# Human Computer Interaction

## Fundamentals and Practice [ SWE - 431 ]

Gerard Jounghyun Kim

## Chapter: 3
## Human Factors as HCI Theories

Mahfuzur Rahman Emon
Lecturer, IICT, SUST

# Human Information Processing

As the main underlying theory for HCI, human factors can largely be divided into:

- **Cognitive Science:** The primary focus of cognitive science is to understand how the human mind works, how people think, learn, remember, and process information. It explains the human's capability and model of conscious processing of high-level information.

- **Ergonomics:** Which elucidates how raw external stimulation signals are accepted by our five senses, are processed up to preattentive level, and are later acted upon in the outer world through the motor organs.

Human-factors knowledge will particularly help us design HCI in the following ways.

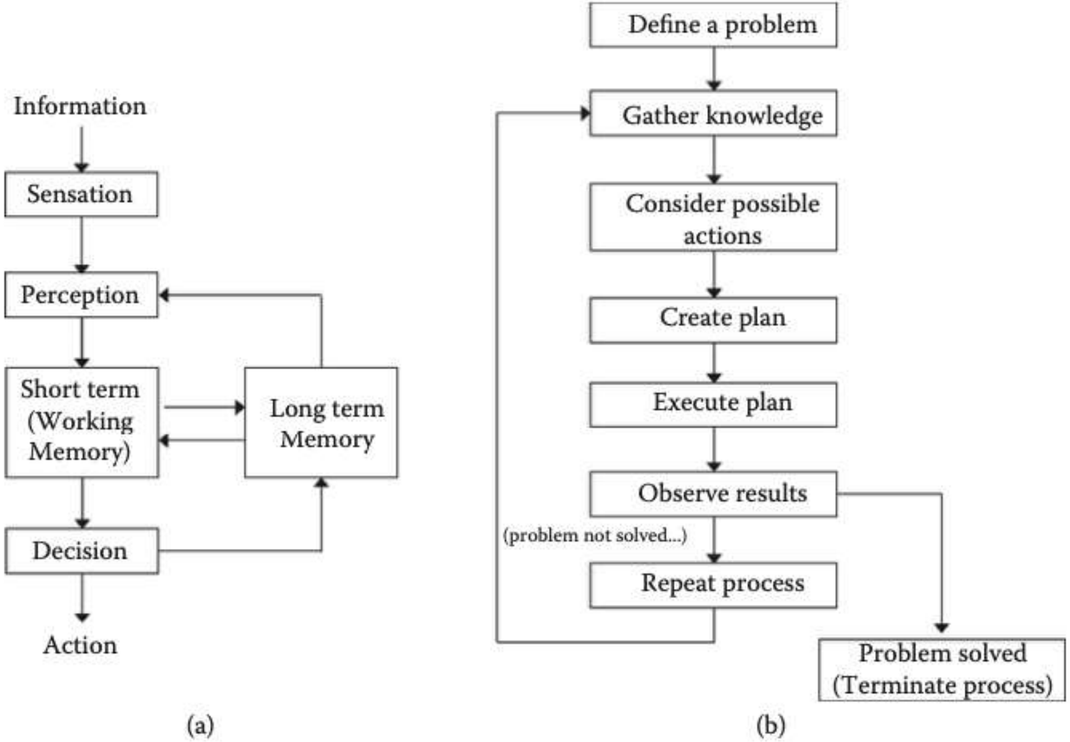
- **Task/interaction modeling:** Formulate the steps for how humans might interact to solve and carry out a given task/problem and derive the interaction model. Designer must obtain the model by his knowledge in cognitive science and direct observation of the user.

- **Prediction, assessment, and evaluation of interactive behavior**: Understand and predict how humans might react mentally to various information-presentation as a basis for interface selection

## Task Modeling and Human Problem-Solving model

- A goal of a word-processing system might be to produce a nice-looking document as easily as possible.
- In more abstract terms, this whole process of interaction could be viewed as a human attempting to solve a “**problem**” and applying certain “**actions**” on “**objects**” to arrive at a final “**solution**.”
- Cognitive science has investigated the ways in which humans solve problems, this helps HCI designers to analyze the task and base the interaction model around the natural problem-solving process.
- Thus for a smaller problem of “fixing the font,” the action could be a “menu item selection” applied to a “highlighted text.”

