LUNDS UNIVERSITET
Ekonomihögskolan

Institutionen för informatik

The value of visual depiction:

A comparison of short-term memory capacity between glyph icons and words

Kandidatuppsats 15 hp, kurs SYSK16 i Informatik.

Författare: Axel Görnebrand

Handledare: Umberto Fiaccadori

Rättande lärare: Björn Svensson
Paul Pierce

The value of visual depiction: A comparison of short-term memory capacity between glyph icons and words

ENGELSK TITEL: The value of visual depiction: A comparison of short-term memory capacity between glyph icons and words

FÖRFATTARE: Axel Görnebrand

UTGIVARE: Institutionen för informatik, Ekonomihögskolan, Lunds universitet

EXAMINATOR: Osama Mansour, PhD

FRAMLAGD: Maj, 2022

DOKUMENTTYP: Kandidatuppsats

ANTAL SIDOR: 52

NYCKELORD: short-term memory, glyph icons, UI

ABSTRACT:
Icons are widely used throughout society; moreover, they play a significant role in computer user interfaces, frequently appearing on websites and applications. In many cases, they are used instead of text labels to make a user interface (UI) more universally understood, among other things. Since icons reside within the UI, they should adhere to UI design principles, such as ”reduce short-term memory load”. This study aimed to research whether the short-term capacity of icons (more specifically: glyph icons) differs from words. A study was carried out with 30 participants of the same population, split into two equally sized groups. One group was asked to recall as many icons as possible. The other group carried out the same task but was instead asked to recall words. The results showed that the icon group could recall an average of 7.73 items, while the word group was able to recall an average of 4.40 items. Users' ability to store more icons in short-term memory than words could provide a basis for future decision-making in UI design, where one has to be chosen over the other.

Content

Figures

1 Introduction

1.1 Background

Icons are found all throughout society. They show us which bathroom to choose, what the weather will be like, if a car's washer fluid is low, if a stove is turned on, or whether a fence is electric or not. In the computer age, icons are primarily associated with the "user interface" (UI) (Hicks, 2011). Here, one finds everything from the simple outline of a heart, as in the like-icon on Instagram, to the very detailed image of a hawk on a post stamp, as in the mail icon of the Mac OS desktop.

Icon-like depictions date back to prehistoric times, as shown in the various cave drawings found throughout Europe. Among these, such as the Cave of Gargas or Pech Merle, one can see the silhouettes of hands surrounded by red ochre and black cinder, which have left them intact some 25 000 years later (Arm, 2010). These stencils of hands perhaps have no more significant meaning than to say: "I was here", as suggested by Arm (2010), yet convey this message in such a powerful and communicative way that it is hard to imagine it being done justice with mere words. This is most likely because of the rich symbolism associated with human hands. They "signify the sovereign, world creating reach of consciousness" while at the same time embodying "effectiveness, industry, adaptation, invention, self-expression and the possession of a will for creative and destructive ends" (Arm, 2010, p. 380). In large, hands have come to signify *man* as a whole (Alpenfels, 1955).

Figure 1: Hand in Pech Merle (Wikipedia, n.d.a)

The visual depiction of the human hand is in no way restricted to prehistoric times. It appears in Hieronymus Bosch's *The Garden of Earthly Delights* (Figure 2), severed and pierced through with a knife; as well as the main subject in one of Dorothea Lange's photographs of an Indonesian dancer (Figure 3). Moreover, the hand appears as an icon in stop signs (Figure 4), usually in a very simplistic, two-dimensional design, not unlike the hands found in the Pech Merle cave. In more recent times, we find the hand icon in the UI of the Windows and

Mac operating systems, first and foremost as a pointer and grab cursor (Figure 6). The purpose of this is to suggest an interaction with the interface, drawing a parallel to the way the user would interact with something in the "real" world, such as grabbing a cup or pointing to a location. On Youtube, as well as Facebook, buttons with a thumbs-up-icon and thumbs-down-icon are used to show either "like" or "dislike" of content (Figure 5).

Figure 2: Detail of Bosch's *The garden of earthly delights* (Bosch,1503-1504)

Figure 3: Lange's photograph of a hand (Lange, 1958)

Figure 4: Stop sign (Myloview, n.d.)

Figure 5: Facebook like icon (Wikimedia, n.d.a)

Figure 6: Mac grab icon (Cursor, n.d.)

Hicks (2011) suggests that icons in the UI can overcome language barriers, offer support for navigation, provide feedback, show status, and help convey mood and emotion. Icons, he continues, are sometimes used as a replacement for text, while in other cases, they are paired with text to emphasize their meaning. The decision between the former and latter alternatives can be based on the available amount of space in the interface or perhaps an artistic vision (Hicks, 2011).

Besides the benefits suggested by Hicks (2011), what other advantages might icons have in contrast to words? Previous research within psychology suggests that pictures are more easily remembered than words, especially within short-term memory (Shepard, 1967) (Pellegrino et al., 1976) (Reder et al., 2006). This superiority of pictures may also apply to icons, although no previous scientific study has bridged this gap yet.

According to Shneiderman (2004), memory is an important issue to consider when designing a UI. He has provided "Eight golden rules of interface design", the last of which is: "Reduce short-term memory load". This last rule is echoed in one of the usability heuristics proposed by Nielsen (1994), which states: "Reduce the information that users have to remember". Furthermore, one of Mandel's (1997) golden rules is: "Reduce users' memory load". According to Shneiderman (2004), reducing the user's memory load makes the user interface easier to use, leading to a more enjoyable experience. But, of course, the same could be said for basically any product: no one would ever buy a vacuum cleaner that requires the user to press ten buttons in a particular order, rather than just one, to start.

Whether it be a hand or cogwheel, icons yield a communicative power different from text. Depending on the circumstances, one probably serves a more important role in functionality than the other. This paper explores how short-term memory capacity between icons and words compares. Since the reduction of the user's memory load strengthens the functionality of the UI, as suggested by Shneiderman (2004), Nielsen (1994), and Mandel (1997), a difference in short-term memory capacity between icons and words may provide a further basis in future decision-making, where one has to be chosen over the other.

1.2 Research question

This paper aims to answer the following question: how do glyph icons compare to words in terms of short-term memory capacity?

1.3 Purpose

The study explores whether short-term memory capacity significantly differs between words and glyph icons. The result of the study may assist in future UI design decision-making since the reduction of user memory load enables a more successful UI. However, one should note that when “reduction of user memory load” is traditionally discussed in the context of UI design, it usually concerns situations where the user may be forced to remember some particular information, such as a number sequence from a previous webpage that has to be repeated. The memory capacity of words versus icons is something very different since they are not usually required to be actively remembered within their UI context. Nevertheless, situations may very well arise where icons or text labels (words) have to be recalled. Here is an example:

> An employee is sitting with a new photo editing tool, navigating through the menu, when his boss suddenly grabs his attention. His boss asks him if it is possible for him to erase a part of a photograph, cut out a part from another photograph, add the cut-out part to the first photograph, and then add a text label over it. Let's say the menu only consists of icons. When talking to his boss, the employee is able to recall four relevant icons: an eraser, a scissor, a plus, and a capitalized T. He is able to define the meaning of these icons and tells his boss that it should be possible. Now, let's say that the menu only consists of words and that the short-term memory capacity of words is 50% of icons. The employee is only able to recall two relevant menu items. Therefore, he is unsure whether all tasks can be performed.

Considering the above scenario, an icon has a reduced load on the user’s memory in comparison to words, which enables the user to store more icon items than word items. One could therefore argue that the icon interface is more effective and user-friendly, which is the main concern of Shneiderman (2004), Nielsen (1994), and Mandel (1997).

1.4 Delimitations

The study was limited to using two types of participant groups: one was exposed to stimuli in the form of icons, while the other one was exposed to stimuli in the form of words. The icons did not originate from a specific UI but were instead selected from the PowerPoint archive. They were judged as mainly being uncommon in terms of usage; therefore, common icons like a cogwheel, thumbs up, heart, etc. were not included. Since *all participants were Swedish*, the words used were *Swedish nouns,* and they corresponded to the icons in terms of meaning.

In both groups, the stimuli, which consisted of 20 items, were presented *simultaneously* in a matrix for a total of 30 seconds. It is not uncommon in STM tests for stimuli to be presented one by one in a flashing instant; however, this was not the case in this study. After the presentation of stimuli, there was a *20-second delay*, consisting of an *auditory distraction.* The distraction meant participants were asked to count backward in steps of three. Following the distraction, the participants were asked to recall as many items as they could remember. They did not have to recall the position of the items.

2 Theory

2.1 Icons

2.1.1 Etymology

According to Lalla (n.d.), “Icon” derives its origin from the Greek word “eikon”, meaning “image”, “likeness”, or “representation”. Its usage, he continues, begins within Ancient Greece, frequently appearing in Platonic texts, such as *The Republic*, where it is used to describe those flickering images upon the cave wall in his famous allegory. This visual-centered understanding of the term persists in its subsequent and contemporary use. (Lalla, n.d.).

Within the context of Eastern Christianities, Lalla (n.d.) writes, the icon is an image, most commonly a painting, that portrays Christ, Mary, Saints, or other religious figures. The work, he writes, is supposed to facilitate a connection between the viewer and the personage represented. This came to be due to Christianity's contact with the Roman Empire in the 5th and 6th centuries, where statues and likenesses of the Roman Emperor figured as stand-ins for his presence (Lalla, n.d.). Drawing on Roman statues, the religious images came to be utilized as sacred tools, being used as objects for prayer and seen as having a divine presence in themselves (Lalla, n.d.).

Figure 7: Icon by Andrei Rublev (Rublev, 1425)

In the field of semiotics, Lalla (n.d.) writes, *icon's* use originates in the works of Charles Peirce where it is one of three classifications of a “sign”, the other two being “symbol” and “index”. The sign is a relational way of being, he continues, and the icon is the classification that shares the most common ground, in terms of features, with its depiction of the signified. On the other side of the spectrum, you have the symbol, which shares no resemblance to its signified object. The index is somewhere in between these two spectrums (Lalla, n.d.).

Icon's most recent usage (and the one most relevant to this paper) is within computing (Lalla, n.d.). In the UI, the icon is an image whose purpose is to help the user navigate a computer system (Hicks, 2011). The term was coined in a Ph.D. thesis by computer science graduate David Canfield Smith (Canfield Smith, 1975). Its use draws on both its religious and semiotic definition, seeing as the computer icon is a small image, sometimes pictographic, which facilitates access to a much larger computational entity (Lalla, n.d.).

2.1.2 History

Before cultures developed any form of written language, icon-like symbols served as a means of communication, which is evident in the cave drawings of Pech Merle and Lascaux (Hicks, 2011).

According to Robinson (1995), many ancient writing systems have used graphic symbols as part of their alphabet and writing. Perhaps the most famous of these is Ancient Egyptian, whose hieroglyphs at first glance may seem completely pictographic, yet are far more complicated than that. For example, he writes that the hieroglyph for water is, in fact, a pictographic image of water; however, this symbol also serves as the letter *n*. Therefore the word "sn" (vocals are not written in Ancient Egyptian), which means "brother", contains the water-hieroglyph. In addition, he continues, while a hieroglyph may appear in a context unrelated to what it depicts, the Ancient Egyptian writing system sometimes uses determinatives, which are logograms (in some cases pictograms) added to the end of words to indicate their meaning. An example of this is the ancient Egyptian word for the Nile River: "itrw". The word contains no n's yet ends with three water signs (plural, meaning "waters") in its hieroglyphic form (Robinson, 1995).

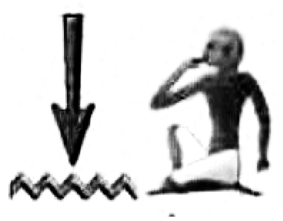


Figure 8: "Brother" in Ancient Egyptian (Pinterest, n.d.a)

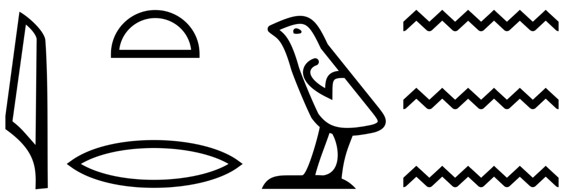


Figure 9: "Nile River" in Ancient Egyptian (Wikimedia, n.d.b)

In more contemporary languages, such as Chinese, Robinson (1995) continues, the pictographic origin of its characters is still visible. One of the most evident of these is the character for "mountain", whose modern three vertical lines have developed from initially being three mountain peaks in its ancient precursor. Similarly, the drawing of a sun (a circle with a line in it) has developed into a standing rectangle, with its line now going from one side to the other (Robinson, 1995).

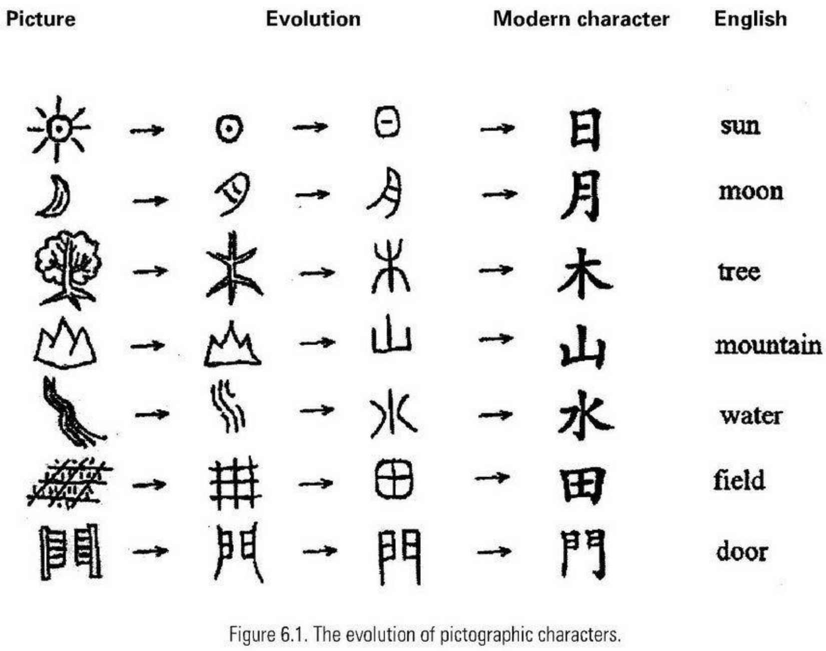


Figure 10: Evolution of Chinese characters (Reddit, n.d.)

