IaSDR07 INTERNATIONAL ASSOCIATION OF SOCIETIES OF DESIGN RESEARCH THE HONG KONG POLYTECHNIC UNIVERSITY 12TH TO 15TH NOVEMBER 2007

Collaborative Design Research: The Visualization of Medical Concepts

Mike Zender, MAF, Associate Professor
Director of Graduate Studies, School of Design
University of Cincinnati, College of Design Architecture Art and Planning,
Center for Design Research and Innovation

Keith A. Crutcher, Ph.D., Professor
Director of Research, Department of Neurosurgery
University of Cincinnati, College of Medicine

ABSTRACT

The proliferation of data is threatening to swamp our ability to convert data into knowledge. Visualization promises to facilitate this conversion. Yet visual communication designers have not been deeply involved. One potential impediment to involvement is the lack of collaboration between visual communication designers and knowledge workers in specialized domains This paper describes a collaborative research project that integrates medical science and visual communication design. The project involves the development of a visual language to represent medical concepts by deriving propositions from papers, breaking propositions into concept objects, designing a visual object system (consisting of icons, glyphs and combinations) to represent the objects, and displaying the objects as a network of concepts with links to the original papers. Prototypes have proven to be highly condensed and accurate yet readable in seconds. If the visualization approach proves successful, the results would be groundbreaking in science and design.

Keywords:
Visualization
Collaboration
Icons

OUTLINE OF A PROBLEM

There is so much to know and there is so little time. Or so it seems. Thanks to the computer and media that flow through it, our lives are filled with information and the data that surrounds it. According to Lyman and Varian, in the year 2000 the world produced between one and two exabytes (a billion gigabytes) of unique information, about 250 megabytes for every man woman and child on earth (Lyman and Varian, 2000). By 2003 this was five exabytes annually. While the quantity of data increases with each passing day, each passing day remains fixed at 24 hours: we have increasing data and fixed time.

DATA QUANTITY

The pressures of data and time are felt particularly by those who specialize in the pursuit of new knowledge: scientists and scholars. For example, many of those working in the scientific and fields use an online library known as PubMed which "contains bibliographic citations and author abstracts from more than 5,000 biomedical journals published in the United States and 80 other countries. The database contains over 15 million citations dating back to the mid-1950's." (NLM/NIH, web, 2007). Suppose a researcher went to PubMed to find out something about Alzheimer's Disease. A recent query (10/24/06) using "Alzheimer*" as the search term returned 54,430 citations. Even for a rapid reader it would take years of uninterrupted reading time to review 54,000 papers. What's more, with papers on Alzheimer's being published at a current rate of over 5,000 per year, as soon as the stack of past papers had been mastered a new stack of unread papers would have accumulated. Clearly, reading every paper is not a viable means of staying current with developments in even one limited field of knowledge. This constant growth in information in medicine and science is not an isolated event. Similar growth also occurs in other fields.

DATA COMPLEXITY

One might imagine dealing effectively with large volumes of data that are homogeneous and simple. Unfortunately, most fields, including the scientific and medical systems noted above, are complex and draw upon specialized knowledge involving data drawn from various levels of organization ranging from atomic particles to populations.

SPECIALIZATION AND ISOLATION

A partial solution to quantity and complexity of data is for individuals to specialize in knowledge domains so that the amount of information needed to be mastered is a small fraction of the total. In fact, this is largely what has happened in most areas of science. Individual scientists have become increasingly knowledgeable about smaller domains. However, a potential negative consequence of such specialization is the loss of insights that come from familiarity with other domains or from understanding the global context of the information.

CONVERTING DATA INTO KNOWLEDGE

In fact, a key to converting data into knowledge is placing it in context. Perceptual Psychologist Rudolf Arnheim affirmed that everything is affected by context, "The mind meets here, at an elementary level, a first instance of the general cognitive problem that arises because everything in this world presents itself in context and is modulated by that context." (Arnheim, 1969) The importance of context is exemplified in the very means we use to encode knowledge: language. To establish the proper meaning of a word requires the context of the sentence. To understand the proper meaning of a sentence, one must know the context of the broader story. Historian and theologian N. T. Wright describes this phenomenon using the example of the sentence "It is going to rain." If one says "It is going to rain," and the context is a story of a planned picnic the sentence means disappointment, while in the context of a prolonged drought in Africa the sentence means jubilation. (Wright, 1996).

A special field has emerged to convert large quantities of data into knowledge using visual means: Information Visualization. The power of visual perception to identify patterns and aid in understanding information has been well established (Tufte, 1983-97). Much of the cerebral cortex is devoted to visual processing. It has been said that we understand more of the world around us through visual perception than through all the other senses combined (Ware, 2004). Information Visualization has grown into a field with its own journals, conferences, theoretical basis and research foci. Science and medicine have embraced Information Visualization with medical scientists, computer engineers, and programmers engaged in Information Visualization sub-specialties such as bioinformatics and biomedical informatics. However, even though visualization is defined as a form of communication (DaFanti, Brown and McCormick, 1989), visual communication designers have not been deeply involved in Information Visualization or in the domains of science where the information is generated. One potential impediment to designer involvement is the relative paucity of collaborative relationships between scientists and communication designers. Science and design are, after all, different disciplines with different tools and traditions. Although science is usually characterized as systematic and rigorous and design tends to be viewed as creative and artistic, these attributes are not unique to each discipline and should arguably compliment each other. Another obstacle is that there are few mechanisms available for funding such collaborative projects. Fortunately for the authors, the University of Cincinnati's forward thinking University Research Council sponsors interdisciplinary collaborative grants. This past year the authors were awarded one of these grants to advance the project reported here.

SKETCH OF A SOLUTION

COLLABORATION

Recognizing the scope and complexity of data and the limitations of specialization, the authors, one a scientist [!!] and the other a designer [**], formed an interdisciplinary collaboration to identify grounds for integrating scientific approaches and design approaches to better visualize scientific concepts. Recognition of the benefits of collaboration has led research-funding agencies to increase calls for interdisciplinary collaboration to solve problems.

PROJECT PROBLEM DEFINITION

The first step toward effective collaboration was to define a specific problem to which two disciplines, science and design, could effectively contribute. One of us (!!), based on frustration with attempts to keep current with the explosion of literature in his field (discussed in the beginning of this paper), and the other of us (**), familiar with non-verbal communication, jointly identified a potential area for investigation in the visualization of scientific concepts. Quoting from a previous report on the author's collaboration:

> "Recognizing that much was being done to visualize data, the authors wondered whether it might be possible to visually represent the key concepts and ideas found in scientific papers in a more immediate way than text-based approaches. The ability of visual form to summarize large data sets is well established (Tufte, 1983; Ware, 2004). The ability of icons to communicate concepts is similarly well documented and a part of everyday life (Arnheim, 1974). The utility of scientific and mathematical visual notation systems is also commonplace, although these systems, like all sign-based systems, require special learning. We wondered if key concepts in fields with controlled vocabularies, such as medicine, might be efficiently communicated with images such as glyphs or icons, and, if so, whether these images might then effectively illustrate the web of conceptual connections spread across hundreds or thousands of journal articles and papers within a specific area of investigation. If such a system were interactive, we suspect that it might lead scientists to insights more quickly than scanning mountains of papers. If such a system also remained linked to individual papers then such a visual display might be an improved means of exploring a literature database such as PubMed."
> (Zender, Crutcher, 2007)

Identifying a problem where both disciplines could reasonably contribute was an important step toward successful collaboration, but far from the only issue to be overcome.