There are several "**human problem-solving**" models that are put forth by a number of researchers, but most of them can be collectively summarized as:



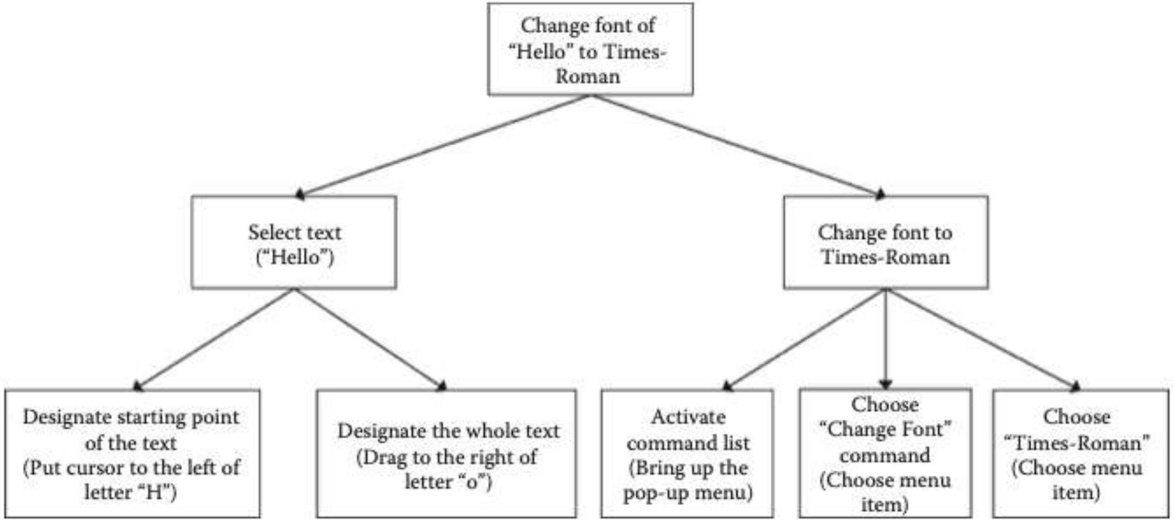

a. The overall human problem-solving model and process
b. A more detailed view of the "decision maker/executor."

Human problem-solving or information processing efforts consist of these important parts:

- **Sensation:** Which senses external information (e.g., visual, aural, haptic), and Perception, which interprets and extracts basic meanings of the external information.

- **Memory**: Which stores momentary and short-term information or long-term knowledge. This knowledge includes information about the external world, procedures, rules, relations, schemas, candidates of actions to apply, the current objective, the plan of action etc.

- **Decision maker/executor:** which formulates and revises a “plan,” then decides what to do based on the various knowledge in the memory, and finally acts it out by commanding the motor system.




The user's intention defines the problem, establishing it as the top goal. A hierarchical plan is then created by breaking down the top goal into subgoals. Subtasks are identified to address these subgoals, considering the external situation. The process involves enacting these subtasks, though success is not guaranteed. The cycle continues with observation, plan revision, and restoration until the top goal is achieved.

## Human Reaction and Prediction of Cognitive Performance

- The success of human interface design relies on understanding cognitive aspects.

- The "**gulf of execution/evaluation**" - introduced by Norman and Draper states that users can be left bewildered (and not perform very well) when an interactive system does not offer certain actions or does not result in a state as expected by the user. The mismatch between the user's mental model and the task model employed by the interactive system creates the "**gulf**."

- Memory capacity, encompassing short-term and long-term memory, plays a crucial role. Short-term memory limitations require interfaces to present information effectively, especially in multitasking scenarios where context switches can degrade performance. Interfaces should assist users by capturing context information during suspension and highlighting it upon resumption.


Gulf of execution and evaluation: the gap between the expected and actual.

## Predictive Performance Assessment: GOMS

- Many important cognitive activities have been analyzed in terms of their typical approximate process time. E.g. for single-chunk retrieval from the short-term memory, encoding (memorizing) of information into the long-term memory, responding to a visual stimulus and interpreting its content, etc.

