a person puts his or her eyes and hands on, ranging from consumer products such as a TV remote control or a web site to enterprise systems such as air traffic control systems or space suits. The fundamental question is how we can focus the testing and evaluation process to fully understand how all sorts of people interact with these products in sometimes difficult contexts and with time constraints. Use of the 6Us assures a comprehensive investigation especially when the depth of analysis is pursued through the 5 Whys.

**Precedents**

Why 6Us? There are many precedents in system operation and analysis that use mnemonics and checklists to guide behavior and assure appropriate performance. One classic tool in accident investigation is the 5Ws – “What? Where? When? Who? Why?” and sometimes a “How?” for good measure. This approach promotes a breadth of analysis, reminding the investigator to address all, or at least most of the possible factors that might have contributed to an accident. For example, a car accident may involve a single vehicle, on a narrow road, at night, with a teenage driver, because of distraction and involving a sideways skid then a rollover. Further analysis could identify the fact that the road was icy and the driver was speeding. This does not lead to the root cause because it lacks depth. There is a need for a follow up process to push the depth of the analysis – the 5 Whys? The 5 Whys process (whose origins along with those of many similar simple analytic tools are to be found in the Toyota Production System) requires the analyst to ask “Why?” about 5 times for every possible sequence leading to a root cause of failure. For example in the car accident case the 5Ws, the “Who?” may lead to the reasons for the driver driving too quickly or erratically. The 5 Whys process can have many branches and often leads to links between branches. By combining the 5Ws with the 5Whys, the testing and evaluation process achieves both depth and breadth. There are many other memory jogging precedents that are customized for particular domains; the aviation world is full of them. One uses “CRAFTE” when getting a clearance from Air Traffic Control – clearance, route, altitude, communication frequency, transponder, and when to expect further clearance. The 6Ts are used when entering a holding pattern – Time (start the stop clock), Turn (to the appropriate heading), Twist (the navigation aids to the correct orientation), Throttle (check the airspeed), Tires (don’t forget to put the wheels down!!) and finally Talk (tell Air Traffic Control where you are and what you are doing). Like all facilitators (instructions, warnings, labels, procedures), mnemonics are generally most useful for less trained and less experienced investigators or operators. However they may also serve to help more experienced people avoid bad habits and shortcuts.

Usability testing has been a fundamental purpose of ergonomics for decades. Jones (1963) coined the term “fitting trials” to describe a process of systematically varying parameters of a workspace to obtain subjective opinions regarding the convenience of a particular item of furniture to complement design based purely on anthropometry. Kirk et al (1969) similarly varied the height of a chair surface, using psychophysical methods developed by Weber and Fechner 150 years ago, to obtain subjective comfort values. Moving forward to 2005 Stanton et al, described a systematic method of user trials for interface design. This method involves a relatively sophisticated process necessary to assure reliable measures of human performance. Almost every recent issue of Ergonomics in Design has an article on the usability of this that or the other product or process (Changzu and Liu, 2007; Klein and Isaacson, 2003; Torres, Heck, Rudd and Kelly, 2008; Law, Yi, Choi and Jacko, 2007; Dong and Cassim, 2007.) The Usability Professional Association now has thousands of members who are particularly, but not exclusively, focused on modern technology. Usability assessment has indeed become a household concept, with both elaborate and time consuming techniques at one end of the analysis spectrum and, at the other end of the spectrum, quick and dirty, but sufficiently reliable methods. Erlandson (2008) wrote an excellent text on “Universal and Accessible Design”, in which he systematically described two dozen strategies to accommodate the human user.

Ergonomics and Human Factors have many analytic processes – both simple and complex. For example the HFACS (Human Factors Analysis and Classification System) (Shappell and Weigman, 2000) reminds the analyst to attend to the primary accident cause (usually an operator error), the context (including the designed interfaces and environment), the operator supervision (including training and monitoring) and finally the organizational safety climate and processes. This analytic process delves down into these major causal categories and identifies in depth the various human factors applicable to each level. Manufacturing ergonomics also has many checklists and worksheets that address such things as posture, force and repetition, with various levels of detail. The NIOSH Lift Equation for example combines half a dozen or so relevant quantities to produce a comprehensive risk index. Cognitive task analysis tools address sensory, perceptual, memory, problem solving, decision making, and effector components. The NASA TLX tool provides a comprehensive set of questions that target the operators’ subjective rating of physical, mental and temporal demands, and associated performance, effort and frustration ratings. Situation Awareness tools address the operator, the situation and the information content of the task. A comprehensive survey of many of these tools is presented in the “Human Performance Measures Handbook” (Gawron, 2000).

