

Interface and Mind

*A “paper lecture” about a domain-specific design methodology
based on contemporary mind science*

Version 05.3, April 26, 2009

Published in a special issue on Smart Graphics, in the journal
it – Information Technology, 51 (2009) 3, pp. 131-141

W. Bradford Paley

Columbia University Department of Computer Science;

Digital Image Design Incorporated

brad@didi.com • (917) 686-3274

Abstract

An emerging interaction design methodology is presented; it addresses the benefits of incorporating domain-specific knowledge into the interface. The methodology is informally described using three real-world examples, exposing how findings from Visual Perception, Psycholinguistics, and Cognitive Science can be applied to the design process. Tying the design process back to science can help us to understand why current standard-widget toolkits are suboptimal for many domains. It can also inform the development of tools and intellectual approaches valuable at various stages of design.

Within the world of interface design there’s a favorite rhetorical gesture of showing a picture of the Xerox Star computer (released in 1981, complete with windowing graphics display, mouse, icons, and pointer; the so-called WIMP interface) and asking people to guess what year it was designed. The joke is that people don’t even get the decade right—that’s how little interfaces have advanced since the early 80s. I believe a new kind of interface is just beginning to emerge: one based on revealing the nature of the task and subject directly, not interpreting it through computer interface guidelines. In a very real sense these interfaces get the computer “out of the way” and let people deal directly with the reality of, say, brokering a trade on Wall Street, or making a hypothesis about a piece of literature. A rich, idiosyncratic, domain-specific reality plays out in the minds of knowledge workers. They may never have actually *seen* it, but they definitely live it; and it can be extracted and illustrated by following a straightforward design methodology.

These new interfaces have their roots in some older computer programs and even pre-computer media because, almost paradoxically, before standardized interfaces were developed toolmakers could concentrate more on fitting how we think and work than

what the computer could do. While standardization of some operations in an interface is beneficial, I believe more things than necessary are standardized, losing the richness of expression that a specific field's domain of practice naturally develops; and the resulting interface doesn't take advantage of the flexibility that evolution built into the human mind. The hobbled expressive range of standardized widgets is compounded by the illusion that following standards makes a good interface. So most projects on Wall Street, for example, are handed directly to a development team who may be brilliant at software engineering but have no preparation for the "cognitive engineering" necessary to connect data on the screen to decision-making processes in an expert's mind.

While the new interfaces are often visually informed by illustration and fine art, their power comes directly from applying the sciences of mind, in a "knowledge-acquisition pipeline" that steps through Visual Perception to Cognitive Science and Human Memory, to Psycholinguistics, culminating in the kinds of cultural specifics studied by Sociology.¹ Interestingly, a strict application of these disciplines leads *away* from the mechanical-looking interfaces one gets from dutifully using the standard widgets in Apple's or Windows' toolbox, or on the Web. It can lead toward interfaces that reveal the nature of the subject being studied or manipulated—revealing it through the eyes of the experts: the culture that best understands it.

Science can directly support design. And design can directly invoke culture. A design might reveal so much of a subject, in such an unusual and transparent way, that it's even seen as that cultural capstone: capital-A "Art." One of my earlier attempts to harness this pipeline is called TextArc; it was recently shown in the Museum of Modern Art's *Design and the Elastic Mind* exhibition, won Grand Prize [non-interactive] at Tokyo's Media Arts Festival in 2002, the year it was released, and was written up in the New York Times Art section (it will also always be at textarc.org; try it yourself if you like). This is somewhat ironic because it was designed purely as a tool for Structuralist textual analysis, a discipline some find as dry as disciplines get.

TextArc shows how the earliest stages of the pipeline can be engaged. Later work, especially designs completed for the New York Stock Exchange and several Wall Street firms, engage farther and farther along that pipeline towards culture and reap greater rewards as a result—culminating in a 20:1 speedup of a key task on the NYSE trading floor. The speedup came from a design that revealed the mental models created by floor brokers in the culture of floor brokers, discovered by listening to broker language with a specific Psycholinguistics objective.