- Based on these figures and a task-sequence model, one might be able to quantitatively estimate the time taken to complete a given task and, therefore, make an evaluation with regard to the original performance requirements

| TYPE OF OPERATION | TIME ESTIMATE |
|---|---|
| K: Keyboard input | Expert: 0.12 s<br>Average: 0.20 s<br>Novice: 1.2 s |
| T(*n*): Type *n* characters | 280 × *n* ms |
| P: Point with mouse to something on the display | 1100 ms |
| B: Press or release mouse button | 100 ms |
| BB: Click a mouse button (press and release) | 200 ms |
| H: Home hands, either to the keyboard or mouse | 400 ms |
| M: Thinking what to do (mental operator) | 1200 ms<br>(can change) |
| W(*t*): Waiting for the system (to respond) | *t* ms |

Estimates of time taken for typical desktop computers operations from GOMS

GOMS (Goals, Operators, Methods, and Selection rules) methodology involves hierarchical task modeling, mapping operators to subtasks, and calculating total task performance time.

Developed for desktop computing, it remains valid with figures for mouse clicks, keyboard input, and mental operators.

| DESIGN 1 | | DESIGN 2[a] | |
|---|---|---|---|
| 1. Point to file icon | P | 1. Point to file icon | P |
| 2. Click mouse button | BB | 2. Click mouse button | BB |
| 3. Point to file menu | P | 3. Move hand to keyboard | M |
| 4. Press and hold mouse button | B | 4. Hit command key: command-T | KK |
| 5. Point to DELETE item | P | 5. Move hand back to mouse | H |
| 6. Release mouse button | B | | |
| 7. Point to original window | P | | |
| Total time = 4.8 s | | Total time = 2.66 s | |

*Note:* The total time is computed by adding the corresponding figures in Table 3.1.
[a] Design 2 is the “expert” version that uses a hot key [7].

Estimates of Time Taken for Two Tasks Models of “Deleting a File”

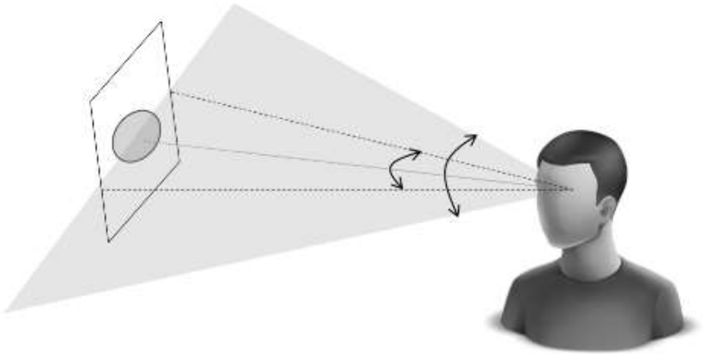
# Sensation and Perception of Information

## Visual

By far the most important information medium. the parameters of the visual interface design and display system will have to conform to the capacity and characteristics of the human visual system

***<u>Visual and Display Parameter:</u>***

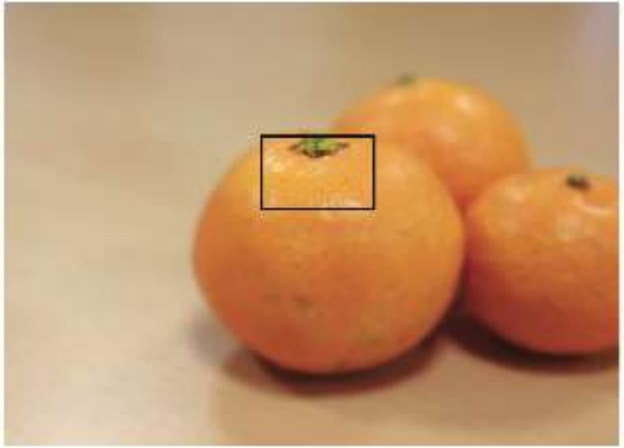
- **Field of View ( FOV ):** Refers to the angle covered by the visible area for a human user, either in the horizontal or vertical direction. The shaded area depicts the horizontal field of view. The human FOV is approximately 180° in both horizontal and vertical directions.



- **Viewing distance:** The perpendicular distance to the surface of the display. Viewing distance may change with user movements. However, one might be able to define a nominal and typical viewing distance for a given task or operating environment.
- **Display field of view:** Refers to the angle covered by the display area when viewed from a specific distance. It is important to note that for a fixed display area, the display FOV will vary at different viewing distances.
- **Pixel**: A display system is typically composed of an array of small rectangular areas called pixels.
- **Display resolution:** This is the number of pixels in the horizontal and vertical directions for a fixed area.
- **Visual acuity:** In effect, this is the resolution perceivable by the human eye from a fixed distance. This is also synonymous with the power of sight, which is different for different people and age groups.