CONTEXTS FOR A SOLUTION

COLLABORATION

Relationship and Philosophy

Before starting work on the specific problem we had informal dialogues on the nature and role of visual communication, art and design on the one hand, and their relationship to knowledge, means of knowing and verification in science on the other. These discussions laid a personal and philosophical foundation for fruitful collaboration. Our philosophical and methodological discussions included the merits of various approaches to forming knowledge from the artistic and poetic to the analytic and scientific. Just as importantly, as we discussed these issues, mutual respect and trust, which had begun through personal friendship prior to this collaboration, were strengthened through dialogue. Each of us took time to study the other's field and methods. As we did so, several specific issues emerged.

Vocabulary

Through our discussions we realized that in many ways we spoke different languages. Whereas science has developed a variety of specialized controlled vocabularies, design has a more limited and loosely defined vocabulary. For example, designers often refer to color in the sense of hue as opposed to value or saturation, yet all three: hue, value and saturation (or chroma or intensity) are different and essential aspects of color. Even when designers are careful in their use of language by using 'hue' when referring to a dominant wavelength of visible light, there are only six to eight consistent hue names among the millions of different hues that humans can perceive (Ware, 2004). The failure to have an agreed upon naming system for color, an essential component of most visual communication design, is very different from the experience of a scientist familiar with a precise

Latin-derived name for nearly every part of the human body, e.g., tibia, femur, humerus, etc., or every chemical structure. When a designer specifies 'red' to represent an important data element a scientist might reasonable ask 'which red?' or 'what is red?' In order to collaborate effectively, we needed to come to grips with the complexity of the vocabulary on the one side (science) and the paucity and vagueness of vocabulary on the other (design). In fairness, the relative scarcity of precise vocabulary in design likely reflects the types of problems that designers have focused on rather than an intrinsic lack of rigor. In most cases, precisely which color of red is chosen to highlight a design feature does not carry the same level of importance as knowing, for example, which chemical is used as a drug. In fact, one of the goals of this project was to increase the rigor with which design principles can be applied to the problem of concept visualization.

Method

Behind the issues of language lay another significant issue, that of methodology. As discussed in another paper presented by one of the authors at this conference (**), scientific method often focuses on one variable at a time while design often create by modifying many variables simultaneously as they work toward a solution.

Approach

While it seems as though science and design employ very different methods the authors found this was in fact not the whole story. Despite the apparent differences in method we found an overwhelming common ground of a problem-solving approach. Both scientists and designers define problems then create solutions to them. Scientists are necessarily creative and designers are also systematic. For example, when it became clear that we would need to find a way to compare the effectiveness of different prototypes, we both agreed that this required a quantitative assessment rather than simply relying on subjective impressions. In other words, where rigor was needed, a rigorous approach would be taken.

LANGUAGE COMPLEXITY

In addition to the challenges of working collaboratively across these disciplines, there are significant obstacles in the task itself, i.e., visualizing concepts. One problem is language. In scientific literature, as in most other areas, findings are reported in writing and the concepts are embodied in words. Yet words are often difficult to define, requiring a context to determine their meaning (Wright, 1992). Fortunately for this project, scientific words are often, although not always, highly specialized and specific. Moreover, in medicine there are well-established specialist vocabularies and these vocabularies are organized in a structured conceptual system called the Unified Medical Language System (find the UMLS at: www.nlm.nih.gov/research/umls/). This system includes 135 semantic types and 54 semantic relations that organize medical knowledge in a hierarchy of parent/child relationships. In addition, PubMed uses a vocabulary known as MeSH (Medical Subject Headings) to classify and categorize the content of papers. This is an open vocabulary designed to adapt while retaining control over the database of indexing terms used. Another related technology is Natural Language Processing (NLP) software, which is able to parse electronic texts and correctly identify key words, such as UMLS terms.

LACK OF VISUAL LANGUAGE

Concepts not embodied in words may be symbolized in visual language in the form of pictures, icons and glyphs or signs. In scientific literature there are often figures that accompany a paper to facilitate communication. However, even a cursory review of scientific literature reveals that there is nothing approaching a standard means of visually expressing key concepts. Though there is no existing visual language for scientific concepts, there are icon systems in other fields, such as Olympic venues and transportation signage. One of us has studied transportation icon systems to identify how contexts such as the environment serve as clues to define meaning (Zender, 2006). The same study demonstrated that existing icon systems successfully visualize physical objects but rarely communicate processes and actions. This is akin to attempting a written language with no verbs. As a result, even the most effective and comprehensive icon systems do not communicate as precisely as written language. Developing principles for a comprehensive visual language became, therefore, one goal of this project.

VISUAL STUDIES

But to develop an effective system for visualizing scientific concepts, there is an even deeper problem than gaps in visual vocabulary and that has to do with the meager knowledge of the parameters of visual form. Designers simply do not know much about the visual forms they use as tools in communication. Recent work by Colin Ware has approached design from a scientific and perceptual perspective in an attempt to define principles that might guide scientists in the use of visual form to visualize information (Ware, 2004). One of the authors, in a paper presented at this conference, has described studies exploring the theories presented by Ware but from a design perspective (Zender, 2007). These studies of preattentively processed visual form begin to describe and define parameters for making form pop-out from its surroundings, thereby controlling hierarchy. These studies support and enlarge the possibilities for the design of a more precise and comprehensive visual language system such as is necessary for the expression of scientific concepts.

FACETS OF A SOLUTION

TRACTABLE SCOPE

Within the context of a collaborative team, we believed developments in controlled vocabulary, the design of visual icons and the new parameters for visual form provided a foundation for representing medical concepts published in papers. A key remaining obstacle to designing a prototype solution was the scope of the problem. Science is a vast field with many specialties as noted above. Even medicine, one branch of science, is highly specialized. We decided to make the problem tractable by focusing on an area of expertise of one of us (!!): the etiology of Alzheimer's Disease. We hypothesized that demonstration of the feasibility of this approach in a large but defined domain such as Alzheimer's Disease could be extensible to other areas of science.

APPROACH

The general approach to the project problem was to identify key concepts, connect those concepts in summary statements, break those statements into their essential conceptual objects, illustrate those concepts using icons and glyphs, and present these visual objects in an interactive concept space where they could be immediately perceived and understood in relation to each other. The perception of concepts in context was expected to facilitate exploration and discovery.

Extract Propositions

A key problem at the outset was how to extract propositions from published papers. As reported previously, "For our project, papers were reviewed manually based upon a random selection of 40 papers from PubMed based on a search with the terms "Alzheimer's Disease" and the protein "ApoE" (one area of Alzheimer's disease research with which one of us [!!] is familiar). From the papers, 20 propositions were extracted that express key concepts." (Zender, Crutcher, 2007.) These statements (Figure 1), in positive and negative terms to include both sides of the proposition, are shown below.

Figure 1

Propositions (or hypotheses) from Papers

–

1) Polymorphisms of apoE are (not) associated with the risk of Alzheimer's disease.
2) Polymorphisms of apoE are (not) associated with the risk of multiple sclerosis.
3) Polymorphisms of apoE are (not) associated with the risk of autism.
4) Polymorphisms of apoE are (not) associated with glaucoma.
5) Polymorphisms of apoE are (not) associated with outcomes following head injury.
6) Polymorphisms of cathepsin D are (not) associated with the risk of Alzheimer's disease.
7) Proteolysis of apoE is (not) associated with neuronal degeneration.
8) ApoE does (not) regulate metabolism of ß-amyloid.
9) Estrogen does (not) modulate the expression of apoE.
10) Polymorhisms of apoE does (not) interact with herpes simplex virus to modify the risk of Alzheimer's disease.
11) Cathepsin D does (not) degrade apoE.
12) ß-amyloid does (not) cause neuronal degeneration.
13) The C-terminal fragment of apoE does (not) bind to ß-amyloid.
14) ApoE is (not) required for plaque formation.
15) Cathepsin D is (not) present in plaques.
16) ApoE is (not) present in plaques.
17) ApoE is (not) produced by macrophages.
18) ApoE is (not) produced by nerve cells.
19) ApoE is (not) produced by astrocytes.
20) ApoE does (not) affect long term poteniation (LTP).