Horton (1994) writes that, besides serving as a precursor to modern writing, graphic symbols also appeared during the Middle Ages, where esoteric signs were developed by practitioners of alchemy, astrology, and masonry. Many of these have a modern icon-like simplicity to them, he concludes.

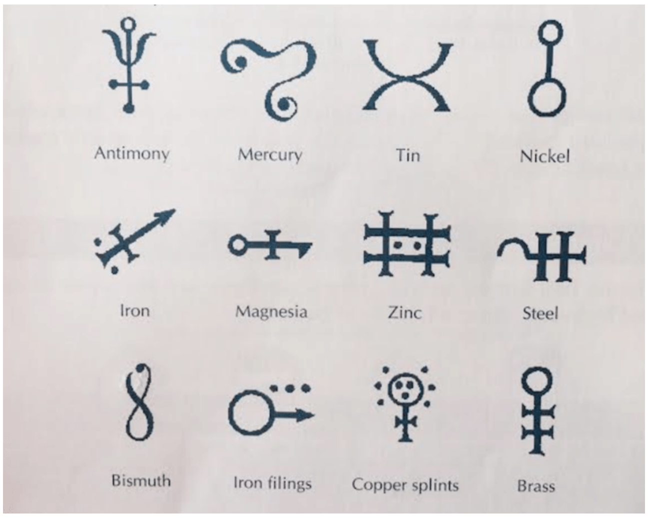


Figure 11: Esoteric signs (Horton, 1994)

Furthermore, from the 1880s until the 1940s, hobos who traveled across the United States consistently carved symbols on benches, street signs, and fences to convey messages to other hobos who would follow the same trail (Hicks, 2011). These signs would provide vital information regarding direction, safety, food, and shelter (Hicks, 2011).

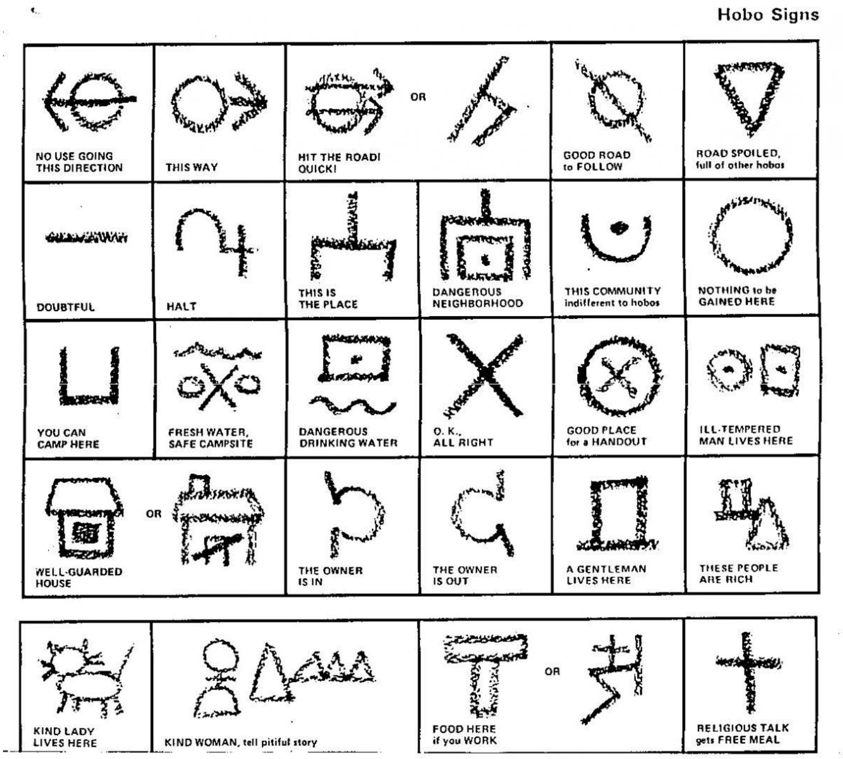


Figure 12: Hobo signs (We find wilderness, n.d.)

Up until this point, icon-like symbols were primarily used in localized efforts, serving as a means of communication within a culture and not outside it, Hicks (2011) writes. This, he continues, changed with the groundbreaking work of Austrian teacher and social scientist Otto Neurath (1882-1945), who developed the picture language “Isotype” (International System of TYpographic Picture Education). The primary purpose of this language was to make complex statistical information universally understandable by using simplified pictographs instead of words (Horton, 1994). Neurath saw picture language as an additional language that, in some instances, fell short compared to words but could express a message more clearly and easily rememberable in other circumstances (Twyman, 1975). The work, designed by Gerd Arntz (1900-1988) and commissioned by Neurath, is a direct precedent of contemporary icons found throughout city infrastructure and the UI. To this day, they still look surprisingly modern (Twyman, 1975).

Figure 13: Gerd Arntz's work commissioned by Otto Neurath (Arntz, n.d.)

Even though icon-like symbols date back to prehistoric times, Hicks (2011) writes, icons are more commonly associated with their use in computer interfaces. The first computer icon appeared in 1973 on the Xerox Alto, the first GUI (Graphical User Interface) based computer (Hicks, 2011). While most of the interface was still text-based, it did contain a cursor and an icon-based tools window (Hicks, 2011). However, Hicks (2011) continues, this was a research computer intended for organizations such as universities and was therefore not available to the public. The first consumer GUI computer was the Xerox Star, released in 1981, which incorporated the now-familiar design concept in which the computer workbench is an office metaphor, thus having files, folders, a desktop, and a wastebasket (Hicks, 2011. The icon set was designed by Norm Cox and marked the first appearance of the document icon with the folded top-right corner (Hicks, 2011).

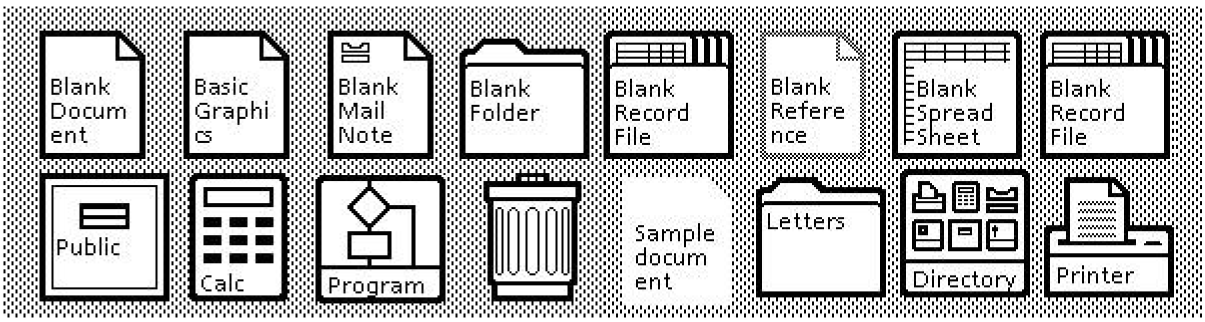


Figure 14: Xerox Star's icons (Stardock, 2020)

According to Stardock (2020), in 1983 Apple launched the Apple Lisa, named after Steve Jobs' daughter, featuring icons that follow the same type of aesthetic as the Xerox Star. The computer was the first of its kind to be aimed at a broad audience of business customers (Stardock, 2020). The Lisa was followed a year later by the Macintosh (Stardock, 2020). Here, Apple took icon design into the realm of an art form by hiring visual artist Susan Kare (Stardock, 2020). The result was a collection of icons that were visually appealing in their simplicity yet at the same precise in their communication (Stardock, 2020). At this point, Kare also designed the Apple Command icon, which is still used and widely recognized today (Hicks, 2011).

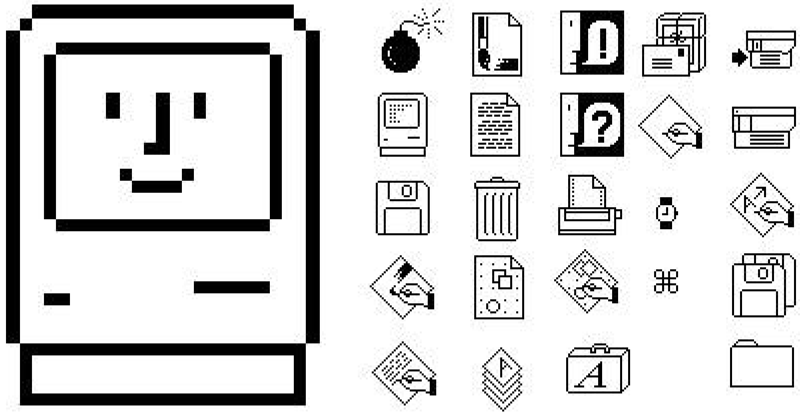


Figure 15: Apple Lisa's icons (Stardock, 2020)

In 1985, a year after the release of the Macintosh, Microsoft released its first computer with a GUI (Stardock, 2020. In comparison to the Macintosh, the icons seem a bit off in terms of consistency and grid (Stardock, 2020).

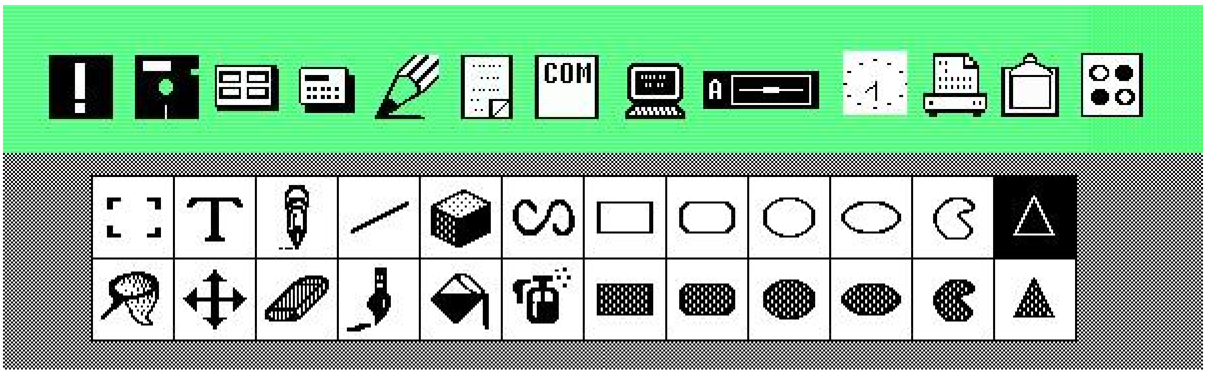


Figure 16: Windows' first icons (Stardock, 2020)

With the release of the Macintosh System 7 in 1991, Apple introduced a very soft color palette to their icons (Stardock, 2020). The icons also carry more visual depth, thus appearing more clickable (Stardock, 2020).

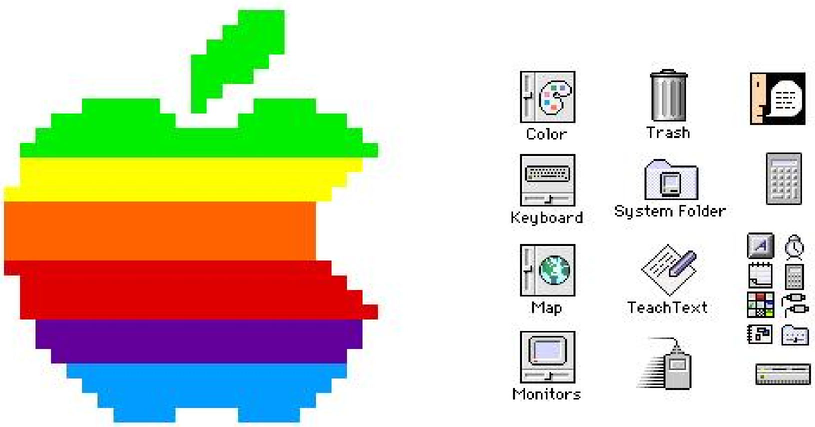


Figure 17: Macintosh System 7's icons (Stardock, 2020)

Windows 95 was released in 1995 and marked the debut of the Start button (Stardock, 2020). The icons also appear more refined, isometric, and realistic than in previous operating systems (Stardock, 2020).

Figure 18: Windows 95's icons (Stardock, 2020)

The appearance of icons took a massive leap in 2001 with the release of Mac OS X (Stardock, 2020). They appear more realistic than ever, showing complex reflections, highlights and textures (Stardock, 2020). In addition, they are rendered either from a straightforward or slightly above point of view (Stardock, 2020). From this point on, the style of the Macintosh GUI does not change significantly (Stardock, 2020).

Figure 19: Mac OS X's icons (Stardock, 2020)

Also, in 2001, Microsoft released Windows XP, whose icons are very saturated and rendered with a semi-transparent drop shadow (Stardock, 2020).