But these methods often miss an important third dimension – scope. The scope dimension accounts for the fact that many products and processes are used in a wide variety of environments and usage scenarios. Different scenarios often bring in different human factors and ergonomics challenges. The 6Us can help focus the analysis in this third dimension. The remainder of this paper will describe the process and how it can build on the 5Ws for breadth and the 5 Whys for depth.

**The Six Us**

The 6Us comprise Utility, Users (and Misusers) , Usage, Utilization, Usability (and Misuse)and User error.

* Utility addresses the purpose of the device or service. This is often the province of a designer, engineer or marketing specialist. If we have mice, we need a mousetrap. But what we really need is a better mousetrap and that’s where the ergonomist can help. Human factors practitioners are particularly effective at identifying user requirements that go beyond the obvious. We realize that the mousetrap has to fit into the junk drawer for the times it is not in use. It also has to be easy to get out the old cheese smell. We also recognize the need for a safety feature that prevents us from snapping it down on our fingers by mistake when setting it up.
* Users (misusers). Here we pay attention to both the intended and unintended users and the misusers. Sometimes we have an individual in mind, but usually we define a “user population” as reflected by the characteristics and behaviors of a representative sample of that population. Frequently as in consumer product design the user population may be very large – e.g. all drivers or all users of the TV remote control. The unintended users may be small children rummaging in the cleaning closet or drivers under the influence of alcohol or drugs. Intended but less abled users may include the elderly or physically (mentally) challenged. Deliberate misusers may use pliers instead of a wrench to loosen a nut.
* The User category is also the focus of a large group of practitioners who espouse the ideal of “universal design.” This grand intention goes beyond the simplistic 5th percentile concept often used in ergonomics (and statistical) decision making and suggests that we find ways to support everyone. If we can create a design that can be used by the cognitively impaired, elderly, short, inexperienced or otherwise least capable person at a particular activity, then everyone else should be covered as well. A completely “universal” design may be an impossibly lofty ideal in practice, but considering these least capable user groups during the design process can improve the eventual usability for everyone. One strategy in product design is to provide a wide range of features that are used selectively by users with different levels of experience. The advanced features can be hidden from novices using <advanced> menu pullouts. Pedagogical vectors (Cooper et al 2007) can be used to encourage intermediate users to explore these advanced features to develop expertise. The Microsoft Windows interface is a good example of this design strategy.
* Usage describes how and in what circumstances the product or process will be used. For example a car may be used on a dark and stormy night; a wrench may be used by a space walking astronaut assembling the International Space Station, while wearing a cumbersome space suit and perching atop robotic arm; a PDA may be used in the dark and without access to a recharging outlet. A chair with casters may be used for sitting on or in order to reach something off a high shelf. These questions beg the question regarding the limitations of expected use and foreseeable misuse. Usage may be the most critical category to consider because it is these difficult usage categories that often result in the most serious consequences. From a business perspective, they also may be the best differentiator of one product from another. Every online bookstore allows the customer to find a book by title, author, or ISBN number. But which one helps me find a birthday present for my cousin that he will actually like?
* Utilization is the frequency or amount of use. Some processes, like filling in your tax returns, are done once a year. The complexity of the rules for this annual chore is such that more than 50% of tax payers send their shoe box full of receipts to an accountant or tax preparer. The problem with infrequent utilization is that the users must be current in training, unless the device or process is extremely simple. With frequent utilization the familiar user may learn to compensate for product or process shortcomings. Web pages are like that; skilled users sail through the flight booking process while unwary infrequent users may book the same flight twice, on the wrong day or to the wrong destination.
* Usability. In recent years ‘usability’ has become synonymous with ergonomics. Indeed there are now many more usability specialists than “ergonomists” largely because of the semantics and methods of practice. Perhaps the best advocate of “usability” is the “Easy Button” seen in Staples advertisements. Ironically, the advertisement acknowledges that there is no such thing as “an easy button,” rather a user with a problem will usually find a product at Staples to solve that problem. Staples, therefore, inadvertently becomes a metaphor for ergonomics – it represents the interface between a product and a user by facilitating the acquisition process. As technology allows products to support more and more functions, usability becomes more important. A useful feature on your new digital camera is useless if you can’t find it on the menu. Or if you can find it, but it takes too many key presses and your new baby isn’t smiling any more.
* User Error. As Murphy said “if it can go wrong someone will make it go wrong.” “To err is human; technology only increases the frequency and consequences of error.” The understanding, anticipation, prevention and mitigation of human error are important human factors purposes. Sometimes errors cause serious damage or injury. Sometimes they are just annoying. But they are always bad for business.