–

Since the start of work on our project, others have developed automated or computer mediated techniques for extracting concepts from papers, notably the Telemakus system (Revere, 2003, www.telemakus.org).

Identify and Organize Objects

The proposition statements were then organized into categories that correspond roughly with the ontological structure of the UMLS and relevant MeSH terms. The category decisions were the result of collaborative discussion between the authors. To the scientist [!!], levels of analysis are significant: from the molecular or genetic level up to the cellular, tissue, organ and organism levels. These levels help organize information and guide analysis. For the designer they become means for supplying important interpretive context for visual form that guides correct interpretation of the icons. We referred to these categories as ‘objects’ in the sense of modular conceptual elements, like individual words that could easily be rearranged to form visual propositional ‘statements’. The objects were placed into three categories: things, processes, and actions, analogous to nouns, verbs and gerunds in language (Figure 2).

Figure 2

Objects found in Extracted Propositions

–

MOLECULAR LEVEL BIOCHEMICAL OBJECTS
polymorphisms of apoE / apoE
C-terminal fragment or apoE
cathepsin D
ß-amyloid
estrogen
herpes simplex virus
CELLULAR LEVEL BIOLOGICAL OBJECTS
plaques
macrophages
nerve cells
astrocytes

DSEASE OBJECTS
Alzheimer’s disease
multiple sclerosis
autism
glaucoma
head injury
herpes

PROCESS OBJECTS
proteolysis
neuronal degeneration
long term potentiation (LTP)

ACTIONS
does regulate / does not regulate
does modulate expression / does not modulate expression
does interact with / does not interact with
does cause / does not cause
does bind / does not bind
is associated / is not associated
is required / is not required
is present / is not present
is produced by / is not produced by
does affect / does not affect
does degrade / does not degrade
to modify

Convert Objects to Icons/Glyphs

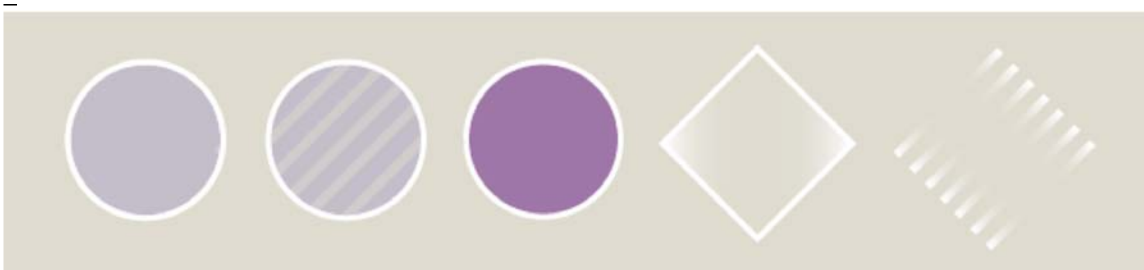
Beginning in 2002, twenty different teams of student designers, working under the direction of the authors, have converted each conceptual object into a visual icon / glyph. As described elsewhere by one of us [**], an icon visually represents a concept through a process of abstraction (Zender, 2006). Unlike an icon, a glyph is a non-representational signifier. Glyphs such as Arabic numerals "communicate content without reference to any particular visual attribute of the thing being represented." (ibid.) Combining graphic forms that serve both icon and glyph functions allowed the icons to signify the specific meaning of an individual object while the glyph form was used to place the icon in a categorical context that could inform and further specify its meaning. In the example below (Figure 3) the icons for objects are all in circular containing shapes that serve a glyph (non-representational) function. The disease objects on the other hand are in diagonal squares that distinguish them from the other objects. Glyphs thus signify category while icons signify meaning within a category. The potential for combining or 'layering' different representational levels within an icon to clarify meaning has also been discussed previously (ibid.).

Figure 3

Sample Glyphs

Circle for object, circle with diagonal lines for object process, diagonal square for diseases.

–



–

VISUALIZATION SYSTEM

The icons/glyphs were not designed to work in isolation. As stated previously by the authors,

> "The icon / glyphs were conceived not as isolated visual objects but as an integrated system of communication objects designed to be read together. Designing icons to work together adds to the context of the entire system so that each icon helps inform the interpretation of every other icon. … The ultimate aim was to combine icons with more abstract visual shapes and icon modifiers in a system that could express complex visual concepts." (Zender, Crutcher, 2007.)

The process for developing these icon/glyph systems was not isolated either. Individual student designers each researched several objects and then visualized them as an icon alone, presented their proposed icon to their peers and placed their individual icons in a 'pool' of icons on a server.

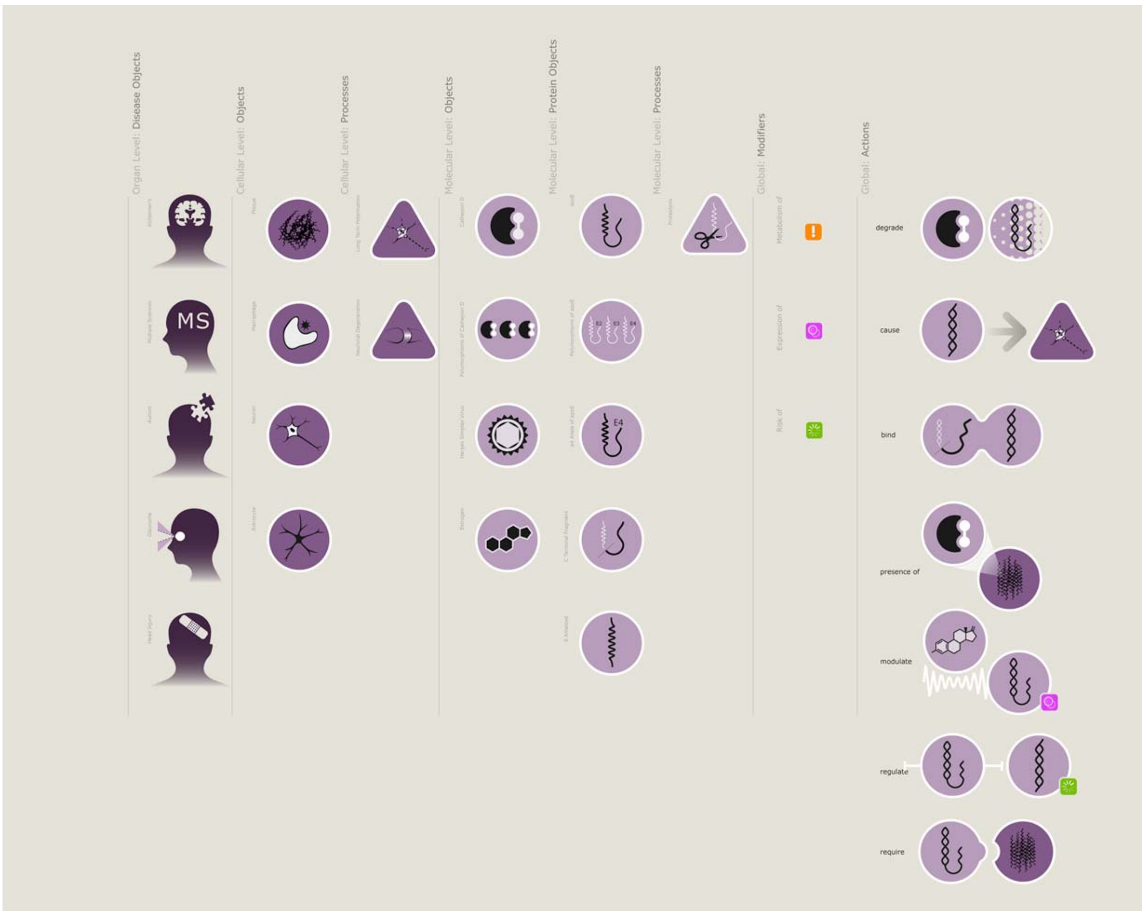
Students then formed small teams of 3 – 4 designers that gathered all of the individual icons together then, as a team, conceived and designed the individual icons into a system by conceiving and designing a glyph system to organize and categorize the individual icons into a coherent system described above. One such system, shown below (Figure 4), was designed by students Sean Gresens, David Kroner, Nolan Stover and Luke Woods.