***Detail and Peripheral Vision:***

- The human eye has cones for color and detail in the central retina (5° in FOV).
- Rods for motion detection in the periphery.
- Displays usually have uniform resolution, unlike the varying sensitivity of human vision.
- Adjusting object details based on user focus can make rendering more efficient.
- Large, high-resolution displays may have parts unused unless viewed simultaneously. Using smaller high-resolution displays at close distances can be more economical. Microsoft's Illumiroom combines a high-resolution display with low-resolution projection for an immersive experience.



( a )

( b )

a. An ideal display that would provide relatively higher resolution in the area of the user's focus
b. A large immersive display as realized by a high-resolution monitor in the middle and low resolution projection in the periphery

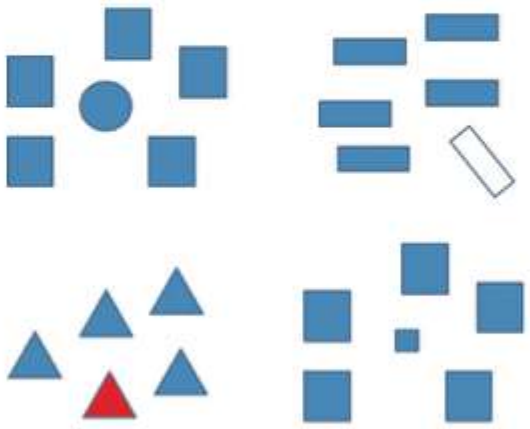
A large, tiled, high-resolution display. Is it really worth the cost?

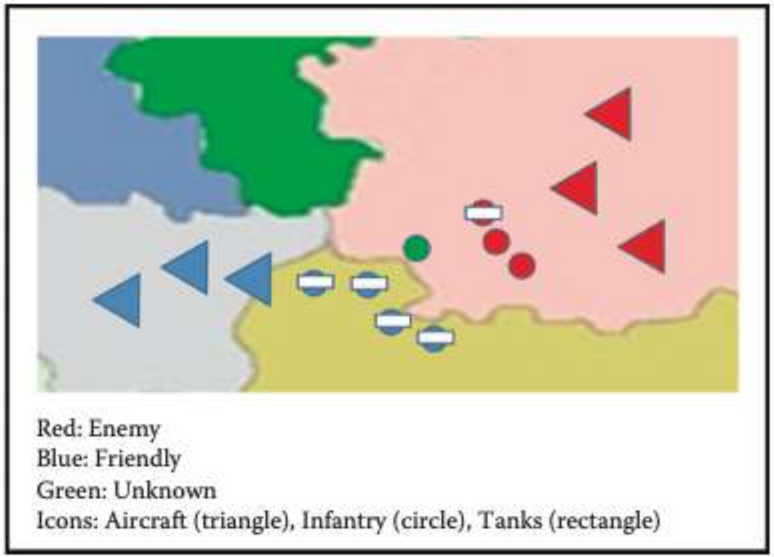
## ***Color, Brightness and Contrast:***

- **Brightness**: The amount of light energy emitted by the object.
- **Color**: Human response to different wavelengths of light, namely for those corresponding to red, green, blue, and their mixtures.
- **Contrast**: The difference in brightness or color between two things you see. For brightness, it's about how much light one thing has compared to another. They suggest a ratio of at least 3:1 for brightness difference. Color contrast is about differences in color hue and saturation. Brightness contrast is better for seeing details compared to color contrast.

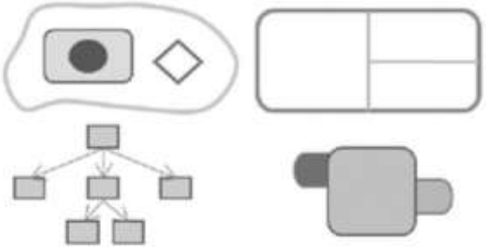
***<u>Pre-Attentive Features and High-Level Diagrammatic Semantics:</u>***

- Preattentive features, like color, size, shape, and motion, quickly catch our attention before we consciously recognize details.
- These composite visual elements are automatically identified within 10 ms, relying on differences in visual properties.
- They form the basis for effective graphic icons. At a conscious level, universally recognized geometric shapes convey complex concepts across cultures, seen in diagrams representing connection, dependency, causality, and more.