Figure 20: Windows XP's icons (Stardock, 2020)

With the release of Windows Vista in 2007, the icons took on a sort of "glassy" look, appearing less colorful and more transparent (Stardock, 2020).

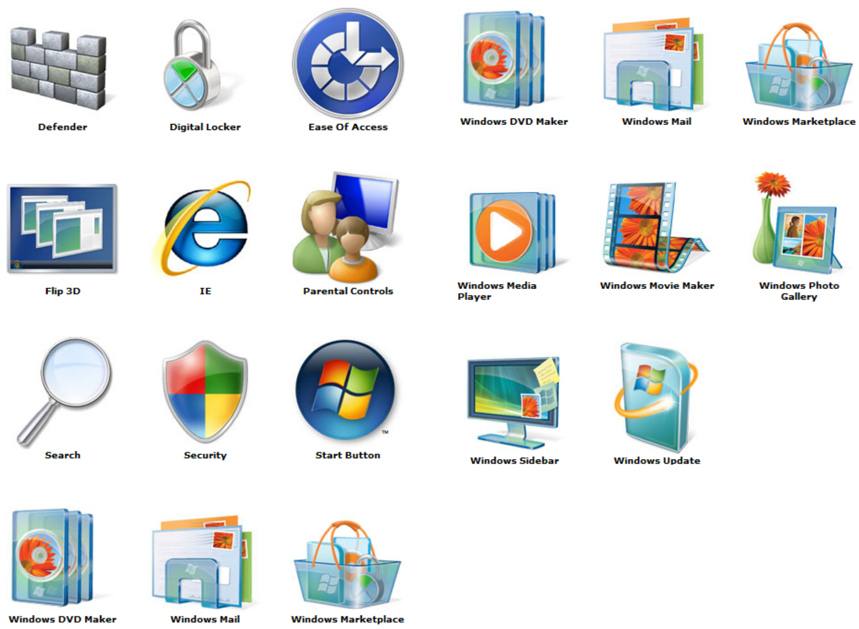


Figure 21: Windows Vistas' icons (Stardock, 2020)

In 2012, Microsoft released Windows 8, whose icons appear very different from their predecessors (Stardock, 2020). The glossy and reflective style found in Mac OS X and Vista is wholly abandoned and replaced by simplistic, concrete, and timeless glyph icons which appear in colorful squares (Stardock, 2020).

Figure 22: Windows Vistas’ icons (Stardock, 2020)

In the years following Windows 8, there was a significant increase in icons, especially glyph and outlined ones, appearing in GUIs. They no longer primarily appear in desktop interfaces but also on websites and mobile applications. One can take the Gmail.com interface as an example; look at a comparison between 2010 and 2022, especially the left panel, in terms of icon use:

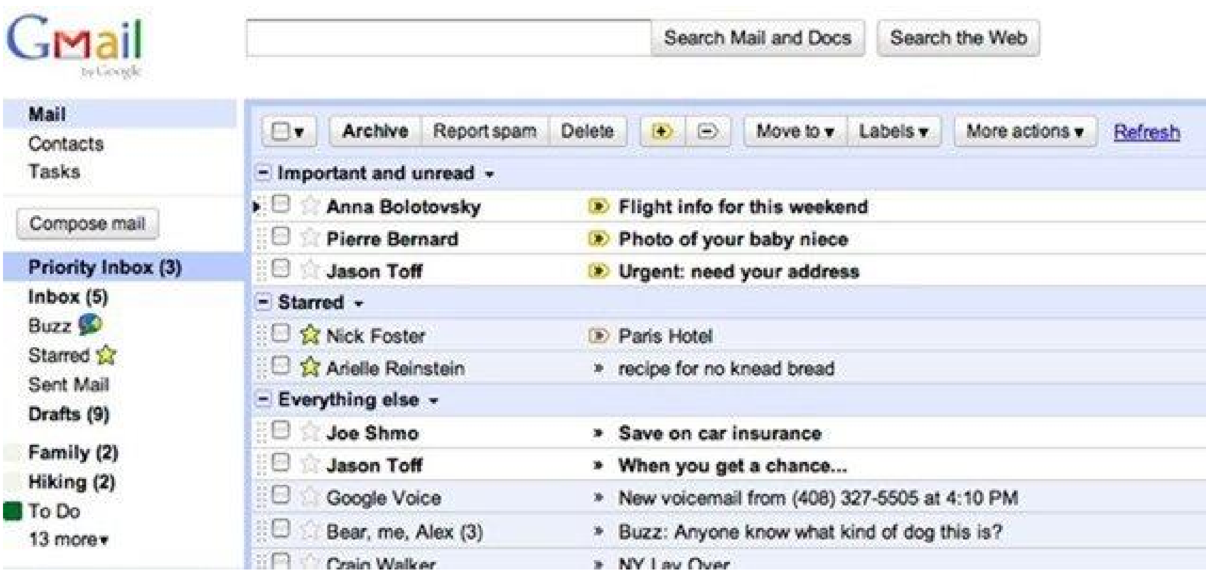


Figure 23: Gmail interface in 2010 (CNN, 2010)

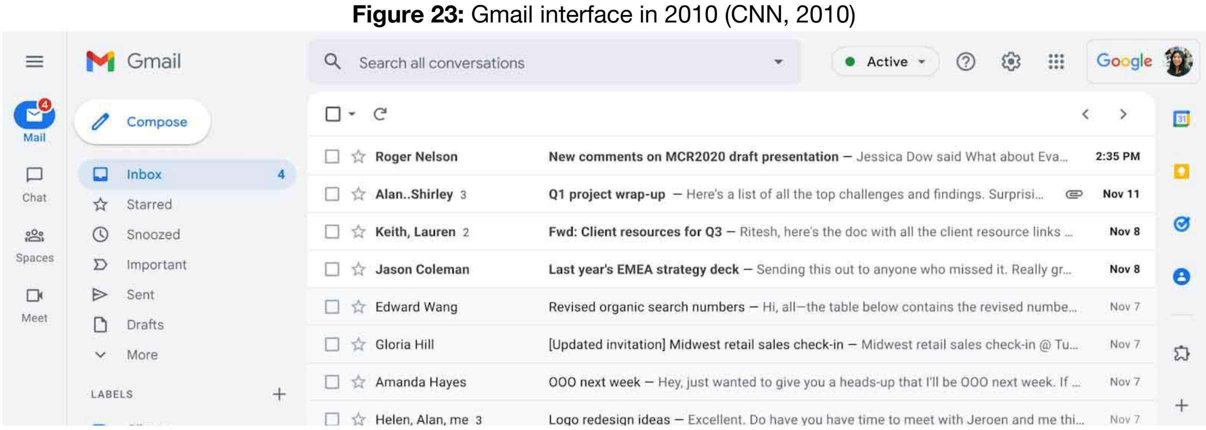


Figure 24: Gmail interface in 2022 (The verge, 2022)

It is hard to determine the exact reason why the increased usage of icons has happened so recently. Is it because the bandwidth has become a less significant issue when designing an interface? Or maybe because mobile interfaces require more effective solutions in terms of space? Most likely, there is more than one major factor. Either way, icons, especially glyph and outlined icons, are broadly used in today's most visited websites and applications, such as Facebook, Instagram, Youtube, and Reddit.

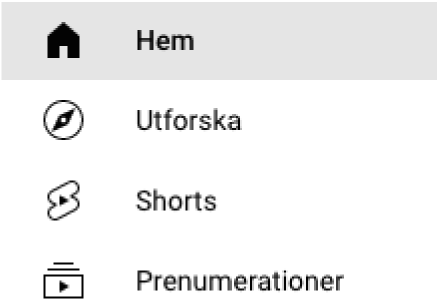


Figure 25: Youtube icons (Youtube, 2022)

Figure 26: Instagram icons (Instagram, 2022)

Figure 27: Facebook icons (Facebook, 2022)

Figure 28: Google maps icons (Google, 2022)

Figure 29: Reddit icons (Reddit, 2022)

2.1.3 Types of icons

There are several ways to categorize icons, and there does not seem to be any consensus on how it should be done. Gittins (1986) uses the classifications “Key” and “Associative”, which in themselves contain further classifications. An Associative icon, for example, can be either “literal” or “abstract”, and an abstract icon can be either “individual” or “metaphorical”.

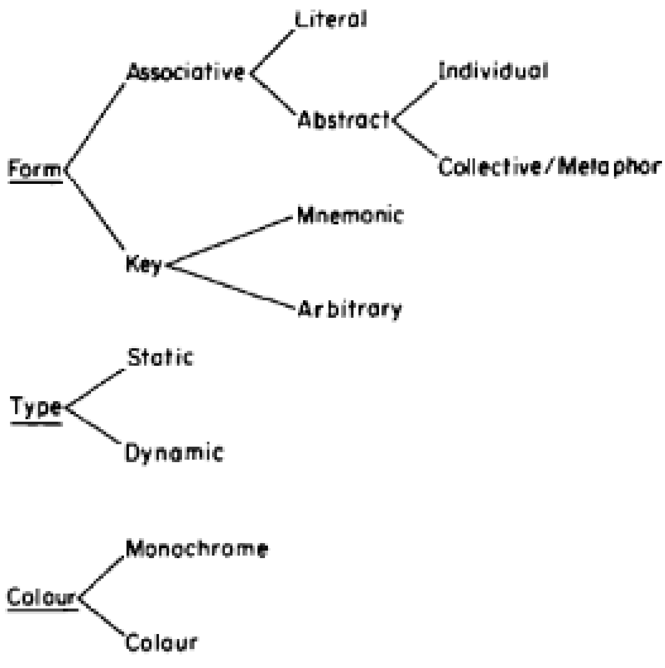


Figure 30: Icon types (Gittins, 1986)

Hicks (2011), on the other hand, categorize icons in the following way:

- **Favicons**. Smaller versions of a website logo are usually found in browsers’ address bars or bookmark views.

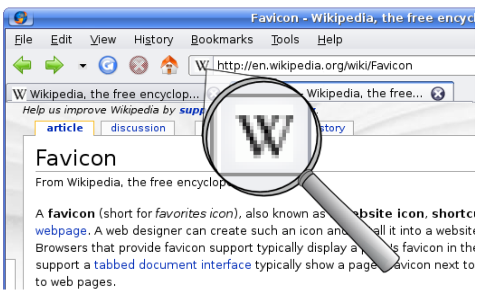


Figure 31: Wikipedia favicon (Wikipedia, n.d.b)

- **Ideograms**, **pictograms,** and **arbitrary icons**. These are the types of icons used in websites and user interfaces. A pictogram depicts something, and an ideogram is an idea of something. The arbitrary icon is an invention that has no real relationship to what it signifies.

Figure 32: Pictogram icon (Vecteezy, n.d.)

Figure 33: Ideogram icon (Max pixel, n.d.)

Figure 34: Arbitrary icon (Free icons library, n.d.)

- **Application icons**. According to Hicks (2011), these are usually more photorealistic.

Figure 35: Macintosh mail application icon (Mac rumors, n.d.)

Finally, Kamenez (2021) classifies icons into these types:

- **Glyph**. Usually completely or partially solid. Sometimes an outlined icon can become a glyph icon, once selected, as shown in the interfaces of Instagram and Youtube.

Figure 36: Glyph icon (Free icon shop, n.d.)

- **Colored**.

Figure 37: Colored icon (Flaticon, n.d.)

- **Duotone**. Contains two colors.

Figure 38: Duotone icon (Svgrepo, n.d.)

- **Outlined**. Created by vector strokes and empty inside.

Figure 39: Outlined icon (Stickpng, n.d.)

- **Universal**. Immediately recognized and usually represent actions such as home, search, or print.
- **Unique**. Represent unique functions or features. A logo icon would be an example.
- **Conflicting**. Icons that represent the same concept but are different. A "like" can be either a heart, thumbs up, or a star.

2.1.4 The use of icons

The icon's primary attribute, according to Lenaerts (2014), is to be a "common visual language" that transcends language barriers. Since software and the Internet is shared by a large number of people stemming from different cultures, he writes, the usage of icons becomes increasingly efficient in its digital context. Horton (1994) argues that icons give the

UI an international look, which reduces the need to translate every menu, button, or command each time a product, website, or application is intended for usage in another country. Furthermore, icons help users avoid reading in general, he concludes. Even though the rate of illiteracy in the world has significantly declined in the last century, 14% of the world's population still remains illiterate, which is more than one billion people (Roser & Ortiz-Ospina, 2018). Icons decrease the difficulty for this population to navigate in a UI (Horton 1994).