Figure 4

Visualization System

Student designers: Sean Gresens, David Kroner, Nolan Stover and Luke Woods

–



–

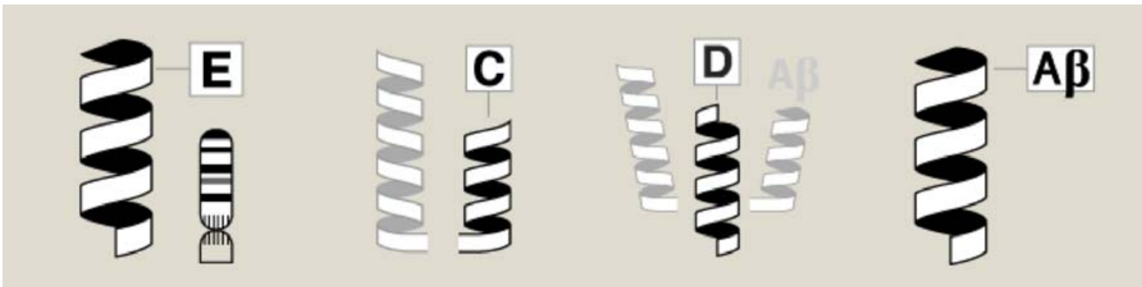
Tangible Objects

As has been noted previously, icons have been proven effective at representing tangible objects, the language equivalent of nouns (Zender, 2006). In the development of the requisite Biological or Biochemical Object icons, designers leveraged the systematic propensities of scientists by designing families of tangible object icons. Icon families mimic the parent/child structure of the UMLS and by doing so created icons with a 'proximate context' that enables one icon's meaning, the parent, to inform the meaning of other icons, the children (ibid.). An example of this can be seen in the ApoE, C-terminal fragment of ApoE and Cathepsin-D and beta-amaloid icons (Figure 5).

Figure 5

ApoE, C-terminal fragment of ApoE and Cathepsin-D and beta-amyloid icons

–



–

In these icons an alternating black and white zigzag form is interpretive of one typical means scientists use to represent proteins. Apolipoprotein E, a C-terminal fragment of this protein, the enzyme cathepsin D, and the peptide known as beta-amyloid all share the property of being made up primarily of amino acids, the sequence of which provides the specific identity. This conceptual relationship is conveyed by repeating the zigzag form in each icon representing a protein related object while modifying that form and adding other icons or graphic devices such as letterforms to it to add to the base meaning. Seen together, these icons form a context for each other that help clarify their meaning. In addition to the icons shown here, we explored making icons for the various levels of the UMLS hierarchy, in this case a generic protein icon, followed by ApoE and it's child C-terminal fragment along with it's sibling Beta-Amyloid (Figure 6). This family icon approach is a clear example of collaborative engagement causing the structure of science to interact with the principles of design, each informing the other: the parent/child structure of the UMLS supports the communication design principle of proximate context.

Figure 6

Unified Medical Language System

–

UMLS Semantic Types (135)

- Entity|A
- *Physical Object|A1 (4)*
- Organism|A1.1
- Anatomical Structure A1.2
- Manufactured Object|A1.3
- Substance|A1.4
  - Chemical|A1.4.1
    - Chemical Viewed Functionally|A1.4.1.1
    - Chemical Viewed Structurally|A1.4.1.2
      - Organic Chemical|A1.4.1.2.1
        - Nucleic Acid, Nucleoside, or Nucleotide|A1.4.1.2.1.5
        - Organophosphorus Compound|A1.4.1.2.1.6
        - Amino Acid, Peptide, or Protein|A1.4.1.2.1.7
          - Protein
            - Apoproteins
              - Apolipoproteins
                - Apolipoproteins E (apoE)
                  - C-terminal fragment
            - Amyloid
              - Amyloid beta

–

Processes

Building on the family approach of the protein icons, process conceptual objects were developed by taking an icon for a thing, such as a neuron, and modifying it to suggest a process: degeneration, hence the process object: neuronal degeneration (Figure 7).

Figure 7

Process Visualization: Neuronal Degeneration

Two different approaches, among many, to modifying the neuron: lines and dots, to signify degeneration.

–

–

As noted previously, "Such (process) objects are roughly similar to gerunds, or verbal nouns, in language. Gerunds in English are generally formed by adding 'ing' to the end of a word, such as 'time'. 'Timing' is a noun that denotes a process (verb) by adding "ing" to "time." In the propositional statements we analyzed, "neuronal degeneration" is one example of a process object: a neuron (thing, noun) degenerates (dies). This would be a conceptual entity in the UMLS." (Zender and Crutcher, 2007) As in all the icons shown, several approaches were developed and evaluated.

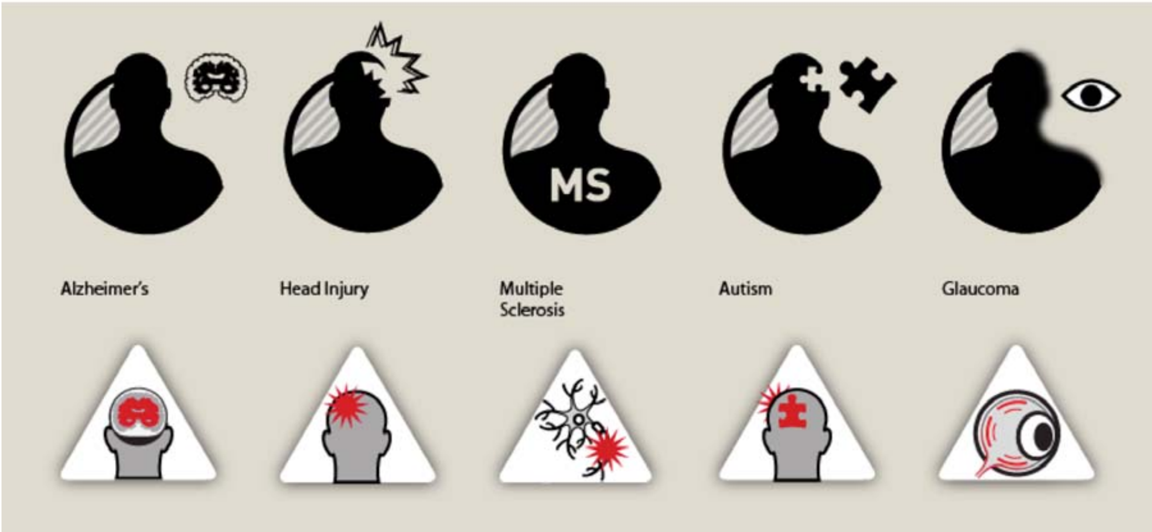
Modifiers

As just noted above, it was found to be useful to develop modifying visual forms to supplement the basic categories of things, actions and processes. These modifiers can perform functions similar to the role adjectives or adverbs serve in language. Or they can serve as more conceptual modifiers when layered with other icons. In the example below (Figure 8) a starburst form was developed to suggest trauma, injury or impairment. Combined with the head the suggestion is "head injury," the same starburst combined with a neuron suggests damage to the neuron's myelin sheath associated with Multiple Sclerosis. Both these are 'disease objects' in our categorization though one, head injury, is not strictly a disease. The use of a modifier to suggest disease led another team of designers to design a disease pattern, a series of diagonal lines in the background of each disease icon and yet another team to develop a disease 'texture' to layer with the relevant body part.