Examples of preattentive features for attention focus based on differences in size, shape, and orientation (left) and application to icon design (right).

| Diagram | Meaning |
|---|---|
|  | Relation/Path between Objects |
|  | Types/Strength of Relations |
|  | Inclusion, Structure, Hierarchy |
|  | Order, Causality, Flow, Process, Dependency |

Examples of diagrams/shapes/objects/figures with universal semantics.

# Aural

The actual form of sound feedback can be roughly divided into three types:

(a) simple beep like sounds

(b) short symbolic sound bytes known as earcons

(c) relatively longer “as is” sound feedback that is replayed from recordings or synthesis

Some important parameters of the human aurul capacity and the corresponding aural display parameters are discussed further.

***Aural Display Parameters:***

- **Intensity/Volume:** Refers to the amount of sound energy. 0 dB corresponds to the lowest level of audible sound and about 130 dB is the highest.

- **Sound:** Can be viewed as composed of a number of sinusoidal waves with different frequencies and corresponding amplitudes. The dominant frequency components determine various characteristics of sounds such as the pitch ( low or high key), timbre (which instrument), and even directionality (where is the sound coming from). Humans can hear sound waves with frequency values between about 20 and 20,000 Hz

- **Phase:** Refers to the time differences among sound waves that comes from the same source. Phase differences occur as our left and right ears may have slightly different distances to the sound source and. Phase differences are also known to contribute to the perception of spatialized sound such as stereo

| INTENSITY (DB) | DESCRIPTION |
|---|---|
| 0 | Weakest sound audible |
| 30 | Whisper |
| 50 | Office environment |
| 60 | Normal conversation |
| 110 | Rock band |
| 130 | Pain threshold |

Examples of different sounds and their typical intensity levels in decibels.

General recommendations for aural feedback is sound signal should be between 50 and 5000 Hz, composed of at least 4 prominent harmonic frequency components ( frequencies that are integer multiples of one another), each within the range of 1000–4000 Hz .

***<u>Other Characteristics of Sound as Interaction Feedback:</u>***

- As sound is omnidirectional, it is most often used to attract and direct a user’s attention. Also, it can be a nuisance as a task interrupter.
- The auditory feedback to be heard effectively, there should be a difference of 15–30 decibels between the volume of the feedback sound and the background noise/sound.
- Continuous sounds are more prone to habituation. and becomes less noticeable or even ignored [ like elevator music ].
- Only one aural aspect can be interpreted at a time.

***Aural Modality as Input Method:***

- The aural modality is not only used for passive feedback but can also serve as an active input method for interactive systems. Two major methods for aural input: Keyword recognition and natural language understanding.
- Technology for recognizing isolated words, especially for simple commands, has become robust.
- Challenges of voice input
  - Segmenting the relevant portion from continuous voice input or background noise.
  - Switching to voice-command mode can be a nuisance for users.
  - Voice input is more effective in situations where hands are occupied or there's minimal background noise.
  - Understanding long sentences and natural-language commands is computationally challenging.

## Tactile and Haptic

**Haptic:** Refer to the modality for sensing force and kinesthetic feedback through our joints and muscles (even though any force feedback practically requires contact through the skin)

**Tactile:** For sensing different types of touch (e.g., texture, light pressure/contact, pain, vibration, and even temperature) through our skin.

**Tactile Display Parameters:**

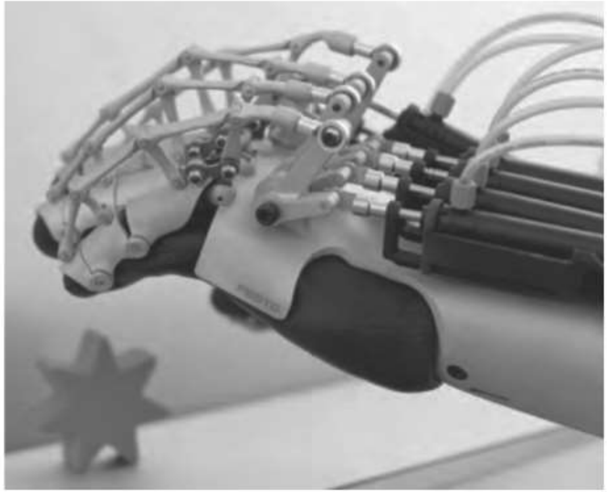
- Tactile resolution: The skin sensitivity to physical objects. As fingertip being one of the most sensitive area can sense objects as small as 40μm.
- Vibration frequency: Signal-response range is 100-300 Hz, Vibration frequency of about 250 Hz is said to be the optimal for comfortable perception.
- Pressure threshold: The lightest amount of pressure humans can sense is said to be about 1000 N/m2. For a fingertip, this amounts to about 0.02 N for the fingertip area .