**An Example - The Notorious Remote Control**

Almost everybody watches television and chooses which program to watch by pressing one or more buttons on a remote control, which may sport anywhere up to 50 or even more buttons, each with a different function. The following qualitative analysis, using the 6Us, 5Ws, some Whys and Hows are not intended to be exhaustive; rather it is intended to illustrate the breadth and depth of the analytic process. A tabular MS Word process will be illustrated here, although a spreadsheet or concept mapping approach may be more versatile, especially where multiple branches and cross links between concepts are needed. An example worksheet is presented in Table 1.

* Utility – Of course the remote is useful – it saves you the trouble of getting off the couch. One reason for the success of the device (like the MS Windows interface) is the sheer number of potentially useful functions for control of the entertainment center. It also has many functions that may or may not be useful to many of the users.
  + Why is the remote useful? Because it saves users from having to get up and walk over to the TV to make adjustments. This is particularly useful for anyone unable or unwilling to easily stand up and sit down repeatedly.
  + Why are adjustments valuable? People seek cognitive variety. Humans have a genetic instinct to learn and a short attention span.
  + Why might people be unable to easily stand up or sit down? They may be physically challenged, fatigued, or lazy.
  + Why do people become physically challenged? They may be disabled or elderly (or watch too much television).
  + Why do people become fatigued or lazy before using their TV? They may have engaged in physical or cognitive activities that temporarily reduced their capability.
  + How do we become aware of the Utility of the remote control? Because millions of people use a variant of this product every day? Focus groups, surveys, and interviews can be used to uncover use cases and scenarios. And yet, much of the utility of remotes is never used because there is no conduit to illustrate nonobvious uses to the user.
  + What can be done to increase the utility of the remote? Add more functions. Improve the interface so each function is easier to access , expand the device so that a single remote can be used to access several electronic devices.
* Users – now here’s the big challenge – the potential users include almost everybody and from the analysis point of view that represents a very large amount of human variability.
  + Why is there so much user population variability? Because the visual and auditory medium of television is conducive to presenting information and entertainment of an almost infinite variety that appeals to users of all ages and walks of life.
  + What is the form of this variability? The population has variability in sensory, cognitive and motor functions. There are also a wide variety of scenarios in which these users choose to interact with their television.
  + What can be done to understand the capabilities of this wide range of users? Human Factors and Ergonomics practitioners have very many tools at their finger tips to extract information about users and they should be able to beat Darwin by anticipating intended use and foreseeable misuse of the product or service at hand. Ethnography and contextual inquiry can be used to identify these scenarios, even when users were themselves unaware of their own habits and needs.
* Usage – Usually the remote is used in a very comfortable and non time restricted environment, although quick reaction may be needed when adjusting the volume when the phone rings or switching back to a favorite show when the commercial ends. An environmental issue regarding usage is that it may be used with a very dim light. Another usage issue is the fact that the TV remote may be just one of many sitting on the coffee table, each of which are dissimilar in function and design, and thus confuse the user. Sometimes one remote has to be used in conjunction with others and this may add to the possible confusion.
  + Where do users use the remote? In their living rooms, kitchens, studies or bedrooms.
  + Why are there so many varieties of remote control? Because the nature of product competition and design, and some differences in function, is such that the remote (like the automobile interface) is a product differentiator. Attempts to offer “universal remotes” have had some success, but functional and esthetic differences hamper these lofty aspirations.
  + Why doesn’t technology, such as voice recognition and command, replace the manual remote control? Because this technology is not yet well developed, error proof or cost effective.
  + What can be done to manage usage? Perhaps usage should not be managed, the richer the usage the greater the information regarding parameters for robust design. Analysts should always explore intended use and foreseeable misuse.
* Utilization – The device is used very frequently by most users, but usually for a very limited number of functions.