Figure 8

Modifiers for Disease Objects

Sample modifiers:: top row, diagonal lines; bottom row, starburst (not shown, texture).



As was discussed above, glyph forms were also used to provide a categorical context for icons, further enhancing the proper reading of the icon/modifier combinations. At this time it is not determined whether there will be confusion if a modifier is used for both a disease and a process. Again, collaboration with science may provide the answer. The UMLS groups diseases under the broad category of "Phenomena or Processes|B2" (Figure 9).

Figure 9

UMLS Diseases

–

Phenomenon or Process|B2
- Human-caused Phenomenon or Process|B2.1
  - Environmental Effect of Humans|B2.1.1
    - Wounds and Injuries
      - Head Injury (Craniocerebral Trauma)
- Natural Phenomenon or Process|B2.2
  - Biologic Function|B2.2.1
    - Physiologic Function|B2.2.1.1
      - Organism Function|B2.2.1.1.1
        - Mental Process|B2.2.1.1.1.1
      - Organ or Tissue Function|B2.2.1.1.2
      - Cell Function|B2.2.1.1.3
      - Molecular Function|B2.2.1.1.4
        - Genetic Function|B2.2.1.1.4.1
    - Pathologic Function|B2.2.1.2
      - Disease or Syndrome|B2.2.1.2.1
        - Nervous System Disease
          - Neurodegenerative Diseases
            - Alzheimer Disease
          - Demyelinating Diseases
            - Demyelinating Autoamune Diseases
              - Multiple Sclerosis
        - Eye Disease
          - Ocular Hypertension
            - Glaucoma
      - Mental or Behavioral Dysfunction|B2.2.1.2.1.1
        - Mental Disorders Childhood
          - Child Development Disorders, Pervasive
            - Autism

–

It should also be noted that some diseases, such as Autism, are not only a challenge to define visually but medically. It may be that until the precise cause of the disorder is found, the visualization of that disease will be necessarily vague as well. This highlights the nature of true collaboration where synchronicity is found and developments in one field directly affect the other.

Actions

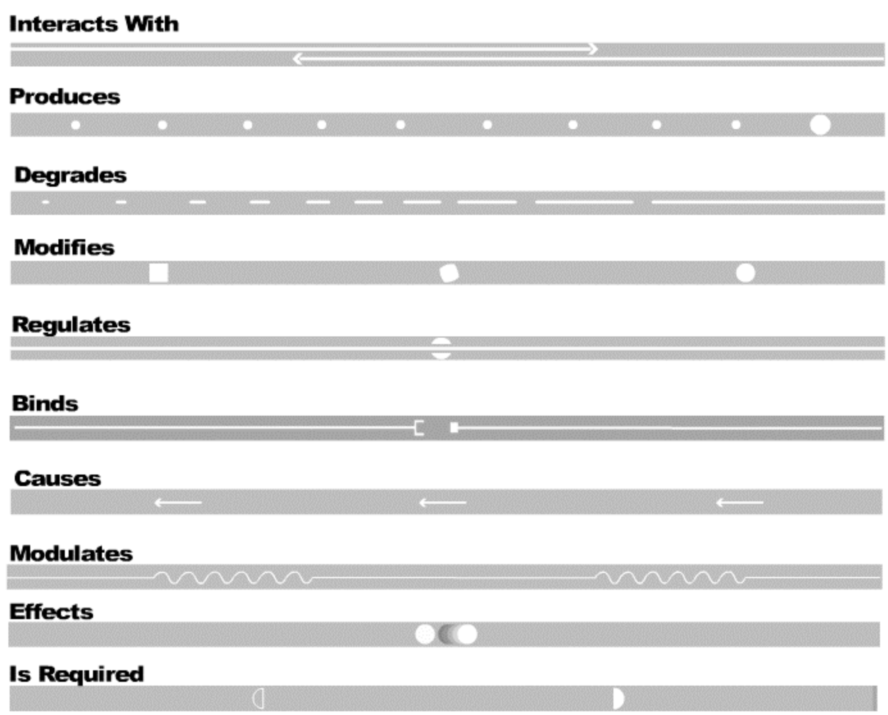
One of the more difficult concepts to represent is an action or state of being (Zender, 2006). This challenge has been approached from several directions, the most current shown below. As the authors developed the Proposition Statements they generally took the form "object A – does something relative to – object B." The "does something relative to" part of the statement is the Action object. The position of the Action object between Biological or Biochemical objects led to the concept of developing a connecting form for the Action icon. In this example the thick line contains a small graphic representation of an Action concept: bind, modulate, produce (Figure 10). The thick line is used to connect two or more Biological or Biochemical objects to represent a complete Proposition statement. Upon interaction with the line (roll-over in this case) the objects animate within the line to more explicitly depict the action. In the final display, these icons and all others have a tool-tip-like name that pops-up when the mouse hovers over the object for more than one half second.

Figure 10

Verb Connection

Most successful verb icon combined a static iconic image in a linear band that animated on roll-over. Movement added to clarity of meaning. More study cold be done with animated icons to enhance meaning.

–



–

DISPLAY

The last step in this project was to use the Visualization Systems described above to create visual equivalents of all 20 verbal Proposition Statements in a single visual Display Prototype. Two of the more than 20 Display Prototypes are shown below (Figure 11 and Figure 12). The Prototypes were designed to gauge user feedback in two areas: accuracy and speed of comprehension of the concept statements. Due to the limited timeframe for each project Prototype developed thus far, the focus was necessarily more on the effectiveness of the visualization system and less on the interactive approach.