Many type of tactile stimulation: texture, pressure, vibration and even temperature.

**Haptic and Haptic Display Parameters:**

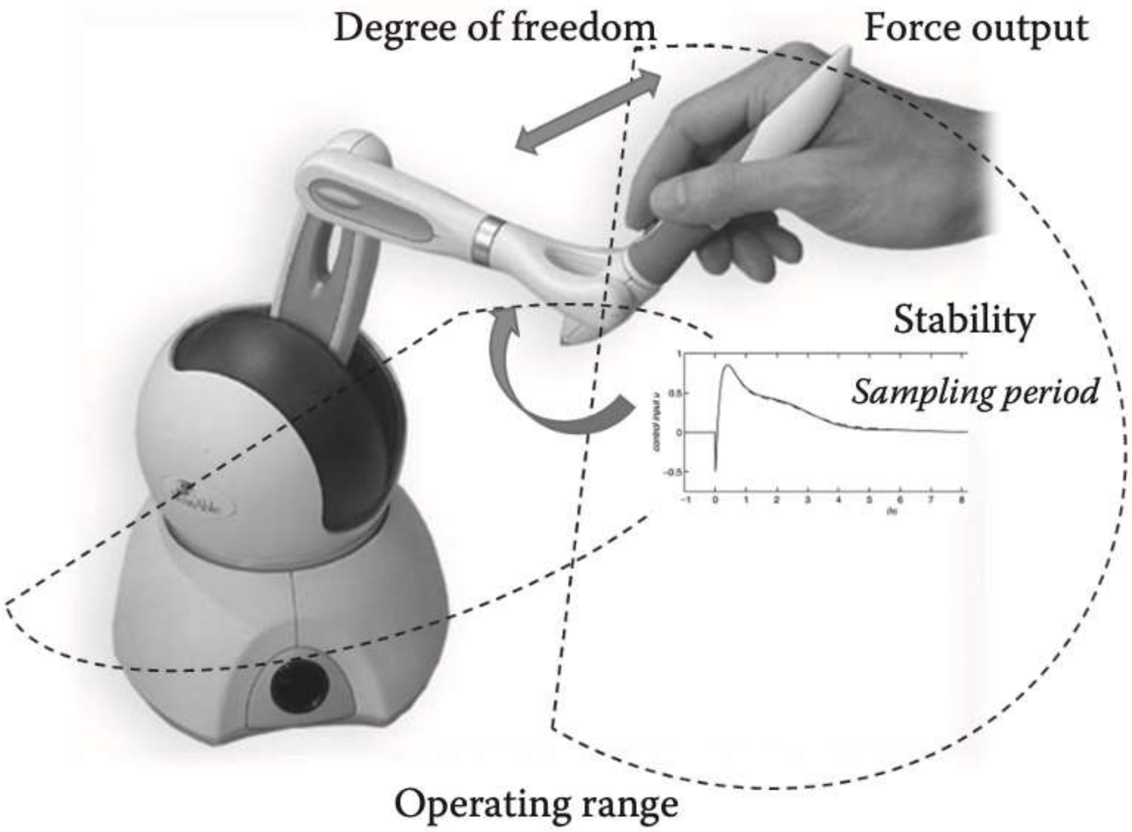
- The degrees of freedom: The number of directions in which force or torque be can displayed
- Force range: should be at least greater than 0.5mN
- Operating/Interaction range: How much movement is allowed through the device
- Stability: How stable the supplied force is felt to be

(a)



(b)

Two types of haptic systems: grounded and body worn




Important parameters for a haptic display system

**Multimodal Interaction**

By employing more than one modality, interfaces can become more effective in a number of ways, depending on how they are configured. Here are some representative examples.

- Complementary: Different modalities can assume different roles and act in a complementary fashion to achieve specific interaction objectives. For example, an aural feedback can signify the arrival of a phone call while the visual displays the caller’s name.
- Redundant: Different modality input methods or feedback can be used to ensure a reliable achievement of the interaction objective. For instance, the ring of a phone call can be simultaneously aural and tactile to strengthen the pick-up probability.
- Alternative: Providing users with alternative ways to interact gives people more choices. For instance, a phone call can be made either by touching a button or by speaking the caller's name, thereby promoting convenience and usability.