Figure 40: Airport icons (Pinterest, n.d.b)

Icons are frequently used in airports to help travelers find their way. This functionality also applies in a digital context. Hicks (2011) argues that navigation text is harder to scan than its iconic counterpart since there is little visual difference between text. Horton (1994) brings forth the same argument, writing that text labels, especially ones in all capital letters, all bear the same rectangular shape and are, therefore less visually distinct than icons, which becomes even more apparent when one looks at them squinting. Zhang (2020a) presents an excellent example of how icons help with navigation by showing a Windows 95 Start menu with and without icons:

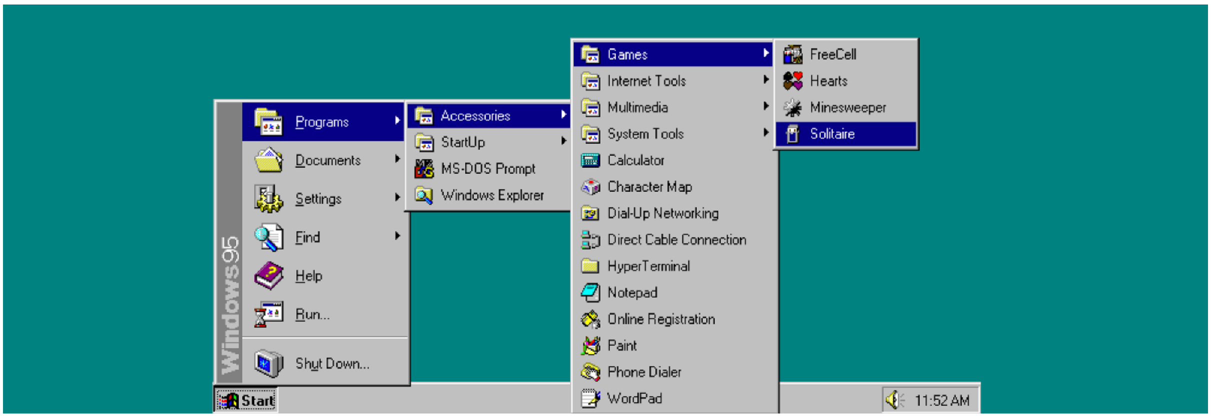


Figure 41: Menu with icons (Zhang, 2020a)

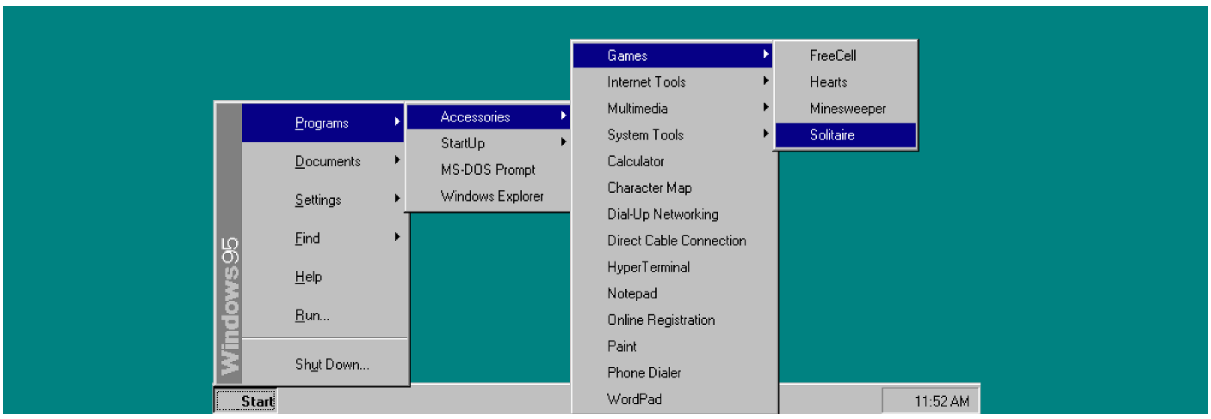


Figure 42: Menu without icons (Zhang, 2020a)

Icons are frequently used to show status, especially in video meeting applications like Microsoft Teams (Hicks, 2011). In this case, they are first and foremost used to show the availability of a user, but also to show whether they are muted or if their camera is being used. Furthermore, operative systems usually display a status bar, where icons are used to show battery and internet status, among other things (Zhang, 2020a) (Hicks, 2011).

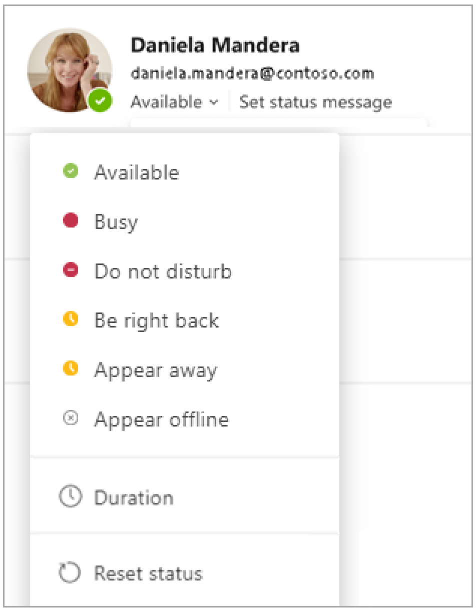


Figure 43: Microsoft Teams status icons (Microsoft, n.d.)

Another great use of icons, according to Hicks (2011), is to provide feedback. This, he writes, is usually done with a combination of colors. For example, when filling out a form, a user may be presented with a red cross if the username is already taken or if a password was not repeated correctly. However, as Hicks (2011) writes, it is best not to rely on color alone when conveying information since people with color blindness will have a hard time interpreting. Colors are also understood differently in other cultures: while red often means danger in Western society, it is instead related to luck in China (Hicks, 2011). With the use of icons, the correct interpretation can be reinforced, Hicks (2011) concludes.

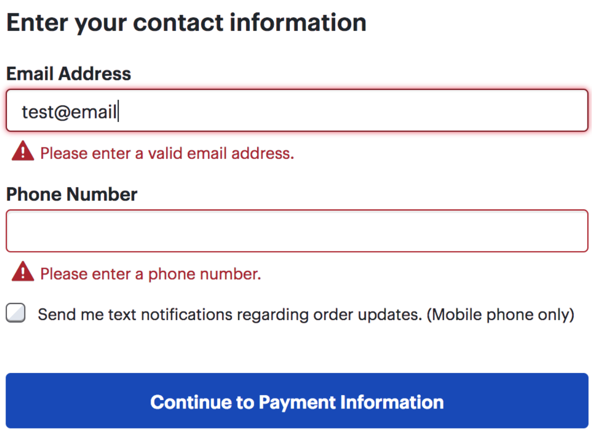


Figure 44: Feedback icons (Nielsen norman group, 2019)

Finally, icons can be used to save space, which is of great importance to mobile interfaces, which often need to be economical in terms of space (Hicks, 2011). Horton (1994) argues that concise icons can say more than their corresponding text labels or headings. Furthermore, he writes, a text is doomed to be in a wide rectangle shape, while an icon can be made into a compact square or another shape more appropriate to the UI. A look at different Bluetooth icons gives a glimpse of how written information can be condensed:

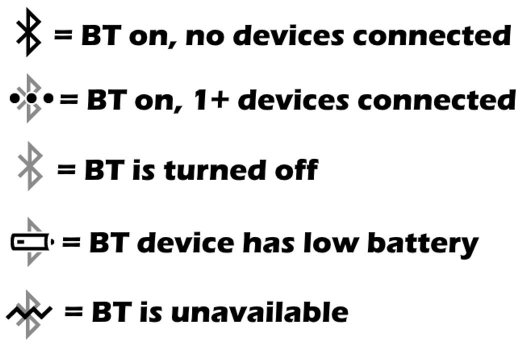


Figure 45: Bluetooth icons (Vira les locos, n.d.)

2.1.5 How icons should be designed

In terms of icon design, Zhang (2020b) has presented seven principles which are:

- **Clarity**. Since an icon's primary goal, she writes, is to convey a message clearly and quickly, clarity concerning its meaning is of utmost importance. With this in mind, diffuse metaphorical icons should generally be avoided. However, sometimes there is no standard image (like a cogwheel for "Settings") of what is supposed to be signified. In these contexts, the designer can successfully create a new type of icon, which may become a standard. A great example of this is the command icon created by Susan Kare, which had no prior image related to its use.

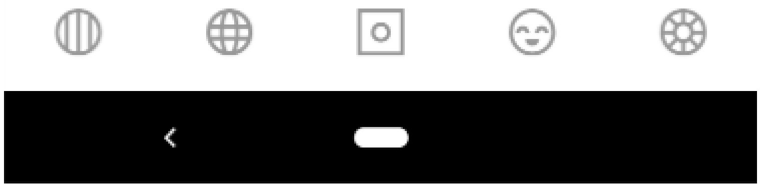


Figure 46: Hard to understand icons (Zhang, 2020b)

- **Readability**. Once the icon is understandable, it should also be easy to read. Zhang (2020b) uses the app Amtrak's icon for "Station" as an example of a hard-to-read icon since its details are difficult to perceive. Google Map icons, on the other hand, are highly readable.

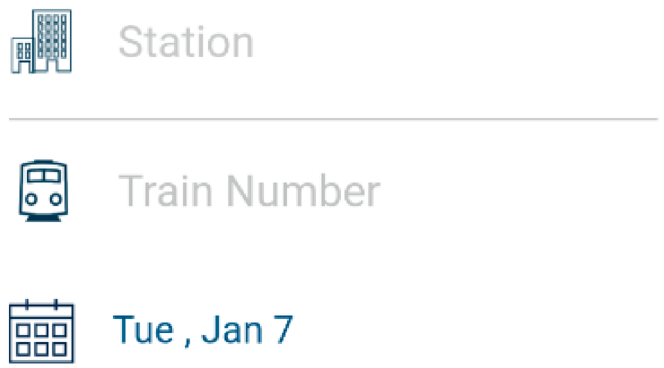


Figure 47: Amtrak icons (Zhang, 2020b)

Figure 48: Google Maps icons (Zhang, 2020b)

- **Alignment**. An icon should be aligned optically and not dependent on being metrically placed in the center. Even though the numbers might say an icon is correctly aligned, our perception can say otherwise. Zhang (2020b) uses the example of a play button; even though the triangle is metrically centered, we perceive it as unbalanced since one side of it bears more weight than the other. Just as typographers create optical illusions of balance in letters, iconographers should make similar adjustments.

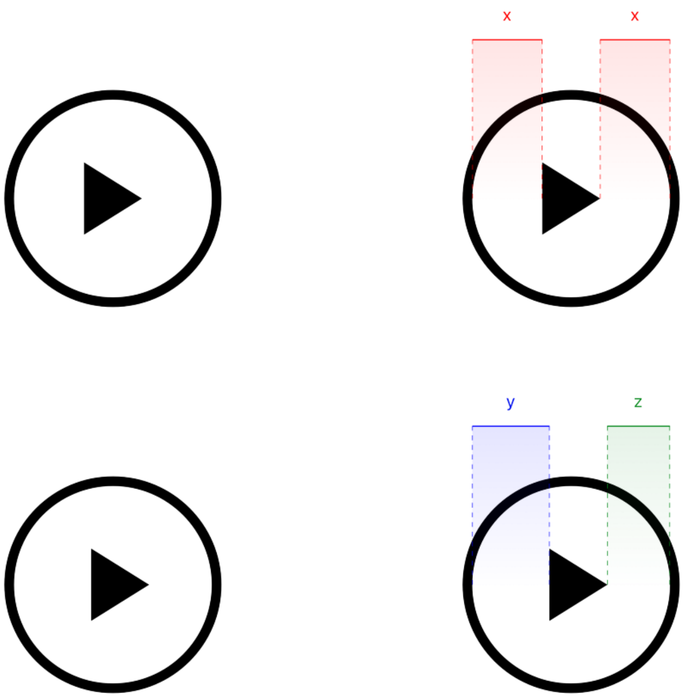


Figure 49: Alignment in icons (Zhang, 2020b)

- **Brevity**. Much like an idea expressed with just a few words seems more elegant than the same idea expressed with a couple of sentences, a simplified icon has more clarity and legibility. Besides, overly detailed icons entail more information that needs to be processed, thus requiring more cognitive effort (Goldstein & Van Hooff, 2018).

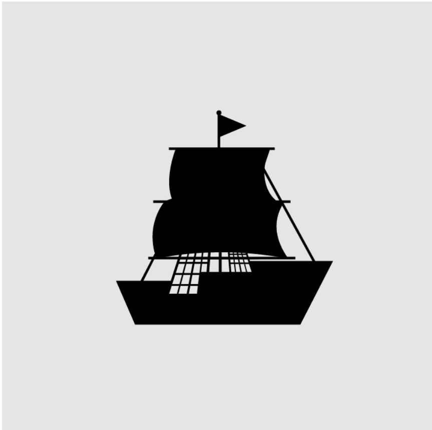


Figure 50: Overly complex icon (Material, n.d.)

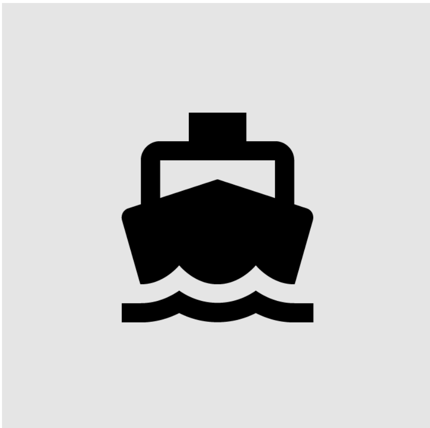


Figure 51: Simple icon (Material, n.d.)

- **Consistency**. Both when designing a family of icons and an icon in itself, there has to be consistency in terms of strokes, filling, and size. Without consistency, the user may start wasting time noticing design fallacies rather than completing a specific task.

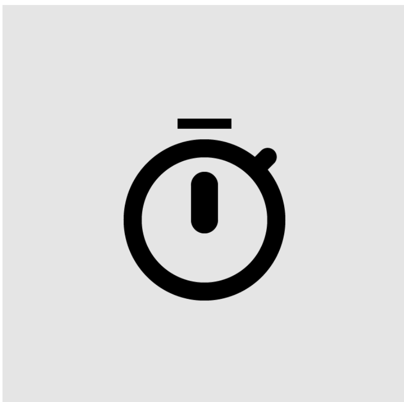


Figure 52: Inconsistent icon (Material, n.d.)