  + Why are only a few functions used frequently? Because as with much contemporary technology (TV, computer, PDA, Cell Phone) the designers and marketers seek to include as much functionality as possible with the aim of satisfying the broadest possible range of utility; consequently the core functions are used frequently and the “elective” functions less frequently.
  + Why are the elective functions used less frequently? Because the particular users either don’t need, don’t know about or don’t understand the interfaces/ procedures for all forms of elective functionality.
  + Why don’t these users understand the interfaces? Because the ancillary training material – users guides, help cards etc are either not available or not easily understood. It is rare that the utility of the remote is valuable enough to warrant significant effort at learning how to use the functions.
  + Why are help cards not available? Because they are not integral to the product and get misplaced.
  + What can be done to manage utilization? The use of remote control devices will continue to increase due to technology innovation in vehicles, kitchen and space craft. Conventional hard buttoned remote control devices will give way to devices with adaptive soft keys or voice interfaces that penetrate more deeply into the cognitive interface and capabilities of a wider variety of users.
* Usability – The principle challenge is a cognitive one – there are very many functions and many of them may be used very rarely, if not at all. The interface is also enhanced by functional grouping, shape, labeling and color of the buttons, although this coding may not be familiar to all users. Timely feedback from the screen (or DVD unit) is usually associated with appropriate commands and sometimes inappropriate commands – e.g. channel up or down. Usability is often enhanced by the provision of facilitators, such as a user’s manual or a descriptive card (both of which may not always be readily accessible.) Users may supplement these by affixing their own cheat sheet on the back of the remote.
  + Why don’t the interface design strategies resolve the cognitive challenge? A short answer is that they do, to some extent, but wide “expertise” in the use of the remote control can only be achieved by study or frequent use of all the features.
  + Why only “to some extent?” The purpose of a user interface is to assist user sensory, cognitive and motor functions with a product. The sensory and motor functions are usually fairly easily addressed. The cognitive interface is always a greater challenge for designers because of the desirability to accommodate an extremely wide range of functions and users.
  + How can a designer resolve this cognitive interface challenge? By putting “intelligence” into the interface that adapts the interface to the current purpose. Such a strategy is achieved with contemporary soft keys in PDAs which can indicate choices and sequences in contrast to the hard keys in most remotes.
  + What can be done to improve usability of the remote control? By carrying out ergonomics and ethnographic studies of usage in a wide variety of conditions. By analyzing the customer feedback and customer errors. As monopolies develop in the entertainment world and the human factors message is spread, there will be a tendency to a reduction of the variety of devices.
* User Error – The first class of user error includes lost device and batteries worn out. User error may be caused by not understanding the function of buttons or sequences. Another error type occurs when the wrong button is pressed. In some cases error recovery sequences may be unclear and the user may have to rely on the time honored process of switching the whole unit off and starting again.
  + Why do users make so many errors with the remote yet continue to use it regularly? They avoid errors by limiting the number of functions that they habitually use. Also the feedback from the screen facilitates recovery in many instances. There are also no alternatives to the remote that support the desired utility.
  + Why do some users make more errors than others? Usually because they had a greater amount of transfer of training from similar devices or because they took time to read the user’s manual or kept the facilitator card close at hand.
  + How can appropriate training be achieved? The first approach is to make the interface intuitive and compatible with the task at hand. The second step is to make the device either very similar to previous devices or very different to avoid negative transfer.
  + Does user error really matter if the devices continue to be used? Perhaps not, the market and Darwin will eventually win. But user error is our bread and butter, we are paid to analyze, anticipate, prevent and mitigate. And as third party universal remotes gain popularity, competition creates a business imperative to create effective designs that users will like.
  + What can be done to prevent error? By having adaptable interfaces and built in instructions. (Did anyone say standardization, I hope not!!)