¹ "Pipeline" implies a linear, one-way process but the real mental mechanisms are definitively not linear and one-way. I present this as a simplification, retaining only the steps that help inform design from the perhaps hundreds of interlocking mechanisms understood to exist by contemporary mind-science researchers.

Let's start at the very beginning of that pipeline: *Sensation*², with photons hitting the retina to trigger what Visual Perception researchers call Early Vision processes. But first, let's look at TextArc and the essential first step of the design process before applying these ideas at all: what was the goal of the interface?

TextArc: the design brief & conceptual design

TextArc had a lofty goal: to make it as easy to get the gist of on-line text as it was to flip through a standard paper codex. The design brief for the yet-to-be-named tool was something like "Your boss gives you a book with no cover, index, or table of contents; and says you need to impress the client with the month of thinking you've been applying to it—oh, and by the way, the client is in the elevator." (I can't start a design without an explicit task statement, and if the brief is emotionally-charged, so much the better.)

It was clear that the tool needed to have some kind of overview, perhaps formed by counting and associating words. But the computational linguistics lists and word-soup displays I'd seen³ led me to no insights about real meaning in the text. At this point in my research the power went off in my SoHo offices, giving me the opportunity to put my feet up on the desk and reconsider the problem from a healthier perspective: that of someone who might use the tool: me, an (admittedly amateur) textual analyst.

I realized existing tools were missing something significant; they threw away exactly what an author actually does: place words in a specific order, in specific places in a text. So I started by imagining a technologically-unconstrained layout that might show only that: draw out the entire text in a very tall vertical galley, then draw each word (again) at its average position. (I was concentrating on words because they're a decent first-approximation to meaning, and can be identified without making any strong assumptions.) If I made more common words larger, that would at least show where each word appeared most, but they would all clump on top of one another. And worse, this approach respected the ordering of the words but completely hid the fine-

² Seven sections of this pipeline are distinct enough from a designer's point of view to be worth naming and addressing individually: *Sensation* (raw sensory input, e.g., light hitting the retina), *Distinction* (deriving sensory input, e.g., lines or texture gradients), *Segmentation* (using distinctions to separate objects), *Recognition* (matching an object to features or a prototype of similar known objects), *Interpretation* (knowing that an object represents something beyond itself), *Association* (grouping or relating objects to form more complex ideas), and finally—the real goal—*Comprehension* (situating the ideas in a context relevant to the task at hand). I suggest that a good interface design employs all sections, in a balanced way. What "balanced" means is determined by the abilities we have evolved as humans, their strengths and saturation points, and the demands of the task being supported. An account of pipeline sections and balance is beyond the scope of this introduction, but suffice it to say I almost never see an interface that effectively uses more than two or three of the seven.

³ Views created using techniques such as animated force-directed placement, multi-dimensional scaling, and self-organizing maps; unreadable though they are often lovely to look at because of the cell-like structure or sensuous animation dynamics

grained distribution of the words: what if a word were used a lot in the first and last chapters? Averaging would display it in the middle, the same place it would end up if it were evenly distributed throughout the text or used only in the middle chapter.

Color-coding, every programmer's knee-jerk reaction to conveying information, might help if I were only looking for a couple of words⁴. But this tool needed to show dozens, hundreds—or almost *all* of the thousands of words in the text if it were to be effective; I wanted an overview that gave the viewer access to virtually the entire text so their focused imagination would not be limited to the “important” issues defined by some hacked metric. So if each word were pulled away from the text galley, say to the right, and rays were drawn from the word to every line in which it appeared we'd see the full, possibly lumpy distribution of the word. But still only a couple of words could show these rays before overlaps made it unreadable. And how far to pull it? Some arbitrary scaling factor needed to be developed. It was time to consider technical constraints: constraints of both the display *and* the human visual system.