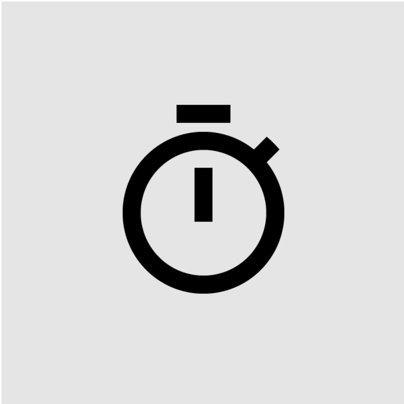


Figure 53: Consistent icon (Material, n.d.)

- **Personality**. When icons are a large part of a company's brand, their unique personality becomes increasingly important. Duolingo's unique, colorful icons are of great importance to their GUI by making the user feel more enthusiastic about learning new languages.

Figure 54: Duolingo icons (Design week, 2019)

- **Ease of use**. An icon set is not done, even when it has been drawn to perfection. Further preparation is needed to accommodate contributors making new icons, to help designers use them in their designs, and for engineers to code them. An icon set should be well organized in a master file, well-documented in terms of principles, and tested with regard to the other principles.

Hodge (2008) has presented another set of seven principles, and even though they do not differ considerably from Zhang's (2020b), a select few provide further insight into optimal icon design. Hodge's (2008) second principle: "Consider Your Audience", states that one has to take into account different cultural perspectives when designing icons, especially if the icons are meant for an international audience. For example, designing a mail icon that depicts a mailbox might be a bad idea since mailboxes look different in different parts of the world. Apple has avoided this problem by having a stamp, which has more cultural universality, as their application icon on MacBooks. Furthermore, Hodge's (2008) third principle: "Design for the Size the Icon will be Used At", states that each output size of an icon needs its own design. An icon that looks good in 576px might look like a blurry mess in 16px. Also, a detailed icon is usually best used in large sizes since it otherwise might be hard to read (as shown in the example of Amtrak's Station icon), while a simple glyph icon is more fitting as small and might appear visually underwhelming in a larger size.

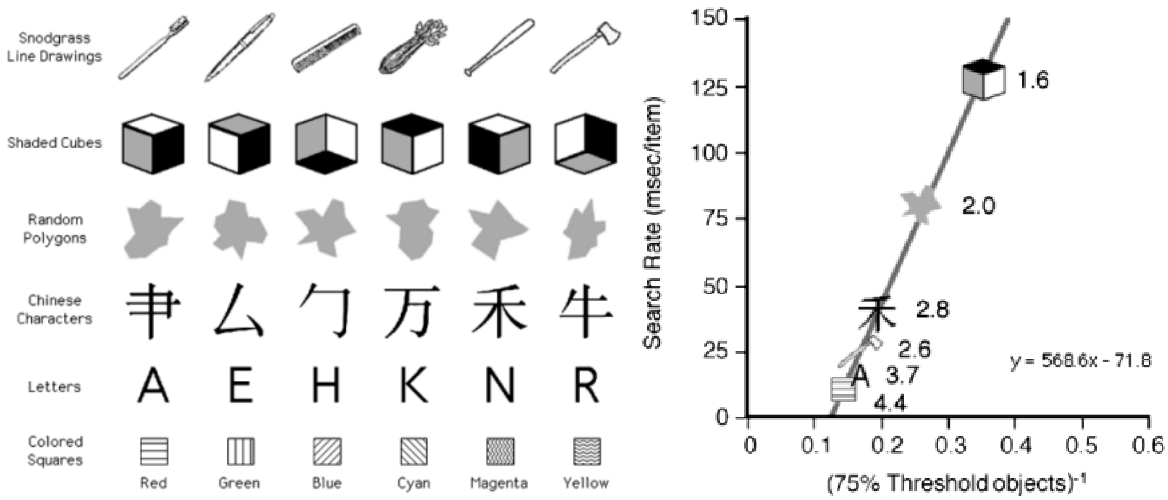
2.2 Short-term memory

Short-term memory (STM), also known as primary memory or active memory, is, according to Goldstein & Van Hooff (2018), the system involved in storing a small amount of information for a short amount of time. STM's brief duration makes it easy to downplay in comparison to long-term memory (LTM), they write, but STM is responsible for a great part of our mental life and can be considered "our window on the present". So whatever you, the reader, might be thinking right now or remember from what you have just read is information that is stored in your short-term memory (Goldstein & Van Hooff, 2018).

The lasting duration of short-term memory is considered to be between 15 to 30 seconds, which was demonstrated by Brown (1958) and Peterson & Peterson (1959) in an experiment on recall. Participants were given random letters paired with a random number, and after the instructor had read the letters and the following number, the participants were asked to repeat the number and begin counting backward in steps of three. For example, when the instructor said: "FOF 500", the participants would start counting: "500… 497… 494…" and would continue doing so until the instructor asked them to recall the original three letters. After counting for three seconds, participants were found to remember 80 percent of the letter groups but remembered only twelve percent after 18 seconds of counting. The results of this study were originally interpreted as evidence of memory decay following the passage of time. However, when Keppel & Underwood (1962) further analyzed the results, they found that participants' memory for the first trials was high, even when following a delay of 18 seconds. It was first, after multiple trials, that the result started to drop sharply. Keppel & Underwood (1962) concluded that the drop-off in memory was due to proactive interference (when previously learned information interferes with learning new information) rather than as a result of decay. In the context of everyday life, interference is constantly happening as one event follows the other. Therefore cognitive psychologists have concluded that the duration of STM, when rehearsal is prevented, is about 15 to 30 seconds (Goldstein & Van Hooff, 2018) (Atkinson & Shiffrin, 1971).

The estimation of how many items can be stored in the STM usually ranges between five to nine (Goldstein & Van Hooff, 2018). This idea was first introduced by Miller (1956) in his paper "The Magical Number Seven, plus or minus two". In this context, "Items" refer to a "meaningful unit or "chunk", which in itself can contain several items. For example, the letter series PFCBCSTM can be more easily remembered if it is split into several abbreviations such as PFC (Prefrontal Cortex), BC (Before Christ), and STM (Short-term memory). In conclusion, a person might be able to recall a sequence of six to nine words, but if the words were rearranged to form meaningful sentences, the capacity could increase to 20 words or more (Goldstein & Van Hooff, 2018).

The idea that the capacity of STM could be measured in terms of a number of items has been contested in later research, writes Goldstein & Van Hooff (2018). Some researchers, they continue, have suggested that STM capacity should instead be looked at in terms of "amount of information", rather than "amount of items". Concerning visual objects, they continue; the amount of information has been defined as the visual features or details of the object in question. The reasoning behind this is apparent when you take into consideration how a USB stick works: there is no limit in terms of how many pictures you can store on it; instead, it depends on the resolution of the pictures (Goldstein & Van Hooff, 2018). With this in mind, Alvarez & Cavanagh (2004) conducted an experiment where participants were shown two pictures consisting of objects of varying complexity that flashed one after the other. Afterward, the participants were asked to identify whether any change had occurred in the second picture. For example, shaded cubes, which were the most complex objects, would be displayed randomly across a white background, followed by a blank interval, and then followed by a picture that was either the same as the first one or in which one of the cubes was different. The result showed, as indicated by Figure 54, that increased complexity of items resulted in fewer items that could be stored in the STM.



Figures 55 & 56 Material and result (Alvarez & Cavanagh, 2004)

2.2.1 Pictures vs. Words

Since this study compares STM capacity between words and their corresponding icons (images), we will now take a look at prior experiments concerning similar comparisons.

Shepard (1967) conducted a series of experiments that compared memory recognition between words, sentences, and pictures. In the first experiment, a selection of 540 common English nouns and adjectives were exposed to the participants. Immediately after, the participants were shown a series of 60-word pairs, each containing a new unseen stimulus, and were asked to identify which of the words were "old". In the second experiment, participants were instead asked to identify the old stimulus out of a pair of short sentences. In this case, 612 sentences were followed by 68 pairs. The last experiment followed the same procedure as the two previous ones and consisted of 612 pictures (including colored prints, photographs, and other illustrations) and 68 pairs. Depending on whether the stimuli population consisted of words, sentences, or pictures, median participants were able to recognize the old stimulus in 90, 88, or 98% of the test pairs, respectively. In terms of pictures, participants best recognized the difference between old and new when the stimuli were colored and meaningful.

Four years after Shepard's (1967) experiments, Paivio and Csapo (1971) looked to explore the difference in short-term *sequential* memory (immediate memory span) between pictures and words. They conducted a series of experiments in which the presentation rate was either slow or fast (5.3 or 2 items per second), and the stimuli were either pictures, concrete words, or abstract words. The two classes of words were almost the same in terms of meaningfulness and frequency; however, concrete words were judged to be more image-conjuring than abstract words. The nine pictures were single-line black ink drawings of the objects that were labeled with the concrete words. In the case of the fast rate of stimuli exposure, words were still able to be read, yet pictures were presented so fast that they were unable to be implicitly labeled; therefore, their memory of them was only kept in a nonverbal storage system. After the participants were presented with the stimuli on a filmstrip projector, they had 30 seconds to put the items (printed on blocks) in the same order as they appeared on the screen. The results reported that performance in all three categories of stimuli was generally poorer at a fast rate. However, the picture reconstruction score was inferior to that of words at a fast rate, with abstract words having the highest score, followed by concrete words. At a slow rate, the reconstruction score was insignificant between the three categories. This result was ascribed to the dual encoding theory, which suggests that visual and verbal stimuli use different storage systems for memory.

In 1976, Pellegrino and colleagues reported a series of experiments in which short-term retention between pictures and words was compared following different delays (three, ten, or 30 seconds) and different types of distractions (acoustic vs. acoustic + visual). All participants were presented with a 24-trial sequence of displays consisting of 24-word or picture triads. The pictures consisted of 72 black line drawings of familiar objects and animals. Each trial consisted of a two-second display of the visual stimuli and was followed by either a three, ten, or 30-second delay of the particular distraction task. After each trial, the participants were asked to recall each item and its position within the triad orally. The acoustic distraction task comprised backward counting in steps of seven from a three-digit number. The acoustic + visual distraction task involved backward counting while simultaneously working on a modified Hidden Figures Test, in which the participant tries to locate a certain shape in a matrix of random lines. The recall was scored under two different criteria: stringent and lenient. Under the stringent scoring criterion, items were only considered correct if recalled and assigned their correct position. The lenient scoring criterion disregarded the positions of the items, and they were considered valid if recalled at all. The results of the stringent scoring system showed that pictures were significantly better recalled under acoustic distraction (.63 vs. .31), while there was no significant difference under dual distraction (.48 vs. .51). The results of the lenient scoring system showed that pictures were significantly better recalled under acoustic distraction (.67 vs. .41) while the opposite was the case under dual distraction (.51 vs. .61). Without exception, recall declined over the delay intervals and did not interact with the two factors. Pellegrino et al. (1976), like Paivio and Csapo (1971), attributed the results to the dual encoding theory. With the acoustic distraction task, semantic recall suffers more than visual recall, while in the dual distraction task, the opposite occurs, perhaps because the Hidden Figure Test is the most demanding of the two distractions (Pellegrino et al., 1976).

Reder et al. (2006) reported a double-blind study, in which participants studied a list of stimuli after receiving either midazolam, a drug that creates temporary anterograde amnesia, or saline (placebo). The list consisted of three types of stimuli: words, photographs, and abstract pictures. Approximately two minutes after receiving an injection, the participants were presented with the three types of stimuli in random order. The items remained on a computer screen for one second each, after which the participant rated their pleasantness. This part lasted between twelve and 15 minutes. After being screened the list of items, the participants were given a working memory task that lasted 20 minutes, after which they were immediately given a recognition test for the earlier judged items. The result demonstrated that midazolam hurt performance for word recognition to a higher degree than for photographs while not affecting abstract picture recognition. Reder et al. (2006) attribute this to the fact that midazolam affects contextual association and episodic traces, which is more relevant to word recognition since words are commonly more familiar than photographs and more so than abstract pictures. The result also showed that participants who received placebo injections had a significantly better recollection of words than photographs and abstract pictures. This, according to Oates & Reder (2011), goes against the popular belief that images are more easily remembered than words.

In conclusion, although the general opinion is that humans remember visual depictions better than words, as stated by Grady et al. (1998), the above studies give a more complex picture of the issue. Although Shepard's (1967) study shows that pictures were more easily recognized than words, the same thing can not be said of Paivio and Csapo's (1971) study, which showed a negligible difference between the two categories at a slow rate and a better recollection score for words at a fast rate. Furthermore, Pellegrino et al. (1976) showed that the type of distraction that follows an experiment affects short-term retention; while an acoustic distraction may interfere more with word retention, a visual distraction may interfere more with picture retention. Lastly, the study of Reder et al. (2006) showed that memory for words, rather than pictures, was more affected by an amnesia-inducing drug. However, when participants were under no influence of a drug, words were more easily recollected than pictures.

2.3 UI design

The UI is the space where human-computer interaction occurs ("UI design", n.d.). In a digital context, this includes the desktop of a computer, a website, or an application window (Indeed Editorial Team, 2022). The role of the UI is to meet user expectations and support the functionality of the program, app, or machine in question (Indeed Editorial Team, 2022). This is achieved through elements such as pull-down menus, buttons, icons, scroll bars, text fields, and containers (Indeed Editorial Team, 2022).