Figure 11

Display Prototype

Student designers: Sean Gresens, David Kroner, Nolan Stover and Luke Woods

–

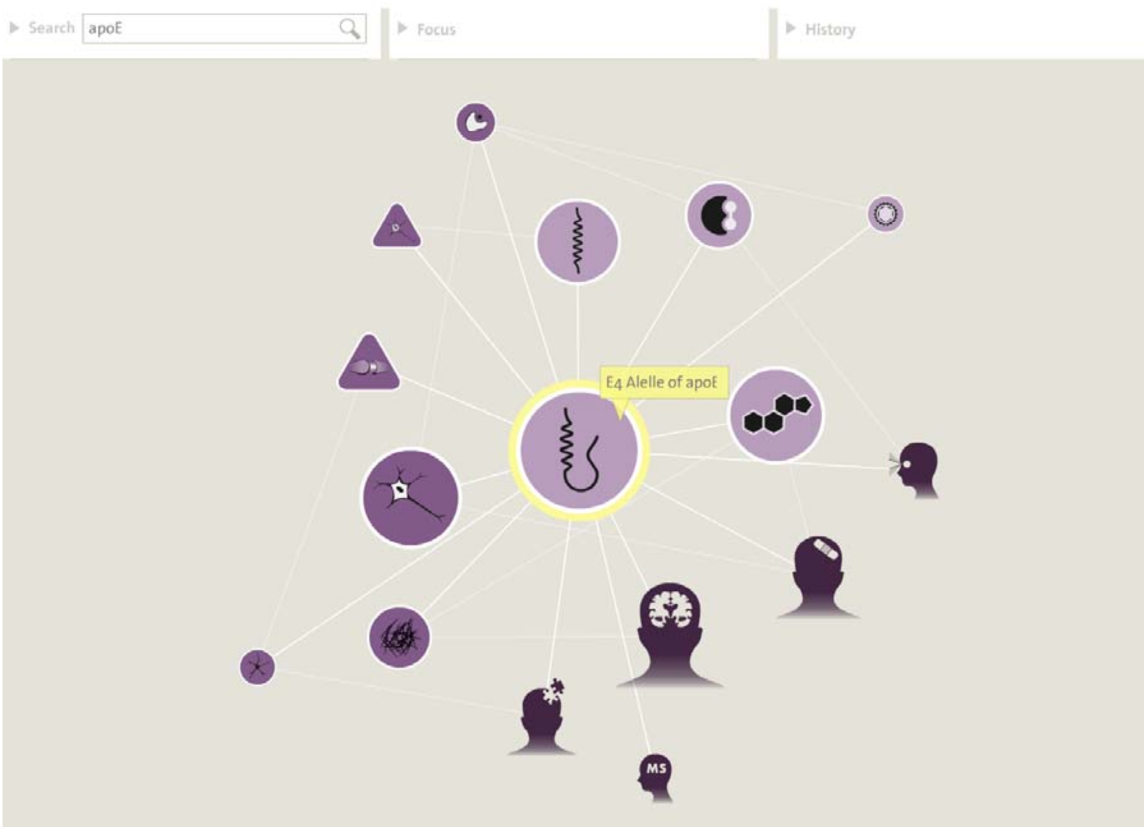
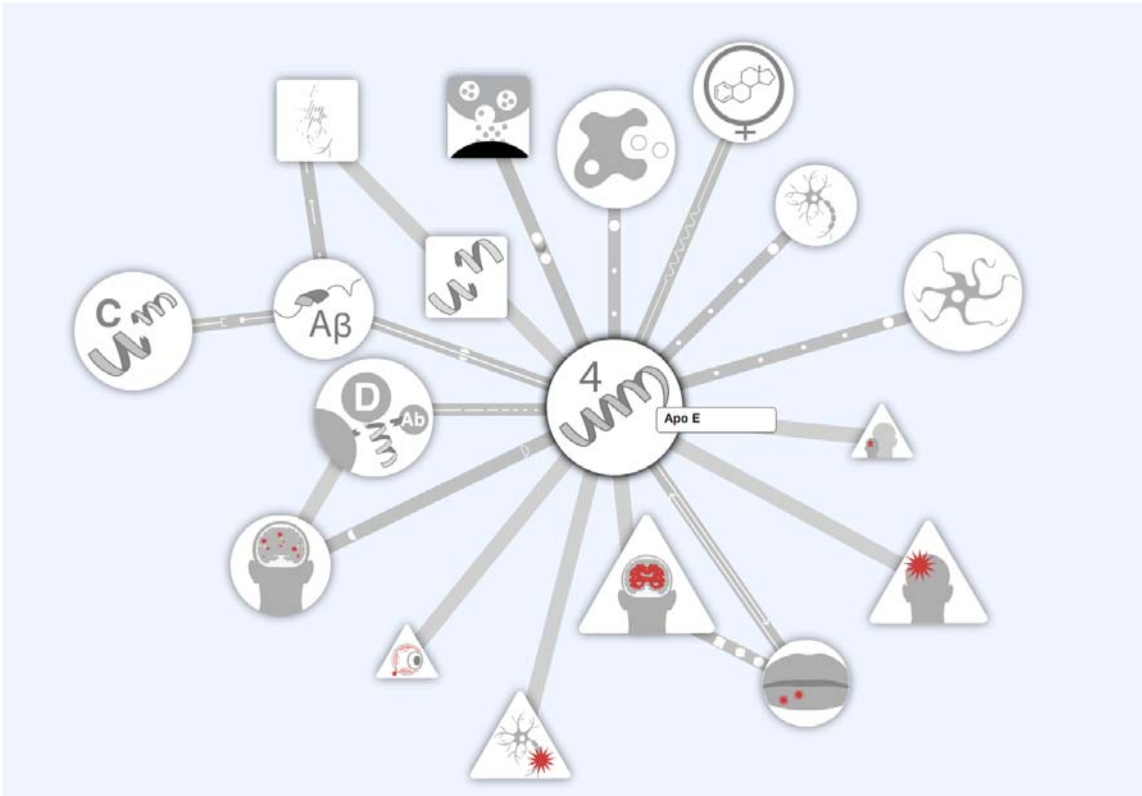


Figure 12

Display Prototype

Student designers: Ian Helm, Nathan Peters, Ryan McAllister

–



–

TESTING

The authors worked with the design teams throughout the project, one advancing design and guiding the application of design principles (**) while the other instructed and consulted with students on science (!!). During the development process several forms of evaluation were used.

EVALUATE ICONS INDIVIDUALLY

While both design and scientific feedback were provided daily throughout the design process, more formal sessions were held to give the scientist [!!] an opportunity to formally evaluate the communication effectiveness first of individual icons then of the Visualization Systems. The evaluations happened in person, via email and via video chat as necessitated by circumstances such as travel. These were qualitative sessions, much like traditional critiques, but with the authors serving as a collaborative team doing the evaluation simultaneously from design, comprehension and scientific accuracy perspectives.

EVALUATE DISPLAYS

Following the design of the Visualization Systems, student teams defined tasks and built prototypes with sufficient interactivity to perform the required tasks. From two to four subjects with domain knowledge in medicine and with varying degrees of knowledge of the subject area of Alzheimer's, but who had no familiarity with the visualization system, or the display, were set before each prototype and given the tasks in writing (see sample task list below, Figure 13). Students observed each subject as they worked through the task list, making observations, recording subject comments and numerically rating subject performance. Some subjects on initial exposure to the project prototype would take 2 – 3 minutes to orient themselves to the concept. After this orientation all subjects were able to 'read' the correct concepts proposition statements from the display. Subjects were encouraged to talk and share their train of thought as they worked. This gave the designers invaluable feedback on icons that were less successful and identified patterns of difficulty in comprehension. In most cases 'tool-tip' pop-up labels were supplied for the animations (as noted above). Overall the results were very promising, so much so that in some cases the student facilitators could not keep up with the test subjects as they rushed through the displays naming icons and reciting concept propositions correctly.

Figure 13

Sample Tasks

–

1. **Find and select the biological icon for ApoE**
2. **Find the disease icon for Multiple Sclerosis**
3. **Describe the correlation between ApoE & Multiple Sclerosis**
4. **Find 3 other diseases associated with ApoE**
5. **Find 1 object that has no correlation with ApoE.**
6. **Find more information on Alzheimer's Disease. Describe said info.**

–

EXPERIMENT DESIGN

Based on the initial success of the informal testing noted above, the authors with the support of a University of Cincinnati University Research Council grant, developed a more rigorous evaluation mechanism to gauge the success of the approach. The evaluation had two phases: first, an evaluation of icons, then a comparison of the icon-based display approach against a similar text-based display approach.

Evaluate Previous Icons

Over the life of the project, spanning five years, many individual icons and icon systems had been developed. To facilitate the development of a single Visualization System for formal evaluation, 4 expert reviewers rated the communication effectiveness of the icon for each concept using the following process (also see Figure 14 below):

Step 1: Collect icons, organize by concept and post on a secure internet site for expert evaluation.

Step 2: Develop evaluation mechanism for each icon category, a 1 – 5 rating scale with 5 being most understandable and 1 being least understandable. Allow for ties.