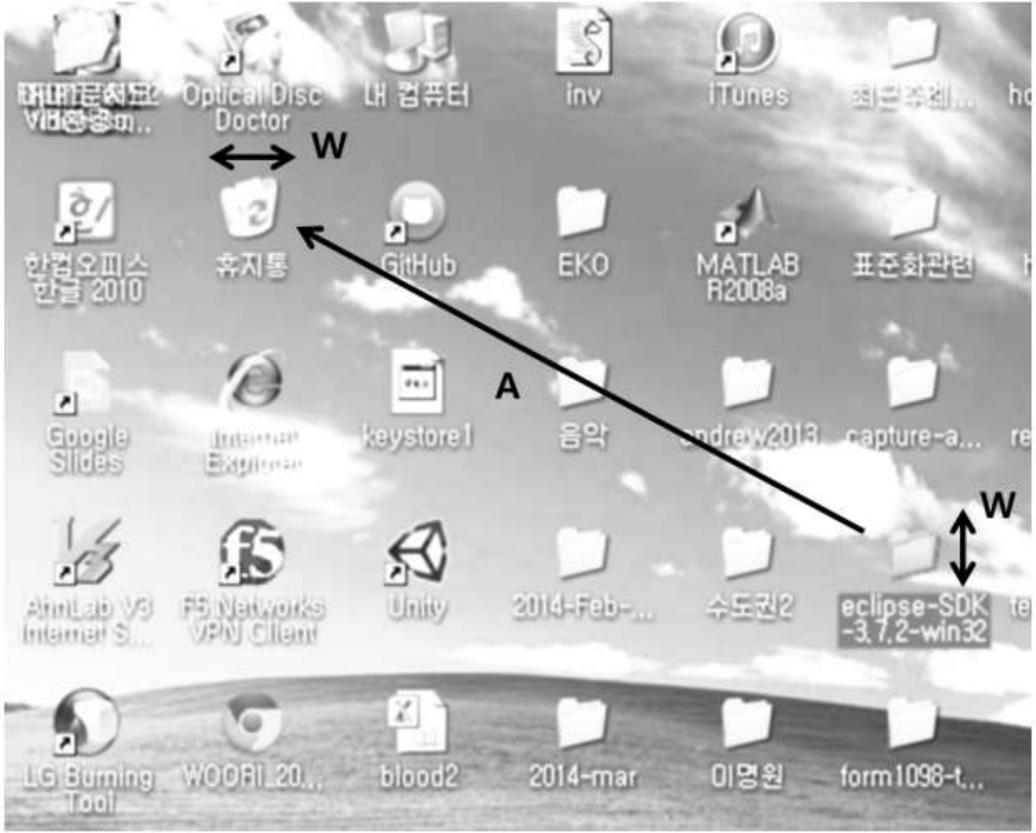
## Human Body Ergonomics ( Motor Capabilities )

**Fitt’s Law:** A model of human movement that predicts the time required to rapidly move to a target area as a function of the distance to and the size of the target.

**MT = a + b * ID and ID = log(A/W + 1)**

Where,
A and b are empirically determined constants
Mt is movement time
A is Distance
W is Size of target

*Targets as large as possible*
*Distances as small as possible*




Applying Fitts's law to a computer interface (dragging a file icon into the trashcan icon)

## Motor Control:

- A significant drop in human motor control performance below a certain spatial-resolution threshold. For instance, while the actual performance is dependent on the form factor of the device used and the mode of operation, the mouse is operable with a spatial resolution on the order of thousands of dpi (dots per inch) or ≈0.020 mm, while the resolution for a 3-D stylus in the hundreds.
- Control-Display (C/D) Ratio:
    - It refers to the ratio of movement in the control device (e.g., mouse) to the movement on the display (e.g., cursor).
    - A low C/D ratio means high sensitivity, resulting in fast travel time across the display. Conversely, a high C/D ratio means low sensitivity, leading to relatively faster fine-adjustment time.

## Others

There are many cognitive, perceptual, and ergonomic issues that have been left out. Due to the limited scope of this book, we only identify some of the issues for the reader to investigate further:

- Learning and adaptation
- Modalities other than the “big three” (visual/aural/haptic-tactile), such as gestures, facial expression, brain waves, physiological signals (electromyogram, heart rate, skin conductance), gaze, etc.
- Aesthetics and emotion
- Multitasking