UI is sometimes used interchangeably with user experience (UX), but they are quite different (Indeed Editorial Team, 2022). Both parts play an essential role in the execution of a project and can not survive without the other (Indeed Editorial Team, 2022). However, UI is primarily concerned with the look and feel of the interface in question, while UX spans the entire process of conceptualization, development, and delivery (Indeed Editorial Team, 2022). Lamprecht (2022) makes the following analogy to explain the difference:

> If you imagine a product as the human body, the bones represent the code that give it structure. The organs represent the UX design: measuring and optimizing against input for supporting life functions. And UI design represents the cosmetics of the body, its presentation, its senses and reactions.

2.3.1 Summarizing UI design

Defining (let alone summarizing) UI design is no easy task since the subject is broad, and the angles of approach can vary significantly. Kreimer (2021) defines UI design as "building interfaces with a focus on styling and interactivity", while UX Design Institute (2022) defines it as "the process of designing the look, feel, and interactivity of a digital product."

According to Lamprecht (2022), the goal of UI design is to make the UI as accessible, attractive, guiding, and responsive for the user as possible. Mandel (1997) views UI design as a tool to help users build self-assurance about how they work with computers. Furthermore, he writes, a well-designed UI should build a teacher-student relationship that facilitates learning and encourages users to strive for a better understanding of the UI and computer. Shneiderman (2004) has a more grandiose view of UI design, seeing it as a mechanism that can change people's lives. For example, he writes, a professional UI assists doctors in making accurate diagnoses, help pilots fly airplanes safely, and provides children with practical learning tools.

Without sufficiently good UI design, the user may find the product to be unusable and impractical (Lamprecht, 2022) (Shneiderman, 2004). A second-rate UI might even make good product qualities pointless. For example, what is the point of a home button that redirects correctly if the user is unable to find it?

What constitutes good UI design? In most cases, authors on the topic provide heuristics, principles, or rules to better help readers understand the most critical issues of UI design. Shneiderman (2004) has proposed the following eight "golden rules" of interface design:

1. Strive for consistency
2. Cater to universal usability
3. Offer informative feedback
4. Design dialogs to yield closure
5. Prevent errors
6. Permit easy reversal of actions
7. Support internal locus of control
8. Reduce short-term memory load

Nielsen (1994) has provided the following ten usability heuristics:

1. Visibility of system status
2. Match between system and the real world
3. User control and freedom
4. Consistency and standards
5. Error prevention
6. Recognition rather than recall
7. Flexibility and efficiency of use
8. Aesthetic and minimalist design
9. Help users recognize, diagnose, and recover from errors
10. Help and documentation

Mandel (1997) offers a total of 24 “golden rules” divided into three categories:

Place Users in Control

1. Use modes judiciously
2. Allow users to use either the keyboard or mouse
3. Allow users to change focus
4. Display descriptive messages and text
5. Provide immediate and reversible actions, and feedback
6. Provide meaningful paths and exits
7. Accommodate users with different skill levels
8. Make the user interface transparent
9. Allow users to customize the interface
10. Allow users to manipulate interface objects directly

Reduce Users’ Memory Load

1. Relieve short-term memory
2. Rely on recognition, not recall
3. Provide visual cues
4. Provide defaults, undo, and redo
5. Provide interface shortcuts
6. Promote an object-action syntax
7. Use real-world metaphors
8. User progressive disclosure
9. Promote visual clarity

Make the Interface Consistent

1. Sustain the context of users’ tasks
2. Maintain consistency within and across products
3. Keep interaction results the same
4. Provide aesthetic appeal and integrity
5. Encourage exploration

Norman (1983) has provided the following seven principles:

1. *Discoverability*. (It should be possible to determine what actions are possible in the UI, and where and how to perform them.)

2. *Feedback.*

3. *Conceptual model*. (Examples of conceptual models are: icons, folders, and files. The design should provide all the information needed in order to create a good conceptual model.)

4. *Affordances*. (Affordances determine *what* action can be performed. In the UI, the proper affordances should exist to make the desired actions possible.)

5. *Signifiers*. (Signifiers indicate *where* an action can be performed. Effective use of signifiers helps with discoverability.)

6. *Mappings*. (Mapping is the correspondence between controls and their actions. This should be enhanced as much as possible through spatial layout and temporal adjacency.)

7. *Constraints.*

McKay (2013), with his communicative approach, has perhaps the most unique UI design principles:

1. *UI is Communication*. (A UI is nothing more than a conversation between users and a product.)

2. *Explain tasks clearly and concisely, as you would in person.*

3. *Every UI element can be evaluated by what it communicates and how effectively it does that job.*

4. *Be polite, respectful, and intelligent*. (The standards of human interaction should also apply to human-to-computer interaction.)

5. *If a UI feels like a natural, professional, friendly conversation, it is probably a good design*.

It is difficult to attempt to summarize the above heuristics/principles/rules since they all vary in scope, approach, and details. Nielsen's (1994) heuristics and Shneiderman's (2004) rules are perhaps the most similar in these terms. In comparison, Mandel's (1997) rules are somewhat smaller in scope yet more detailed. At the same time, the principles of Norman (1983) have a very wide scope while at the same time being slightly obscure. Finally, McKay's (2013) principles have probably the most comprehensive scope of them all while offering no concrete UI design suggestions. As a matter of fact, his second principle (which is basically "clear communication") more or less entails the complete set of heuristics/principles/rules suggested by the previous authors.

If one were to try to pinpoint the most featured heuristics/principles/rules they would boil down to the following four:

1. *Feedback*. According to Shneiderman (2004), every action should receive system feedback. Minor and frequent actions may only necessitate a modest response, but a

major or infrequent action should require a more substantial response. Mandel (1997) suggests that feedback and progress indicators serve a vital role in the comfort and enjoyment of users. One should give users some indication that an action has been performed, either by showing the results of that action or acknowledging that it has been performed. Norman (1983) adds that feedback should be immediate. If the feedback is delayed, the user might become annoyed and not stay for the response, thus making it wasteful. Furthermore, he writes that too much feedback can be more annoying than little or none. If the UI has too many announcements, it might cause the user to ignore, or even disable, all of them, which means important ones are apt to be missed.

2. *Consistency*. This is a hard rule to follow, according to Shneiderman (2004), since there are many forms of consistency. There should be, he writes, consistency in action sequences of similar situations; terminology in prompts, menus, and help screens; and color, layout, font, etc. Nielsen (1994) writes that one should be mindful of both external consistency (industry conventions) and internal consistency (consistency within a single product or family of products). Mandel's (1997) view on the topic is in accordance with previous authors. A consistent and user-friendly UI, he adds, encourages exploration without fear of negative consequences.

3. *Reduce short-term memory load*. Since human short-term memory is limited, writes Nielsen (1994), the UI must be designed in a way that accommodates this. This is done by making elements, actions, and options visible. The user, he concludes, should not have to remember information from one interface to another. Shneiderman (2004) writes in accordance with Nielsen (1994). Mandel (1997) writes that information provided by the user in an online form, for example, should be remembered by the system once a user has entered it, especially if it is required that the user retype the information a few screens later.

4. *User control*. Shneiderman (2004) writes that experienced users strongly desire to feel in charge of the interface and that it accurately responds to their actions. Surprising interface pop-ups, boring sequences of data entries, inability to obtain information, or inability to produce desired actions, will leave the user frustrated. Nielsen (1994) writes about the need for digital "emergency exits". Since users often make mistakes, he continues, they should have easy access to an undo- or redo-action since this helps users remain in control of the system and avoid getting stuck. According to Mandel (1997), allowing users to be in control of the UI and observing their behavior may provide vital information regarding future design implementations.

2.3.2 Icons and UI Design

When reading the works of Horton (1994) and Hicks (2011), one might get the impression that icons play a very large role in the UI. However, the major books on UI used in this paper (Shneiderman (2004), Norman (1983), Johnson (2010), and McKay (2013)) tell a slightly different story. Shneiderman's (2004) work only has three pages dedicated to icons. Even though he sees the value of icons in certain situations (like a road curve), he does not delve any deeper into other uses of icons. At the same time, Johnson (2010) only dedicates a brief section to icons. He seems to have a favorable outlook on them, admitting that they are generally effective in usage but does not go into further detail. McKay (2013) has a two-sided opinion on icons. Nonstandard icons, he writes, are usually a very poor way to communicate and are best replaced by text labels. However, well-recognized icons, he continues, can be quite effective; especially if they are in accordance with one or more of the following types:

- Standard symbols
- Preview of results
- Simple nouns
- Simple nouns doing simple verbs
- Simple nouns showing simple adjectives
- Well-known logos

Even though icons are frequently used in UIs, they are only brought up briefly in the previously mentioned books on the subject. There is no apparent reason why this is the case. It may lie in the fact that not many academic sources have been written about icons. Also, no concrete scientific research on icons' value in comparison to text labels has been conducted, which makes it hard to determine their ideal uses. Furthermore, internet articles and blog posts concerning icon usage are still relatively sparse and newly published. In the future, perhaps, the subject will take up more space in books on UI.

2.3.3 Conclusion

A professional UI requires good UI design. An effective way to find the essence of good UI design is to consult the principles, heuristics, or rules provided by professionals with years of experience with UI. Through analysis of different principles/heuristics/rules by various authors, four prominent ones have been established: feedback, consistency, reduce short-term memory load, and user control.

Since computer icons reside within the UI, they should naturally adhere to UI design guidelines, including the four above. By looking at icons' different functionalities, as provided by Hicks (2011), Horton (1994), and Zhang (2020a), we notice that they can support conformity towards these guidelines. For example, icon usage can more easily achieve universal usability since icons transcend language barriers. Furthermore, icons can be used to help offer feedback, prevent errors, provide visual cues, and explain tasks clearly.

This paper aims to explore whether icons may provide support to the principle: "reduce short-term memory load", featured by Shneiderman (2004), Nielsen (1994), and Mandel (1997). When these authors write about short-term memory load reduction, it traditionally does not concern whether the information is best directed through icons or text labels, but rather that users should not have to remember information from one interface to another. Therefore, this paper explores a unique approach to the principle in question.

3 Method

3.1 Epistemological starting point

The following study used quantitative data. According to Lancaster (2015), this type of data can be expressed numerically or be classified by some numerical value, unlike qualitative data, which comes in the form of hard-to-quantify descriptive accounts or observations. Quantitative data, he continues, is generally thought to be more objective and reliable than qualitative data and is associated with more traditional scientific approaches to research found in physical science and psychology. According to Humans of data (2018), other advantages of quantitative research are that they provide a relatively concise and conclusive answer to research questions. Moreover, they write results can be generalized to an entire target group when statistically significant sample sizes are used. However, they continue, unlike its qualitative counterpart, quantitative data collection does not account for the participants' thoughts and perceptions about what is being evaluated. Furthermore, they conclude that it does not explore the "why" and "how" behind a phenomenon.

In this case, choosing a quantitative research method came naturally since measuring STM capacity through qualitative data collection is difficult, if not impossible. Furthermore, all previously mentioned studies of STM, which served as an inspiration for this study, used quantitative data.

3.2 Literature collection

In order to gather the appropriate literature, there were three topics that had to be considered: iconography, STM, and UI design.

There is a limited amount of information concerning iconography. The reason for this is not apparent, but most likely, it has to do with the fact that icons function within the UI and that UI/UX design has been overlooked, in terms of importance, until very recently. Therefore, finding many scientific sources on the topic is difficult, if not impossible. As a result, some of the references cited in this paper are blog posts written by designers, iconographers, etc. The primary reference used for the topic of iconography is John Hicks' *The Icon Handbook* from 2011, followed by William Hortons *The Icon Book* from 1994; the latter is somewhat outdated. Both sources were found by searching on Google. Other sources on the subject were found either through Google Scholar or by searching the Web.

Broader research has been conducted on the topic of STM, but pictures' superiority over words in STM is also somewhat limited and not completely understood. The main source of reference on STM came from the book *Cognitive Psychology* by Goldstein and Van Hooff (2018), which was part of the course literature in my previous semester studying neuroscience. The book references a lot of previously conducted research (which could be found by searching on Google) on the topic.

One of the main sources used for the topic of UI design was Shneiderman's *Designing the user interface: Strategies for effective human-computer interaction* (2004), which is a common book in university courses. There are later versions of this book, but I could not find any for a reasonable price. I also used *The design of everyday things* (1983) by Norman, *Designing with the mind in mind* (2010) by Johnson, and *UI is communication* (2013) by McKay. The latter three I found in a list titled: "Top 14 UI design books every designer should read" (Alban, 2022). Mandel (1997) and Nielsen (1994) I found by searching after design principles on Google.

3.3 Data collection

3.3.1 Sampling

30 Swedish participants, the vast majority of whom were between the ages of 20-26, were used in the study. They consisted of 20 males and ten females. All of the participants were chosen using convenience sampling, which is never ideal, but since the testing occurred in my student dorm room, it would have been challenging to convince strangers to participate. The participants, therefore, consisted of people living in the dorm, friends, parents, and friends of friends.

The 30 participants were split into two equally sized groups: one was to be exposed to icon stimuli, the other one to word stimuli. Both groups consisted of an equally sized ratio of females and males, and both had, more or less, the same average age.

Since the study was conducted in a slightly controlled environment, it can be said to be of an experimental character.

3.3.2 Material

Two PowerPoints were used in the study.

The first PowerPoint comprised four slides, of which the first slide provided instructions to the participant. The instructions did not include any information on the purpose of the study; instead, it told the participant that they would see a matrix (Figure 57) for 30 seconds, which would contain 20 icons. They were told to inspect each icon quietly by themselves. The following slide comprised the 5 x 4 matrix. There was a black glyph icon in the center of each pane, which depicted everyday objects such as a car, key, umbrella, sword, moon, etc. The icons had been selected from the PowerPoint application's own picture archive and were judged to be simple and easy to understand in terms of what they depicted. The slide was set to a timer of 30 seconds and automatically switched to the third slide after the time had passed. The third slide, which had a timer of 20 seconds, instructed the participant to count backward in steps of three from the number "650" until the next slide appeared. The fourth and last slide asked the participant to recall as many items they could remember from the matrix.