**Conclusions**

This article is an attempt to describe a systematic, logical, yet qualitative process of investigating any product or service activity that is used by a variety of people in a variety of contexts. The illustrated analysis of the remote control focuses on the exploration of potential challenges to the design. Of course an important next step is to develop potential solutions to these challenges. The approach can be applied both in error / incident / accident investigation and in the design of error proof products and services, or at least products and services where the effects of errors are mitigated. The approach can be used for kitchen utensils and space craft, and web pages and nuclear plant control rooms. The approach can be quantified, at least to the ordinal level by adding probability, exposure and consequence scores to each of the questions, thus leading to a versatile risk – benefit analysis. The approach can be framed as lists, tables or branching networks. But most of all the approach is ergonomic; it is simple, intuitive and effective.

**Product or Process Evaluation Sheet**

**Product Analyst Date Reference#**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Product or Process Description** | **Why?** | **Why?** | **Product or Process Analysis** | **Why?** | **Why?** | **Why?** | **Why?** |
| **Utility** | Why is the product or process useful? |  |  | Why should the product be improved? |  |  |  |  |
| **Usage** | In what way will the product or process be used |  |  | In what context will the product be used? |  |  |  |  |
| **Utilization** | How often and by how many people will the product or process be used? |  |  | What are the probability and frequency of failure? |  |  |  |  |
| **User** | Who is the intended user? |  |  | How can the intended user be selected or trained to use the product or process? |  |  |  |  |
| **Misuser** | Who is the expected misuser? |  |  | What kinds of users will be associated with these failure modes? |  |  |  |  |
| **Usability** | How easy is it to use the product or process |  |  | How could the product or process be changed to make it easier to use? |  |  |  |  |
| **Misuse** | How easy is it to misuse the product or process |  |  | How could the product or process be changed to prevent and mitigate the effects of misuse? |  |  |  |  |
| **User Error** | What kinds of failure modes can be predicted? |  |  | What are the consequences of failure? |  |  |  |  |

Table 1. A checklist / worksheet for product or process evaluation

**References**

Cooper, A., Reimann, R., and Cronin, D. (2007) About Face. 3rd Edition. Wiley.

Dong, H and Cassim, J ((2007) Critical User Forums for Inclusive Design, Ergonomics in Design, Vol. 15, No 4

Erlandson, R.F. (2008) Universal and Accessible Design for Products, Services and Processes”. CRC PressLondon.

Jones, J. C. (1963) Fitting Trials: A method of fitting equipment dimensions to variations in the activities, comfort requirements and body size of users, The Architects Journal (6 February 1963), pp. 321-325

.

Kirk, N. S., Ward, J. S., Asprey, Baker, E. and Peacock, J. B. (1969) "Discrimination of Chair Seat Heights": Ergonomics, Volume. 12, no. 3, pp.403-414

Klein, H. A. and Isaacson, J. J., ((2003) Making Medication Instructions more Usable Ergonomics in design Vol. 11, No 2

Law, C. M., Yi, J. S., Choi, Y. S., and Jacko J. A. (2007) Unresolved Problems in Accessibility and Universal Design Guidelines, Ergonomics in Design Vol 15, No 3

Shappell, S. A. and Wiegmann, D. A. (2000) The Human Factors Analysis and Classification System, DOD/FAA/AM-00/7 Office of Aviation Medicine, Washington, DC.

Stanton, N.A., Salmon, P.M., Walker, G.H., Baber, C and Jenkins, D.P (2005) “Human Factors Methods – Chapter 10 “Interface Design” , Ashgate, Aldershot, UK.

Torres , R.J. Heck, M. P., Rudd, J. R. and Kelley, J. F. (2008) Vol. 16, No. 2

Wu, C. X and Liu, Y (2007) Usability Makeover of a Cognitive Modeling Tool, Ergonomics in Design Vol. 15, No 2

**Chapter 26**

**Time for Bed: Shift Work**

This article is about sleep, factors that facilitate good sleep and the personal and operational challenges of shift work.

*"Early to bed, early to rise, makes a girl [with apologies to Benjamin Franklin] healthy, wealthy and wise."*

"It’s bed time."

"I’m not tired."

"You will be in the morning."

"I’m thirsty."

"Have a drink of milk and then go to bed."

"It’s still light outside."

"I’ll close the curtains."

"I’m scared of the dark."

"I’ll put the night light on."

"May I stay up late on Saturday?"

"All right, if you are a good girl."

"Will you read me a story?"

"If you wish. Once upon a time there were two fairies called Melatonin and Zeitgeber."

"Those are silly names for fairies."

"Zeitgebers are things like light and dark and suppertime and catching the bus to school at seven o’clock in the morning, and Melatonin is a chemical that fills your blood at night when you sleep."

"This is a boring story, I’m going to take a sleeping pill and dream about interesting things like REM."

Except at the poles or in space, day and night are facts of life. Sleep cycles are driven by an internal clock, our external context – zeitgebers – such as light and dark (Boyce, P. R. 2003)., and our daily habits. If we travel to the other side of the world, after a few days, we will adapt to our new sleep cycle, as shown by the physiological indicators of our circadian rhythms, such as body temperature and melatonin levels (Colquohoun, W. P., & Rutenfranz,J., 1980, Scott, J. A. 1990).

**Sleep and Sleep Disorders**

Normal sleep lasts between 4 and 10 hours; 8 hours is the commonly accepted normal duration, but there are considerable individual differences. Sleep is necessary for recuperation. Sleep consists of repeated (four or five) cycles of five stages, with each cycle lasting about 1.5 to 2.0 hours. The stage lengths vary as the night progresses.