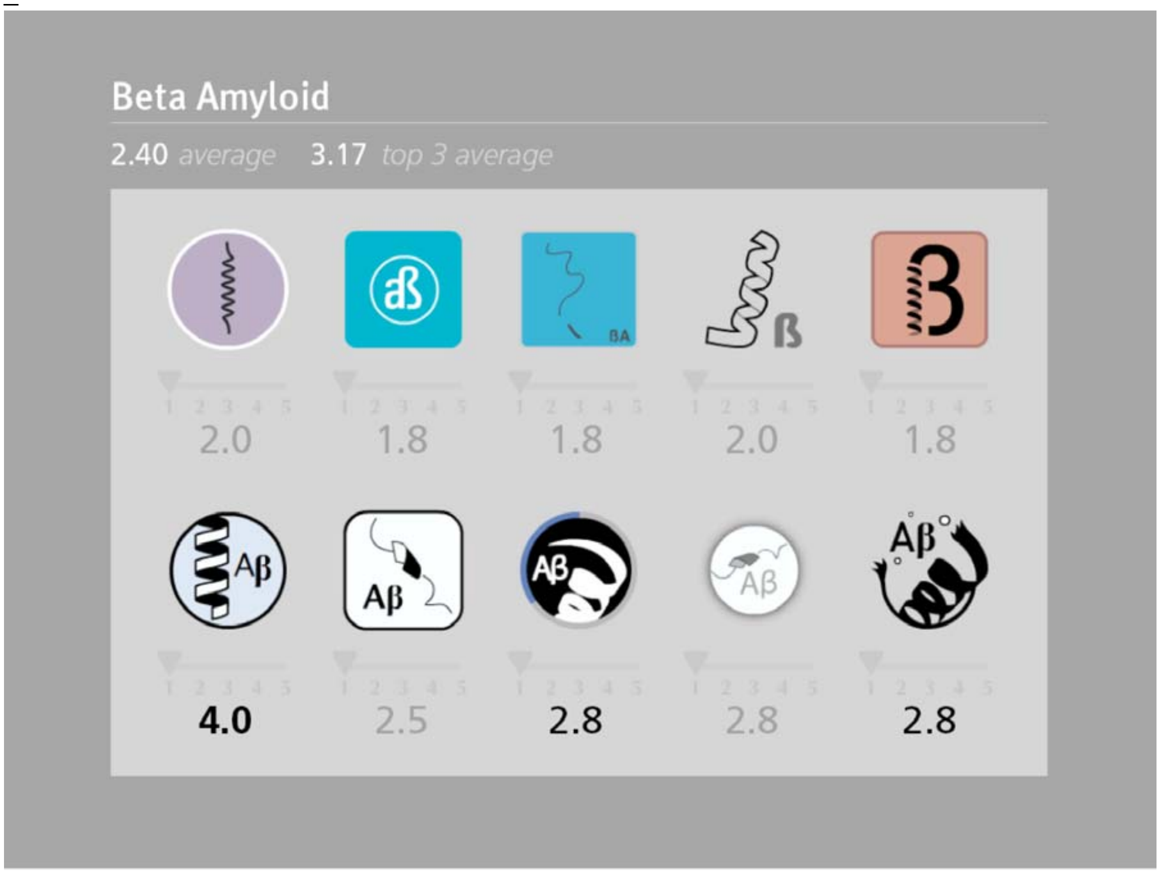
Step 3: Evaluate the icons by one designer and two domain experts, average rating scores, analyze results

Step 4: Redesign icons based on evaluation

Figure 14

Evaluation Mechanism and Results

One of 18 icon evaluation result report pages showing the ratings on a 1 – 5 scale, 5 being the most effective. The ghosted bar with numbers 1 2 3 4 5 is from the evaluation form.



Comparative Evaluation of Two Display Prototypes

Following the development of a consolidated Visualization System based on the input above, an experiment was designed to evaluate the effectiveness of an icon-based display made up of the selected icons (Figure 15). To guide the test an IRB protocol was developed with the following features:

The two displays identical in content (30 concepts consisting of 18 Semantic Types and 12 Semantic Relations in UMLS terms) but different in form, presented to different groups of subjects: a between–subjects design.

One display text-based, the other display icon-based (using an icon/glyph Visualization System).

Text-based display only words enclosed in rectangular boxes and linked by lines. Text-based display a variant of a link-node diagram used to display results of the existing Telemakus concept display system.

Icon-based display system of static and / or moving icons contained by glyphs and associated via graphic forms: shapes, lines or some combination thereof (a Visualization System).

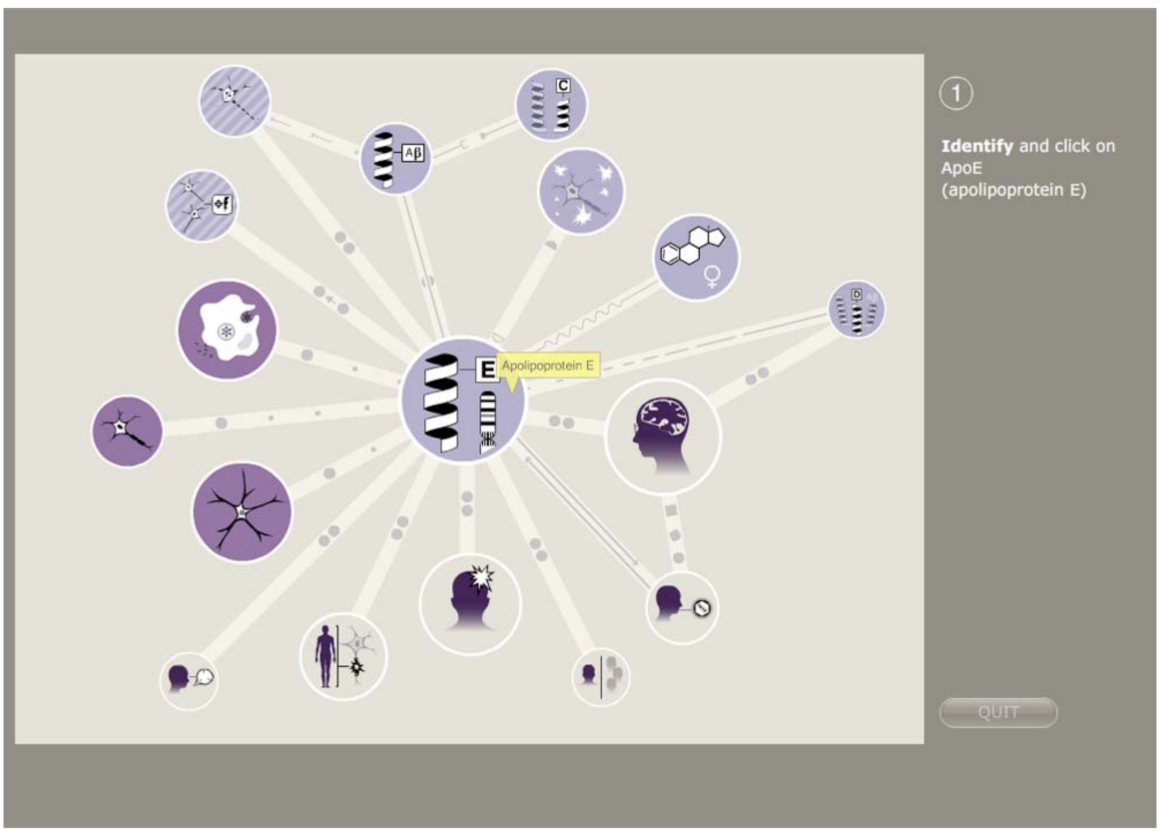
Two pools of 20 subjects, evenly balanced for training, experience and age, perform the same tasks with each display. Tasks designed to measure three effects:

1. speed of recognition of concepts
2. speed of identification of related concepts
3. speed of identification of the type of relationship between concepts