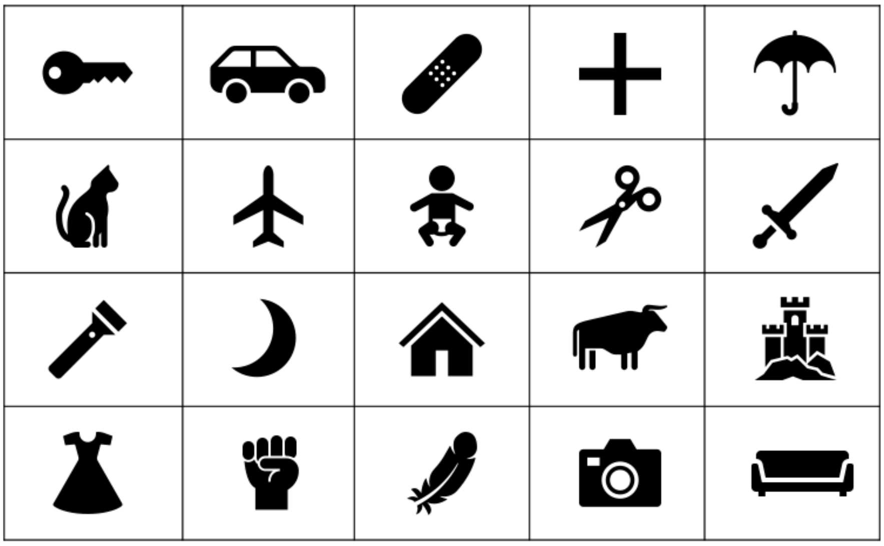


Figure 57 Matrix with icons

Nyckel	Bil	Plåster	Kors	Paraply
Katt	Flygplan	Bebis	Sax	Svärd
Ficklampa	Måne	Hus	Tjur	Slott
Klänning	Näve	Fjäder	Kamera	Soffa

Figure 58 Matrix with words

The second PowerPoint was virtually the same as the first one, having the same amount of slides. The only difference was that the panes of the matrix (Figure 58) in slide two consisted of Swedish words instead of icons. The words corresponded with the glyph icons of the other PowerPoint in terms of meaning and were judged to be common and easily understandable. For example, the pane with an umbrella in the icon matrix instead had the word “Paraply” in the word matrix.

There were a few reasons for displaying all words and icons at the same time, instead of, for example, having each of them displayed individually. First of all, the test was aimed to be simple and easy to understand; having flashing words or images presented to oneself without much context could have been confusing. However, the main reason was to mimic how one normally views icons. Icons in a UI are almost never flashed quickly in front of oneself; instead, they are viewed for a longer amount of time. Naturally, they do not usually appear in a matrix, but the purpose of this was to separate each item from the other more easily.

Having the display of items limited to 30 seconds was judged to be an appropriate time to process all of the items fully. The delay of 20 seconds that followed the stimuli was judged to be of an appropriate length since one naturally does not have to remember something immediately in a UI, but rather after some passing time. However, one can certainly argue against this point. The purpose of the distraction was to inhibit any sort of mnemonic practice. This type of distraction (counting backward) was something I had used in a previous cognitive psychology study and was familiar with.

3.3.3 Procedure

The study was conducted in my student dorm room at Parentesen, Lund. In order to minimize distractions, the room was quiet and kept clean. The only two individuals in the room were the participant and me. Each participant was asked to sit down by a desk facing the wall; in front of them was a computer that displayed either one of the PowerPoints. When the participants were asked to recall as many items as possible by the last slide, I wrote them down in a notebook. There was no time limit to how long they could sit and recall but in every case, this ended after approximately 30 seconds. The whole test took two to three minutes to complete.

3.4 Analysis

A paired t-test was performed with the numerical data collected from all the conducted tests.

According to Hayes (2022), a t-test is a type of statistical test that looks to determine the significant difference between two groups. There are several different types of t-tests, and the most appropriate one is determined by the data and type of analysis required (Hayes, 2022). A paired t-test is used when you want to determine the difference between two variables from two groups of the same population (Hayes, 2022). In this case, the two variables were “icons” and “words”. The two groups of 15 participants each were from the same population and did not differ. For example, one was not completely male while the other one was only female, or one did not consist of old participants while the other one consisted of young participants.

In order to perform the paired t-test, I used a t-test calculator offered by graphpad.com.

3.5 Data quality

A key issue in data collection is the quality of data. Lancaster (2005) lists three dimensions of data quality: validity, reliability, and generalizability. Validity relates to the extent to which the data collection relates to what it is supposed to measure. Reliability relates to the extent to which the data collection approach can yield the same results on different occasions. Finally, generalizability is basically another aspect of validity and relates to the extent to which the result from the sample population is present in the wider population from which the sample is gathered.

All three of the previously mentioned dimensions of data quality have been carefully considered during the course of this study. In terms of validity, the purpose of the study was to measure both STM capacity of words and icons in order to make a comparison. This was done by having two groups of the same population perform two types of tests. In terms of reliability, the data collected is assessed to be decently reliable since the sample population consisted of 30 participants. If the sample population only would have consisted of two participants, it would have been impossible to make any reliable conclusion from the data collected. In terms of generalizability, convenience sampling is not ideal since it usually entails difficulty in making a broader conclusion of the collected data. However, in this case, convenience sampling is not a large issue since STM capacity does not vary significantly between different groups of people. If the study had concerned “opinions about web cookies” and consisted of a questionnaire, convenience sampling would have been a more significant issue since there is a high risk that individuals in one’s close social circle have similar opinions on the topic in question.

3.6 Ethics

According to McLeod (2015), ethics refers to correct rules of conduct when overseeing research. The researcher, he writes, has a moral obligation to respect the rights and dignity of the participants. In psychological research, he continues, the main issues of ethics include informed consent, debriefing, protection of participants, deception, confidentiality, and withdrawal.

- Informed consent entails that the participants agree to participate in the study.
- Debriefing entails that the participants should be able to discuss the findings and procedure with the researcher.
- Protection of participants means that one should not embarrass, frighten, offend, or harm the participants.
- The issue of deception entails that participants should not be deceived about the nature of the research unless necessary.
- Confidentiality entails that data from participants should be anonymous unless otherwise agreed.
- Withdrawal entails that participants should be able to leave the study at any time if they feel uncomfortable.

In order to follow the ethical guidelines provided by McLeod (2015), all participants were asked to give their informed consent to participate in the study. They were also told that they could leave the study whenever they wanted to, that participation was voluntary, and that no personal information would be collected. After the test had been conducted, I was willing to answer any questions they had regarding the findings and procedure.

4 Results

Figure 59 presents the average score of both participant groups (icon vs. word). The participants of the icon group were able to recall an average of 7.73 items from the matrix, while the word group was able to recall an average of 4.40 items. Therefore, icons had a 75.68% higher recall rate than words. The scores were analyzed through a paired t-test which gave a two-tailed p-value of 0.0008 ($p < 0.05$). By conventional criteria, this difference is judged to be highly statistically significant.

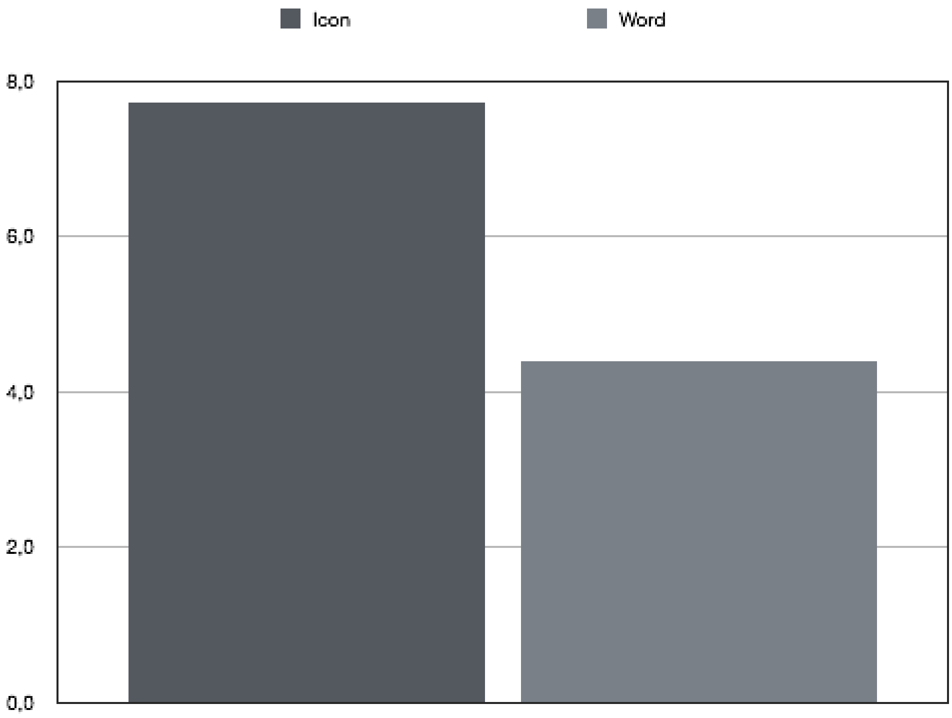


Figure 59 Average items remembered in both categories

5 Discussion

The result of the study suggests that STM capacity for glyph icons is greater than that of words, having a 75.68% higher recall rate after a 20-second delay. The results may seem neither unexpected nor especially profound considering the general opinion, as stated by Grady et al. (1998), that pictures are more easily remembered than words. Moreover, if one compares the results to Pellegrino et al.'s (1976) results from the lenient scoring system (which did not consider positions of items), pictures were likewise significantly better recalled than words following a similar auditory distraction, having a 103.22% higher recall rate. However, while Pellegrino et al. (1976) showed the subjects each stimulus for two seconds, this study presented the participants with the complete set of 20 items for 30 seconds. Even though the difference may seem small at first, one should note that different types of memory tests can yield significantly different results. For example, Shepard's (1967) study did not show any large difference in recognition between words, sentences, and pictures (90% vs. 88% vs. 98%); and Reder et al.'s (2006) placebo results even showed a greater recognition of words rather than pictures.

Glyph icons are widely used in computer UIs, frequently appearing in the world's most used applications and websites, such as Facebook, Instagram, Google, and Reddit. In these circumstances, they provide the user with navigational support, feedback, and status while also transcending language barriers. Since computer icons reside within the UI, they should cohere with essential principles/heuristics/rules of UI design. This paper has established four frequently occurring principles/heuristics/rules, one of which is: "Reduce short-term memory load". Traditionally, this refers to users not having to remember information from one interface to another. However, this paper has introduced a broadened approach to this guideline by exploring how differently "packaged" information may influence short-term memory load; in this case, the different "information packages" were glyph icons and words (or text labels).

According to the findings of this study, glyph icons may offer a reduction of STM load when used instead of words. This is a natural conclusion based on the fact that STM is limited. If more items in the form of glyph icons, rather than items in the form of words, can be stored in STM, it suggests that each glyph icon puts less strain on the STM than each word, even though they provide the same amount of information. With this in consideration, a UI designer deciding to use a glyph icon instead of a text label may enable a UI that more closely adheres to the principle/heuristic/rule of reducing STM load, thus creating a UI that is more efficient and user-friendly.

6 Conclusion

The results of this study suggest that STM capacity of glyph icons is greater than that of words, something which comes as no great surprise considering previous research on STM capacity between pictures and words. Since more items in the form of glyph icons can be stored in STM, rather than items in the form of words, one can conclude that glyph icons put less strain on the STM. This provides a valuable basis for future UI design decisions, where one has to be chosen over the other since reducing STM load adheres to the major UI design principle/heuristic/rule of reducing STM load.

The reason why pictures (or glyph icons in this case) may be superior to words in terms of memory is still a matter of debate. Grady et al. (1998) conducted a study where they sought to determine the different brain activity patterns during memory encoding of words and pictures, respectively. To do this, they used positron emission tomography, which measured regional cerebral blood flow (rCBF) in different brain regions during encoding. The results showed different patterns of activity during encoding between pictures and words. During picture encoding, there was a greater activity of bilateral visual and medial temporal cortices compared to words, while word encoding had a greater activity in prefrontal and temporoparietal regions related to language function. Grady et al. (1998) suggest that pictures induce a more elaborative or associative encoding than words and that this may be the reason why pictures are more easily remembered.

Why pictures would induce a more elaborative brain activity pattern is unclear, but the answer may lie in the fact that humans, and primates in general, are highly visual beings. We strongly rely on our vision to find shelter, food, and mates, but also for more complex behavior, such as parental care and the formation of social hierarchies. Throughout primate evolution, our dependence on visual cues has increased according to adaptive advantages, and consequently, we have developed a brain whose structures involved in visual processing are far more complex than that of other mammals (Kaas & Balaram, 2014). Our primary visual cortex, for example, is much larger in proportion compared to other mammals, and more than 50% of our brain surface is devoted to processing visual information (Hagen, n.d.). Considering how visual we humans are and how important it is to remember various aspects of one's environment for survival, it is not surprising that we may have developed a particularly good memory for pictorial material (Grady et al. 1998).