Stage 1 is a light sleep, characterized by slow eye and body movements and low-frequency EEG waves (2–7 Hz). Stage 2 represents about 50% overall of a night’s sleep; during this stage, the eye movements stop and K complexes (sharp negative waves followed by a slower positive one) are observed. *Sleep spindles* occur in the 12–14-Hz range. Stage 3 is the first stage of deep sleep and is characterized by very slow (less than 2-Hz) waves interspersed with spikes of faster waves (spindles, 14–16 Hz). Stage 4 is similar to Stage 3 but with a larger number of slower waves. Stage 5 sleep involves rapid eye movements (REM), an increase in blood pressure and heart rate, and, often, dreaming **(Susmakova, K. 2004).**.

The electrical activity during wakefulness is variable but usually between 20 and 30 Hz. It is common during the day for individuals to go through similar cycles of alertness and drowsiness, although these are usually dominated by physical, cognitive, and environmental stimulation. If allowed to run without strong external stimuli (e.g., light), the *circadian cycle* is about 25 hours. Detailed experimental work of circadian rhythm adaptation is carried out at the Brigham and Women’s Hospital Center for Sleep Research in Boston, Massachusetts (Dijk & Ceisler, 1994).

There are numerous sleep disorders, some of which are clinically serious and some of which may have considerable detrimental effects on performance (Pilcher & Huffcutt, 1996, Cavallo, A., Jaskiewicz, J., & Ris, M. D., 2002). These disorders may be treated by medication, but many are responsive to behavioral intervention associated with diet, exercise, and sleep routines. Chronic sleep loss or disruptions by other temporal changes, such as shift work, are best addressed by removal of the underlying lifestyle causes.

**Adaptation, Flexibility, and Variability**

Modern society demands services and productivity that do not heed the general human requirement for stable sleep patterns. We like our hospital, police, transportation, maintenance, and energy services to run 24/7. Expensive factory equipment must run around the clock to manufacture its products economically. Astronauts on the International Space Station must deal with a 90-min light and dark cycle by adapting to planned sleep and work cycles and using artificial (blue) light. Because people are adaptable – they can permanently change their habits and rhythms after a few days – and flexible – they can tolerate short-term deviations from their circadian schedule – shift work has become a universal and generally acceptable practice. But there are good and bad shift schedules, and tolerable and dangerous outcomes.

Adaptation and flexibility have a price. Deviation from habitual rhythms or zeitgebers affects peoples' mood, behavior, performance, and, eventually, health. Classic jokes suggest that after a late night, a cup of coffee may be needed both to start a peristaltic wave and to turn us into civil humans being ready for what the day will bring. Coffee at the other end of the day, such as before a long drive home, may have its merits, but eventually the "sleep debt" must be paid. Sleep loss is cumulative. Performance deteriorates.

One certainty, until cloning becomes viable, is that all people are different in their chronobiology and adaptation to shift work (Smith, Jeppesen, & Boggild, 2007). These differences are to be found in physical and mental characteristics, in behaviors, and, in the present context, in their circadian rhythms. Much has been written about "larks" and "owls," or morning and evening types. Whereas there are measurable physiological differences between these two types on a continuum, it is likely that habit is the most important determinant of these circadian differences in behavior. These adaptations are important because they suit people for different circadian responsibilities; however, mismatches can lead to degraded performance in critical tasks. Nowhere is this clearer than when a habitual "morning type" is asked to work or drive late at night.

**Shift Scheduling**

A shift schedule is primarily dependent on operational requirements, although physiological and social factors may be considered to make the arrangements more tolerable. Two principal variables need to be taken into account in a schedule: First, there is the length of a shift, usually somewhere between 4 and 12 hours, although they may be as short as half an hour or up to 24 hours. The second variable is the rotation cycle. This may vary from a fixed cycle, (e.g., permanent nights) to a one-month or one-week rotation between the day and night shift.

From the physiological point of view, it is best to have a permanent shift assignment, because this imposes fewer demands for change in the circadian cycle – provided that daytime sleep is not hindered by light, noise, and other distractions. It is common to see one month, one week, or two- or three-day cycles.

Physiologically, the one-week cycle is probably the worst arrangement, as this may cause undulating changes in the circadian cycle. The one-month cycle is generally preferred, but a very short cycle – such as two days, two nights, and one day off – is usually tolerable because it minimizes the effect on the underlying physiology. Miller (2008) presented a very detailed account of scheduling principles.