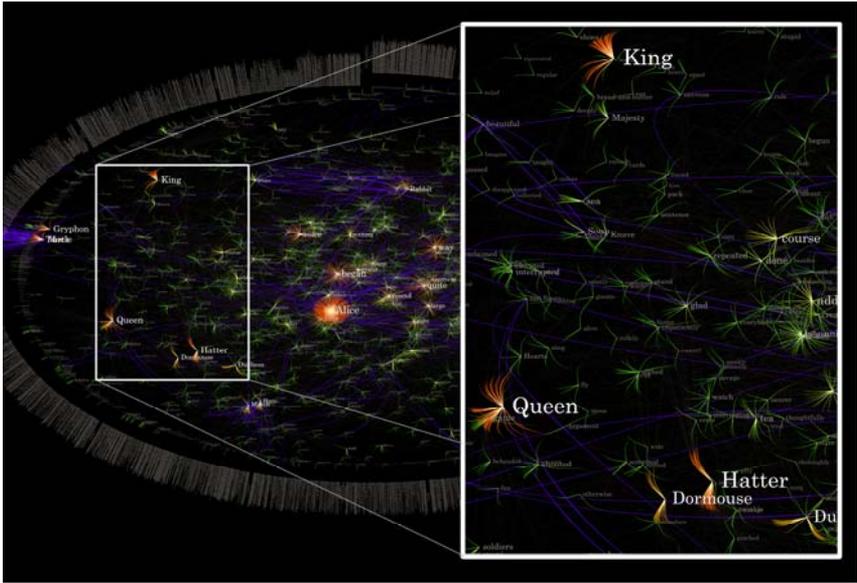
Fitting the concept to the screen & eye

The lights came back on a couple of hours later and I stared again at my small screen, wondering how to fit a whole text into such a small rectangle—and discovered all in a rush (it felt more like a discovery than an invention, it held so effortlessly together) that if I bent it around into a circle not only would any size text fit, but the averaging rule pulled words away from the perimeter automatically. I coded a quick trial and came up with a mess on my screen: words were indeed pulled away from the perimeter, only to layer on top of each other in an unreadable thicket of characters. It fit the screen, but not yet the eye & mind.

Recalling foggy nights in the Adirondacks (and early computer graphics depth-cueing techniques), I realized that the eye was very good at distinguishing different layers of gray. So I mapped the intensity of brightness of each word to how much it was used (a visual/spatial/conceptual metaphor, as we'll see later, bright = visually salient = close = important) and could immediately distinguish hundreds of words. Adding an “armed” state that told words to brighten when the cursor hovered over them gave me access to thousands, with just a scan of the cursor. Likewise, making the rays from the word to the lines in which it was used visible only for the armed word let me see its full distribution in a clear and unambiguous way.⁵

⁴ This might be a valid use of color: the Gestalt Psychologists found in the 30s that color was good at allowing people to associate separated objects into one group. But most current applications of color-coding are far less effective: some arbitrary meaning is attached to a color and people must constantly look back at a color key—if one is provided!—to remember what's meant.

⁵ Ironically, the print TextArc (reproduced here) is more easily scanned by the eye than the interactive version: in a very real way it's more interactive. The vastly higher resolution of print, more appropriately fitting human visual acuity, can feed hundreds of times more information to the eye than contemporary screens. The printed TextArc shrinks the rays that link each word to its



There are dozens of low-level perceptual tricks that one can use to layer information on the screen: from dark to bright, from gray to colorful, from small to large; and people are still doing research on things like which colors pop out more, or how faint a grid needs to be to stay in the background; the user interface and information visualization communities have been addressing near-retinal issues like this for over a decade. But light/dark, color/gray, line/area distinctions may only have been evolved by the human visual system to identify objects in natural scenes. Might there be a more natural way—a way cognitive scientists would call a more “ecologically valid⁶” way—to use these low-level visual attributes?

usages until the starburst is about the size of the word. This prevents them from overlapping yet retains all the distribution information: the eye is extremely sensitive to clustering lines and tiny variations in line angle. As a result the eye can quickly and effortlessly scan for distribution matches or anomalies many times faster in the print than the mouse can move the cursor in the interactive version. In the call-out you can see how “Hatter” and “Doormouse” have similar distributions, but the nearby “Duchess” appears in completely difference places in the text. Likewise “King” and “Queen” have similar distributions—except for the first four references: a visual trace of foreshadowing.

⁶ The trend toward ecological validity is a second-stage development in perceptual studies, where the researcher tries to set up an experiment that tests responses to situations that engage, e.g., the visual system, in the same way it might be engaged in the “real world;” in an ecology. So: measure how light a leaf needs to be to distinguish it from other leaves, not just how much you need to change the intensity of a spot of light in a darkened room to just barely notice the difference.