6.1 Reflections

As McKay (2013) suggested, UI design boils down to communication. This type of communication is primarily visual. Among humans, the basis of visual communication is found in icon-like depictions, as suggested by ancient cave drawings. These two facts combined formed the idea behind this paper. Since icons are continuously used to this day, I figured there to be something powerful about them, something that might not be completely understood. To convey this to the reader, I wrote extensively about the history of icons. I figured that if the reader does not feel engaged by icons, they will neither feel motivated to learn about the study's results. Words are already known to be effective in communication; there is no need to convince anyone of that. However, icons are something one typically does not know much about. If one wants to present a “challenger” to words, one also needs to show that it is a worthy one.

At the beginning of this paper, I knew that Theory would take up a substantial part of the overall work. I feel like this was necessary since the paper concerns three critical topics (UI design, STM, and icons), and also fruitful since these topics have not, prior to this, been thoroughly tied together. Result and Discussion, unlike Theory, do not take up a large part of the overall work. This is a natural consequence of the type of paper that was written. Quantitative results are usually straightforward and concise. In this case, I am looking for a simple straight answer: how do icons compare to words in terms of STM capacity? A single graph with a short paragraph suffices; anything more would be redundant. When it comes to Discussion, I, unfortunately, do not have any previous results to compare my own results to. By this, I mean that there have been no previous studies on STM capacity between icons and words. Instead, I had to compare my results to previous research on pictures vs. words. Furthermore, since it is not completely obvious what the findings of the study may provide for future UI design, it also limits non-speculative conclusions.
Even though the study went well and provided statistically significant results, a few sources of error need to be considered, along with a couple of lessons learned. Firstly, there were a few cases where the participant seemed to assume that the test was a memory test, even though the instructions did not indicate this. If a participant assumes that they are doing a memory test, there is a chance they will do some sort of mnemonic practice that can increase the number of items they can remember. There is certainly room for this with a 30-second display of all items at once. Furthermore, there was a difference between participants when they started counting backward in steps of three. Since the slide after the matrix had a timer set to 20 seconds with written instructions, there was a difference when these instructions were fully understood, and counting commenced. However, this is a minor issue since the main significance of the distraction is to keep the participant's mind occupied.

In terms of method, a couple of points also need to be reflected upon. First of all, convenience sampling is never ideal when performing a study. In this case, however, it probably saved me enough time and work to make it worthwhile, especially considering that a completely randomized sampling would have little or no impact on the results in question. Secondly, I wish I had done more research before I commenced with the testing. Designing a perfect cognitive test is not easy, especially if the type of test (icons vs. words) has not been done before. Therefore, it is necessary to be well-read on the topic of previously performed cognitive tests, which I was not. In hindsight, I might have constructed a slightly different test. First, I would have reconsidered using only an auditory distraction since it may have decreased the participants' ability to recall words more than icons, which was evident in the study by (Pellegrino et al., 1976). However, one should note that having different distractions, auditory and visual, etc., for icons and words, give rise to another issue: how does one know if the two types of distractions are equally distractive? Moreover, the fact that the instructions for the distraction were written in words (introduced directly after the stimuli) may have influenced participants' memory of the stimuli words.

References

Alban, S. (2022, 31 March). Top 14 UI design books every designer should read. UserGuiding. https://userguiding.com/blog/best-ui-design-books/

Alpenfels, E. (1955). The anthropology and social significance of the human hand. *Artificial Limbs*. p. 4-21.

Alvarez, G.A. & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological science, 15*. p. 106-111.

Arm, K. (Ed.). (2010) Hand. *The book of symbols: reflections on archetypal images*. (p. 380-383). Cologne: Taschen.

Arntz, G. (n.d). [Image]. The marginalian. https://www.themarginalian.org/2011/03/09/gerd-artnz-graphic-designer/

Atkinson, R. & Shiffrin, M. (1971). The control processes of short-term memory. *Psychology series*.

Bosch, H. (1503-1504). *The garden of earthly delights* [Painting]. Pinterest. https://www.pinterest.se/pin/382665299563579805/

Brown, J. (1958). Some tests of the decay theory of immediate memory. *The Quarterly Journal of experimental psychology.* p. 12–21.

Canfield Smith, D. (1975). Pygmalion: A creative programming environment. Computer science department: Stanford University.

CNN. (2010, August 31). [Screenshot]. CNN. http://edition.cnn.com/2010/TECH/web/08/31/gmail.priority/index.html

Cursor. (n.d.). [Image]. Cursor. https://cursor.in/

Design week. (n.d.) [Image]. Design week. https://www.designweek.co.uk/issues/2-6-january-2019/duolingo-owl-receives-makeover-as-app-becomes-more-gamified/

Ellwood-Hughes, P. (2020, August 10). James Dean: The enduring legacy of a cultural icon. https://entertainment-focus.com/2020/08/10/james-dean-the-enduring-legacy-of-a-cultural-icon/

Flat icon. (n.d.). [Image]. Flat icon. https://www.flaticon.com/free-icon/trophy_188906

Free icon shop. (n.d.). [Image]. Free icon shop. https://freeiconshop.com/icon/lock-icon-glyph/

Free icons library. (n.d.). [Image]. Free icons library. https://icon-library.com/icon/nuclear-icon-png-29.html

Gittins, D. (1986). Icon-based human–computer interaction. *International Journal of Man-Machine Studies, 24*(6), p. 519–543.

Goldstein, E. & Van Hooff, J. (2018). Short-term and working memory. *Cognitive psychology*. (p. 122-160). Andover: Cengage Learning EMEA.

Google Maps. (2022). [Screenshot]. Google Maps. https://www.google.se/maps/@58.3664257,12.1690005,954m/data=!3m1!1e3?hl=sv

Grady, C., McIntosch, A., Rajah, M., & Craik, F. (1998). Neural correlates of the episodic encoding of pictures and words.

Hagen, S. (n.d.) The mind's eye. https://www.rochester.edu/pr/Review/V74N4/0402_brainscience.html

Hayes, A. (2022, March 12). *T-test.* https://www.investopedia.com/terms/t/t-test.asp

Hicks, J. (2011). *The icon handbook*. Penarth: Five Simple Steps.

Hodge, S. (2008, April 12). *7 principles of effective design*. https://design.tutsplus.com/articles/7-principles-of-effective-icon-design--psd-147

Horton, W. (1994) *The icon book: Visual symbols for computer systems and documentation*. New York: John Wiley & Sons, Inc.

Humans of Data. (2018, June 11). *A complete guide to quantitative research methods.* https://humansofdata.atlan.com/2018/06/quantitative-research-methods/

Indeed Editorial Team. (2022, March 30). *What Does a UI Visual Designer Really Do? (With Job Duties).* https://www.indeed.com/career-advice/finding-a-job/what-does-ui-visual-designer-do

Instagram. (2022). [Screenshot]. Instagram. https://www.instagram.com/

Johnson, J. (2010). *Designing with the mind in mind*. Burlington: Elsevier.

Kamenez, G. (2021, March 24). *Understanding icons in UI design*. https://uxcel.com/blog/understanding-the-icons-in-ui-design

Kaas, J. and Balaram, P. (2014). Current research on the organization and function of the visual system in primates.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of verbal learning & verbal behavior, 1*(3). p. 153–161.

Lalla, N. (n.d.) Icon. *Oxford Companion to Art*. p. 555.

Lamprecht, E. (2022, January 3). *The Difference Between UX and UI Design – A Beginner's Guide*. https://careerfoundry.com/en/blog/ux-design/the-difference-between-ux-and-ui-design-a-laymans-guide/

Lancaster, G. (2005). *Research Methods in Management A concise introduction to research in management and business consultancy*. Oxford: Elsevier Butterworth-Heinemann.

Lange, D. (1958). *Hand, Indonesian dancer* [Picture]. Moma. https://www.moma.org/collection/works/56399

Lenaerts, S. (2014, January 7). *Use and abuse of icons in the modern age*. https://webdesign.tutsplus.com/articles/use-and-abuse-of-icons-in-the-modern-age--webdesign-17064

Mac rumors. (n.d.). [Image]. Mac rumors. https://www.macrumors.com/how-to/out-of-office-replies-apple-icloud-mail/

Material. (n.d.) [Image]. Material. https://material.io/design/iconography/system-icons.html#design-principles

Mandel, T. (1997). *Golden rules of user interface design*. https://theomandel.com/resources/golden-rules-of-user-interface-design/

Max pixel. (n.d.). [Image]. Max pixel. https://www.maxpixel.net/Ban-On-Bringing-Food-Not-To-Bring-Food-1746278

McKay, E. (2013). *UI is communication*. (p.15-16). Waltham: Elsevier.

McLeod, S. (2015). *Psychology research ethics*. https://humansofdata.atlan.com/2018/06/quantitative-research-methods/

Microsoft. (n.d.) [Image]. Microsoft. https://support.microsoft.com/en-us/office/change-your-status-in-teams-ce36ed14-6bc9-4775-a33e-6629ba4ff78e

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review, 63*(2). p. 81–97.

Myloview. (n.d.) [Image]. Myloview. https://myloview.se/canvastavlor-hand-stop-icon-vector-sign-symbol-vector-nr-BA50133

Nielsen norman group. (n.d.) [Image]. Nielsen norman group. https://www.nngroup.com/articles/errors-forms-design-guidelines/

Nielsen, J. (1994). *10 usability heuristics for user interface design*. https://www.nngroup.com/articles/ten-usability-heuristics/

Norman, D. (1983). *The design of everyday things*. New York: Basic Books, 2013.

Oates, J. M., & Reder, L. M. (2011). Memory for pictures: Sometimes a picture is not worth a single word. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork*. p. 447–461.

Paivio, A., & Csapo, K. (1971). Short-term sequential memory for pictures and words. *Psychonomic Science, 24*(2), 50–51.

Pellegrino, J., Siegel, A., & Dhawan, M. (1976). Short-term retention of pictures and words as a function of type of distraction and length of delay interval. *Memory & Cognition* 4(1). p. 5-11.

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of experimental psychology, 58*(3). p. 193–198.

Pinterest. (n.d.a). [Image]. Pinterest. https://www.pinterest.co.uk/pin/110267890848036914/

Pinterest. (n.d.b). [Image]. Pinterest. https://www.pinterest.se/pin/203436108148649947/

Reddit. (2022). [Screenshot]. Reddit. https://reddit.com/

Reddit. (n.d.). [Image]. Reddit. https://www.reddit.com/r/etymology/comments/da2cs3/evolution_of_chinese_characters/

Reder, L. M., Oates, J. M., Thornton, E. R., Quinlan, J. J., Kaufer, A., & Sauer, J. (2006). Drug-induced amnesia hurts recognition, but only for memories that can be unitized. *Psychological science*, *17*(7). p. 562–567.

Robinson, A. (1995) *The story of writing: Alphabets, hieroglyphs & pictograms*. London: Thames & Hudson Ltd.

Roser, M. & Ortiz-Ospina, E. (2018, September 20). *Literacy*. https://ourworldindata.org/literacy#citation

Rublev, A. (1425). *Trinity* [Painting]. Catholic link. https://catholic-link.org/andrei-rublevs-icon-of-the-holy-trinity-explained/

Shepard, R. N. (1967). Recognition memory for words, sentences, and pictures. *Journal of verbal learning & verbal behavior, 6*(1). p. 156–163.

Shneiderman, B. (2004) Use the eight golden rules of interface design. *Designing the user interface: Strategies for effective human-computer interaction.* (p. 74-75). Boston: Addison-Wesley

Stardock. (2020, November 11) *The evolution of computer icons*. https://www.stardock.com/blog/502254/the-evolution-of-computer-icons

Stickpng. (n.d.). [Image]. Stickpng. http://www.stickpng.com/img/electronics/phone-icons/outlined-phone-icon

Svgrepo. (n.d.). [Image]. Svgrepo. https://www.svgrepo.com/svg/363399/camera-duotone

The verge. (2022). [Screenshot]. The verge. https://www.theverge.com/2022/1/31/22911287/gmail-redesign-workspace-rollout-april-q2-chat-meet-spaces

Twyman, M. (1975) *The significance of Isotype.* https://isotyperevisited.org/1975/01/the-significance-of-isotype.php

UI design. (n.d.). https://xd.adobe.com/ideas/process/ui-design/

Vecteezy. (n.d.). [Image]. Vecteezy. https://www.vecteezy.com/vector-art/2103774-hand-writing-with-pen-black-glyph-icon

Vira les locos. (n.d.). [Image]. Vira les locos. https://viraleslocos.com/como-activar-bluetooth-en-tu-mac-y-emparejar-nuevos-dispositivos/

We find wilderness. (n.d.). [Image]. We find wilderness. https://www.we-find-wildness.com/2010/05/hobo-signs/

Wikimedia. (n.d.). [Image]. Wikimedia. https://commons.wikimedia.org/wiki/File:Facebook_Thumb_icon.svg

Wikimedia. (n.d.b). [Image]. Wikimedia. https://commons.wikimedia.org/wiki/File:Egyptian_hieroglyphs-Iteru.svg

Wikipedia. (n.d.a). [Picture]. Wikipedia. https://en.wikipedia.org/wiki/Pech_Merle

Wikipedia. (n.d.b). [Image]. Wikipedia. https://sv.wikipedia.org/wiki/Favicon

Youtube. (2022). [Screenshot]. Youtube. https://www.youtube.com/

Zhang, H. (2020a, April 30). *Foundations of iconography: Definition, benefit, and purpose*. https://uxdesign.cc/foundations-of-iconography-f95d7233a3e6

Zhang, H. (2020b, January 29). *7 principles of icon design*. https://uxdesign.cc/7-principles-of-icon-design-e7187539e4a2