Social and operational factors also play a part in choice of schedule. For example, in the 1980s, at the Saturn auto plant in Tennessee, the associates opted for a one-week rotation between days and nights (the worst physiological cycle), which was implemented out of a desire to be fair to all the workers and allow them to share the domestic and social benefits of day and night work. While in the military in the 1950s, I had a three-week cycle between mornings, afternoons, and nights, with a long weekend between the night and morning phases. This long weekend plus the "forward rotation" created a tolerable pattern.

A municipal police force asked for an investigation of the advantages and disadvantages of changing from a twelve-day, 8-hr cycle to an eight-day, 12-hour cycle (Peacock, Glube, Miller, & Clune, 1983) It was impossible to carry out reliable operational measures (such as number of arrests), but it was feasible to do a study, before and after the change, of physiological factors (blood pressure, exercise tolerance) and before and after-shift surveys of cognitive performance, sleep quantity and quality, and social and domestic desirability of the new 12-hour shift pattern. Urinary catecholamine measures proved to be both unpopular and not useful predictors of the circadian physiology. One notable observation was that in this short-cycle arrangement, circadian physiological factors trumped sleep deprivation factors in the mornings. Apart from body temperature, the most reliable performance differentiation measure was critical flicker fusion frequency. The conclusion of the study was that the 12-hour, eight-day cycle was not worse than the 8 hour, twelve-day cycle from the physiological and performance viewpoints, and it was the generally preferred choice.

Shift changeover and work handoff may present problems in various organizations. It may be necessary to schedule a certain amount of overlap for continuity purposes, at least for key people. On the other hand, in large organizations, it may be appropriate to have a gap between shifts to allow the parking lot to clear before the new shift arrives. The number of crews needed to maintain coverage over the period for which the operation is required also produces scheduling constraints. One solution to this problem is to offer variable work duration options, such as 20-, 30-, and 40-hour work weeks for different employees.

The choice of shift schedule may be affected by less obvious operational and social factors. For example, some people opt for the night shift because of a wage premium or because the management oversight is less stringent. Others choose a particular shift to fit in with a spouse's employment. Many people prefer a four-day work week, with longer shifts, in order to pursue recreational activities or other jobs. These choices suggest that shift workers will often weight social and operational factors over physiological needs.

**Jet Lag**

The progress of globalization brings requirements for transoceanic and transcontinental travel, despite the promise of the Internet. This travel may involve 6 to 12 time zones. Even transcontinental travel over 3 to 5 time zones may be noticeable. People who make these journeys usually encounter the combined effects of circadian disruption and sleep loss caused by the difficulty of getting adequate sleep in an airplane seat. Also, the length of stay in the remote location will vary from a few days to a few weeks, thus incurring a circadian rhythm inversion.

The effects of jet lag are insidious but real. Driving a car or attending a meeting following a long flight may result in subpar performance. During the late 1990s, I often traveled between Detroit and the Opel headquarters in Germany. Autobahn driving, meetings involving both English and German language, and different historical car assembly processes all demanded considerable alertness. Where possible, rest or recovery days should be inserted following a long flight.

International or east-to-west-coast business activities may be accomplished by telecommuting. However, occasionally there needs to be video or telephone conferencing that compromises circadian physiology. Design meetings at 3:00 a.m. are not fun and may put one party at a psychological or cognitive disadvantage. Pilots also suffer jet lag, but scheduling regulations for flight and cabin crew attempt to minimize the physiological disruption. However, as the range of flights becomes greater, attention to schedule and monitoring of performance and health will become more imperative.

**Type of Work**

All work is not equal, at least not in terms of the need for sleep. Unusually physically demanding work creates physical fatigue, which may hasten the desire to sleep. Heavy lunches are also followed by a feeling of drowsiness. Some work, such as police patrols, involve long periods of relatively low stimulation while cruising the streets, followed by occasional, sometimes sudden, demands for high levels of alertness and physical effort. The same is true for firefighters. Long-distance truck drivers and airline pilots spend much of their time with very low levels of workload, except for the demands of dealing with bad weather or, in the case of pilots, during the approach and landing phases.

On the other hand, jobs such as nursing or automobile assembly require continuous walking. Parcel delivery drivers alternate between the high attention demands of city traffic and the high physical demands of package delivery. Perhaps the most attention-demanding shift work job is air traffic control. The archaic requirement of medical residents to spend 24 hours on duty has fortunately changed in recent years to a more physiologically appropriate and safe schedule (Landrigan et al., 2004).