Information objects

I believe there is: we can draw our information on the screen not as tables or lists, but as objects that exhibit the perceptual richness that objects do in the real world. If we can then connect this rich perceptual variation to the most important channel of information needed for a given task, we might see hundredfold speed increases in task performance.

Let's imagine the reverse just to illustrate the principle: take an orchard of apple trees a month before harvest time, when only a few apples are ready to be picked. Start—like any good contemporary knowledge worker—by making a spreadsheet that reduces each apple to a row of attributes; with a column for average color, min & max redness, standard deviation of color variation around a mean, and perhaps size and x, y, z coordinates of each. Now choose: try to pick the early apples from that spreadsheet, or by looking at the orchard itself. Which do you think will be faster; which will be less error-prone?

If you picked the spreadsheet, you're an irredeemable quant and can stop reading this article now... Lists and tables have their uses as structured representations of structured data; metrics and procedures might be designed to pass through such data and do some tasks, quickly and well without human intervention. But then we don't need an interface. This design methodology was developed to support decision-making processes where the human being is an inescapable part of the process, either because algorithms have not yet been designed to replace them—or because the decision requires information or trade-offs that can't be programmed at all, as in political or value decisions. Even there, a list or table might be appropriate in the interface—but *only* if the person for whom we're designing naturally thinks of the domain as rows or columns of text or numbers (as in accounting). Everywhere else we can do better.

The relevant question for the next generation of interfaces is this: if current technology lets us manipulate every pixel on the screen well enough to show a photograph of an orchard, why are we still showing the raw numbers? The answer is twofold: first, not all tasks lend themselves to this kind of analysis. The human visual system is amazing at certain kinds of isolation, classification, and association tasks; but not all knowledge work needs these capabilities.⁷ And second: it's hard to develop meaningful mappings from data to image, hard to draw pictures that support specific tasks in useful ways.

⁷ Working computer models of visual perception are advancing at such a rate that I can foresee a day when information visualization tools relying on humans just for pattern recognition will be obsolete: just plug the visualization into the perception model and ask *it* what the patterns are. In fact, it's not just the human visual system we're talking about, but that of other animals, too—some are better at certain tasks than we are. So non-animal-rights advocates can go even farther: develop displays tuned for sharper-eyed animals and get *them* to find the patterns for us.

The information visualization community has hundreds or thousands of researchers, some of whom have been applying quite brilliant minds to this opportunity for two decades or more. But despite the lure of riches from Wall Street or fame on the Web, almost none of these tools have been adopted into widespread practice. I believe it's because they all focus on developing general-purpose solutions. In a rare alignment of interests general-purpose tools are sought by both industry (who want to apply the product to the largest market) and academia (where one gets a PhD for powerful generalization, not enumerating case studies). But they have several flaws when measured against the human perceptual/cognitive system: we aren't optimized to process data in the abstract⁸; nor are we able to easily transfer mental skills from one domain to another.⁹

So if it's hard to read an abstract graphic and even if you can read it it's hard to apply to what you need to do, how can design help? How can we tap into the deep processing capacity that let us effortlessly pick apples? We can design applying two distinct practices, supported by two mind sciences. We can make it easy to see and identify objects (studied in the overlap of two fields: Visual Perception and Cognitive Science), then make it easy to connect those visual objects to the mental objects people already use (which we find out by applying Psycholinguistics). And both are much easier—perhaps only optimal—when we have the luxury of restricting our focus to a specific domain of practice.

Concrete graphics

The human visual system evolved to identify real-world things, not abstract glyphs like numbers or letters, or abstract charting elements like bar charts or line graphs.¹⁰ Things in the real world have many properties that are rarely used to convey information in computer interfaces. They have shadows, textures, subtle variations of color, vastly varying sizes, different outlines, and different behaviors. We are exquisitely tuned to extract meaning from almost invisible variations in these properties: Ripe or rotten? Thinking or angry? Dangerous or bluffing? Object or background?