Figure 15a

The Icon-Based Display Test

–

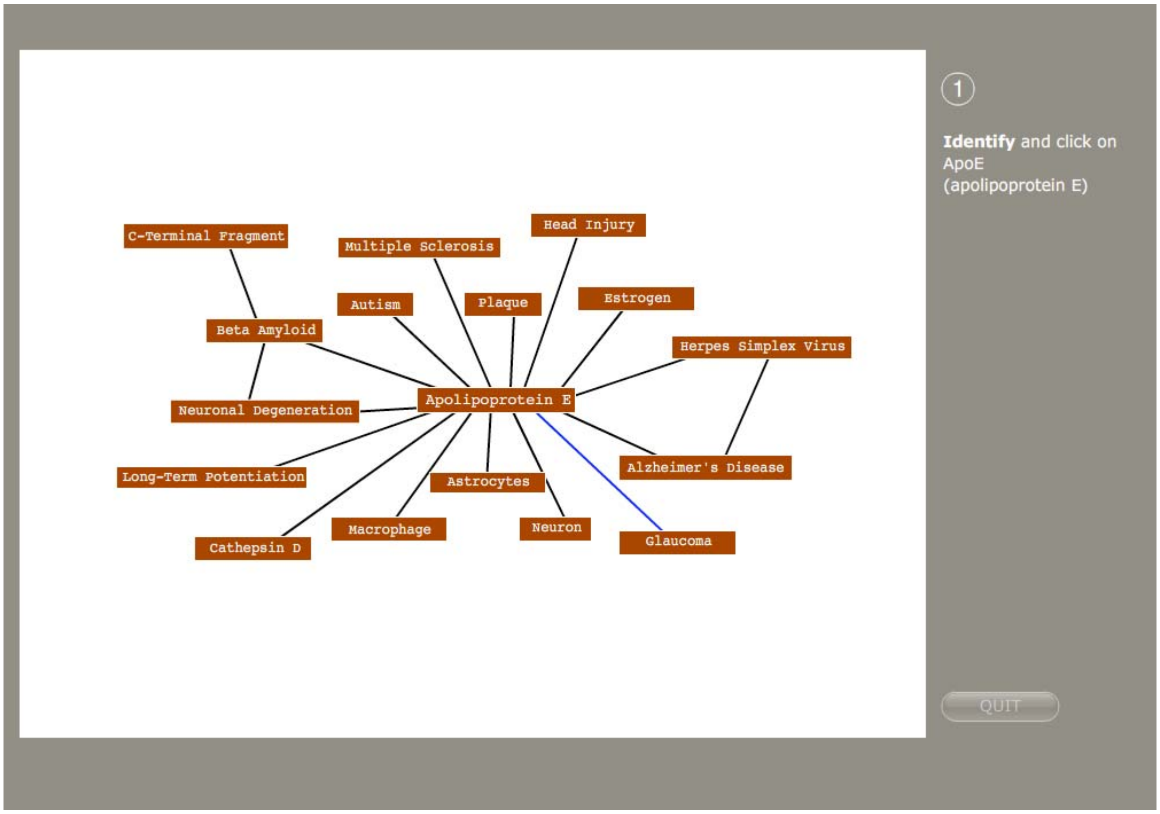


–

Figure 15b

The Text-Base Display Test

–



–

Figure 16

Task List

–

TASK LIST

1. Identify and click on ApoE (apolipoprotein E).
2. Identify and click on Alzheimer's disease.
3. Count the number of Diseases represented in the display. Select the correct number below:
4. Identify the relationship between ApoE and Alzheimer's Disease. Select the correct answer below:
5. Count the number of items that have the same type of relationship you identified between ApoE and Alzheimer's in question 4 (above). Select the correct number below:
6. Identify and click on the concept which the display implies is most closely related to ApoE.
7. Identify and click on the C-Terminal Fragment of ApoE.
8. Identify and click on two additional molecular level concepts (gene or proteins or organic chemicals) in the display.

–

The specific tasks for the two subject groups were identical (Figure 16). Some required simple recognition and identification (numbers 1. and 2.) while others requires interpretation and association (number 3.).

Quantitative and Qualitative Data

The results are being analyzed and compared using ANCOVA. Subjects are allowed as much time as needed to perform the tasks but time to perform and accuracy are measured. In addition to these objective measures, user impressions of their ability to find novel relationships in the display are being gathered using a questionnaire with a rating scale.

RESULTS

Testing was conducted in June 2007 with 27 subjects, 13 with the VL-based prototype and 14 with the text-based prototype. Subjects were a representative population of domain experts. Testing followed the experimental method and IRB protocol described above. Testing was accomplished on-line using a web-based testing tool (screen grabs of test shown in Figures 15a and 15b). There was no moderator intervention in the testing. A single page of copy provided informed consent and gave brief instructions. The test software automatically recorder time and accuracy and reported the results in aggregate and individually for each subject.

The results were promising. Overall, the icon-based display was both faster and more accurate. For simple identification tasks the two approaches were nearly equal in speed of identification (see tasks 1 and 2, Figure 16). For identification tasks requiring reasoning or association the icon-based display was overall 18% faster. For the last task, one requiring identification of similar concepts, the icon-based display was nearly twice as fast as the text-based display. Accuracy of the icon-based display was equal to the text-based display on simple identification tasks but far more accurate on tasks that required the recognition of relationships. On task 3, "Count the number of diseases in the display" the icon-based display was 4.43 times more accurate than the text-based display.

While the results thus far are promising, they are far from comprehensive. For example, one wonders how much of the gain in speed and relational association was due to graphic encoding of shape and color and how much was due to iconic picturing. How will the visual display perform when the quantity of objects increases into the hundreds? How much will interaction with the VL display stimulate discovery?

In addition to quantitative and qualitative results, many unusual visualization techniques have been developed and described above, including: object-based visual / verbal language concept, extensive use of graphic modifiers, systematic integration of glyphs and icons and integration of icons and motion to represent actions.

OBSERVATIONS ON COLLABORATION

While it was important to recognize and discuss differences in use of language, method and conceptual space to make collaboration work, the authors found that in discussing and experiencing the differences we became more aware of both strengths and limitations in our own fields. Designers became painfully aware of the limitations not only of our vocabulary but also the very limited understanding we have of visual form of which lack of vocabulary is but a symptom. Scientists, in turn, discovered that they have missed out on important tools for enhancing communication and managing the explosion of information. Improved awareness within one's own discipline may turn out to be one of the greatest benefits of collaboration.

Thus far, this project has focused on the speed and accuracy of comprehension of a visual display as a necessary first step. However, the most significant pay-off for such a display may lie in the ability to interact with it performing multiple related exploratory searches leading to insight and perhaps discovery. Complete interaction, along with other features of this project, continues to be explored.

After all, it is encouraging to have demonstrated that design can effectively communicate such complex scientific content. While the demonstration at this point is sparse and far from attempting a full display of all the concepts and relationships in a data rich domain such as medicine, there is hope based on evidence that such iconic visualization may be possible, and if the possibility is realized design may help usher in a whole new era of groundbreaking communication of medical information, advancing science and stimulating discovery. Collaboration has made it possible.

REFERENCES

Arnheim, Rudolf. 1969. *Visual Thinking.* Berkley: University of California Press.

DaFanti, Thomas A. Brown, Maxine and McCormick, Bruce H. 1987. Visualization, NSF Report, Computer Graphics / IEEE Computer

Lyman, Peter. Varian, Hal R. 2000. How Much Information, report, University of California at Berkeley, School of Information Management and Systems, 2000
(see also 2003 edition) excerpt from http://www.press.umich.edu/jep/06-02/lyman.html

NLM/NIH. NLM - National Library of Medicine, NIH – National Institutes of Health web site. http://www.ncbi.nlm.nih.gov/entrez/query/static/overview.html

Revere, Debra. 2003. "A New System to Support Knowledge Discovery: Telemakus," Proceedings of the American Society for Information Science and Technology Annual Meeting.

Tufte, Edward. 1983. *The Visual Display of Quantitative Information.* Cheshire, CT: Graphics Press.

Tufte, Edward. 1997. *Visual Explanations.* Cheshire, CT: Graphics Press.

Ware, Colin. 2004. *Information Visualization: Perception for Design*. Morgan Kaufman.

Wright, N. T. 1996. *Jesus and the Victory of God*. Minneapolis, MN. Fortress Press.

Zender, Mike. 2006. Icon Systems for Global Non Verbal Communication. Visible Language 40.2. 177-206

Zender, Mike. Crutcher, Dr. Keith. 2007. Visual Language for the Expression of Scientific Concepts. Visible Language 41.1. 23-49

Zender, Mike. 2007. Basic Design Research: Parameters of Visual Form. Proceedings of the IASDR 2007, Hong Kong.