Shift schedules for these different work types may differ considerably. Where there is substantial and continuous physical or mental demand, shorter shifts are in order. Where the work offers alternating periods of low and high physical or mental demand, longer shift lengths may be acceptable. In such cases, there are many proponents of "power naps." The siesta concept that involves afternoon rest between morning and evening work and play may return to acceptability in modern society.

**Conclusions**

Society-driven operational demands for shift work will continue because of the general tolerance and motivation of individuals toward these arrangements. However, the physiological disruptions to circadian rhythms are such that in the long term, mood, behavior, performance, safety, and health may deteriorate.

The following general rules should be considered when designing or evaluating a shift system to make these disruptions more physiologically tolerable:

1. Recognize the physiological, domestic, social, performance, safety, mood, and health effects of circadian disruption and sleep deprivation.
2. Consider the physical and cognitive work content.
3. Allow employees some individual choice in workday or shift assignments.
4. Choose longer shift durations (10–12 hours) when the task involves a variety of physical and cognitive work.
5. Choose shorter (4–8-hour) shift durations when the physical or cognitive load is high and continuous.
6. Institute a semipermanent (at least one-month) shift assignment with volunteers for the night shift.
7. Institute a rapidly forward rotating schedule
   1. 8 hour – DDEENNOO – 7am – 3pm, 3pm – 11pm, 11pm – 7am
   2. 8 hour - DDEEOOO – 7am – 3pm, 3pm – 11pm
   3. 12 hr – DDNNOOO 7am – 7pm, 7pm – 7am
8. Start and end shifts to avoid rush hour, but do not start or end shifts during the early morning hours of 12:00 to 6:00 a.m.
9. Routinely monitor the performance, safety, mood, and health of shift workers.

"May I have my sleeping pill now, please?"

"No, sleeping pills are not usually the answer, just pay attention to your zeitgebers, and, if all else fails, try counting sheep."

**References**

Boyce, P. R. (2003). *Human factors in lighting.* In *The circadian system* (pp. XXX–XX).. London: Taylor & Francis.

Cavallo, A., Jaskiewicz, J., & Ris, M. D. (2002). Impact of night-float rotation on sleep, mood, and alertness: The resident's perception. [*Chronobiology International: The Journal of Biological and Medical Rhythm Research*](http://www.informaworld.com/smpp/title~content=t713597233~db=all), *19,* 893–902.

Dijk, D. J., & Ceisler, C. A. (1994). Paradoxical timing of the circadian rhythm of sleep propensity serves to consolidate sleep and wakefulness in humans. [*Neurosci Lett*](javascript:AL_get(this,%20'jour',%20'Neurosci%20Lett.');)*166*(1), 63–68. Yes

Colquohoun, W. P., & Rutenfranz,J. (1980). *Studies of shiftwork*. London: Taylor & Francis.

Landrigan, C. P., Rothschild, J. M., Cronin, J. W., Kaushal, R., Burdick, E., Katz, J. T., Lilly, C. M., Stone, P. H., Lockley, S. W., Bates, D. W., & Czeisler, C. A. (2004). Effect of reducing interns' work hours on serious medical errors in intensive care units. *New England Journal of Medicine, 351*, 1838–1848.

Miller, J. C. (2008). Fundamentals of shiftwork scheduling. *Ergonomics in Design,* *16*(3), 13–17.

Peacock, B., Glube, R., Miller, M., & Clune, P. (1983). Police officers response to eight and twelve hour shift schedules. *Ergonomics, 26*, 479–493. 1983 is correct

Pilcher, J. J., & Huffcutt, A. I. (1996). Effects of sleep deprivation on performance: A meta-analysis. *Sleep, 19*(4) 318–326

Scott, J. A. (1990). *Shiftwork: State of the art reviews in occupational medicine*. Philadelphia: Hanley and Belfus.

# Smith, L., Jeppesen, H. J., & Boggild, H. (2007). Internal locus of control and choice in health service shift workers. [*Ergonomics*](http://www.informaworld.com/smpp/title~content=t713701117~db=all)*,* [*50*](http://www.informaworld.com/smpp/title~content=t713701117~db=all~tab=issueslist~branches=50#v50), 1485–1502.

# Susmakova, K. (2004). Human sleep and sleep EEG. *Measurement Science Review, 4*(2), 59–74.

****