⁸ Contrast the five-million-year evolution of the generalization-capable human forebrain with the 250-million-year evolution of the mammal that carries it. What happened for those first 246 million years?

⁹ In perhaps my grossest generalization ever, I can summarize in two words the research results presented in Singley & Anderson's excellent book *The Transfer of Cognitive Skill*⁹¹: it doesn't.

¹⁰ A remarkable and remarkably readable reference into the Vision Science literature was compiled by Steven. E. Palmer;⁸¹ it has overviews of research from hundreds of scientists focusing on vision, and can be skimmed for an understanding that's usually fine for an engineer's or designer's needs. Of course he also cites the work by the original authors themselves in case deeper understanding is desired.

NYSE e-Broker version 7.01
 Mailboxes: Buy 5,0 IBM 79.25 M
 Cancel Adm Msg Looks Multiple Stocks
 Menu Configure Help Create new CAP Book Quote

BAC PRX FCS
 43,0 GE 31.11 more behind GS gs
 5,0 GE 31.08 GTC 9c FBC adc
 8,0 IBM MKT CLO 2s ML mlb
 50,0 IBM 79.28 go along 2s M m
 5,0 IBM 79.25 2s M m P
 14,5 IBM 79.20 2s M m
 120,0 IBM 79.00 2s M m
 GM JCP KBH LH IMU NEW PRB PSS RSH USB WMT XRX ZLC

Cancel orders at post
 Cross Trades
 Leaves Type Price point
 2,9 q 5,8 q 31.11
 2,5 q 5,0 q 31.10
 10,0 • 10,0 • 31.08
 Annual Reports

ANF FJC PRA
 BS bs 7,5 GE 31.26 GTX 9c
 MGS mgs 58,0 GE 31.28 SSE OC 9c
 10,0 GE MKT 9c
 5,0 IBM 31.50 2s
 24,0 IBM 31.50 2s
 7,0 IBM MKT 2s

43,0 • 43,0 • 43,0 q 43,0 ioc
 100% Lvs. c All of it
 50% Lvs. c
 25% Lvs. c 1/4
 2,0 •
 5,0 •
 10,0 • 10,0 • ioc

GE High 32,85 Low 30,22 +0,11 (0,33%)
 Hit 31.11 31.26 Take
 20,0 90,4

DECK	ACTIVE	BIDS	PRICE	OFFERS	ACTIVE	DECK
			31.66	424,4	23,9	
			31.65	400,5	12,6	
			47	287,9	15,6	
			33	272,3	7,6	
			28	264,7	25,1	
			27	239,6	94,7	15,0 q
			26	124,9	34,5	5,0 •
			31.25			
			MKT			
			31.12			
		5,0 d 15,8 s	20,0	.11		
		5,0 q	64,0 84,0	10		
				09		
		10,0	65,0 149,0	08		
		25,0 c	6,0 155,0	07		
			24,0 424,8	30.84		
			40,0 464,8	30.83		

Commit → ← Cancel Cancel all Cancel → ← Commit

By drawing objects with attributes that fit the way our visual processes work we can more clearly distinguish different kinds of information. Take this design for a handheld device, to be used on the floor of the New York Stock Exchange. There are two major entities in play here, an accurate reflection of the two main types of orders driving trading activity on the floor: orders from clients to the broker (You or me to a broker: “Buy me 25,000 shares of IBM”) and “post orders” that a broker makes active in the market by getting them to the “post”: the place on the floor where a floor “specialist” is charged with maintaining a liquid market for a given stock (Broker to specialist: “Find me 5,000 shares of IBM; there may be more to follow”). Client and post orders fulfill a completely different role in the trading process; they’re as different as bulk grain and puffed what cereal—shouldn’t they look different?

They’re drawn differently to help brokers distinguish them without having to think about it: user interface jargon calls this “pre-attentive” distinction; we actually make the distinction before it comes to our conscious attention. The time frames in which these responses happen are small, but not insignificant. We’re all living in the very recent past, to some extent: consciousness lags behind the real world some 300ms, about a third of a second; enough time to get into a car accident if you’re tailgating. But researchers studying rapid visual categorization find that objects, and even entire scenes, can be recognized in less than a tenth of a second. This is remarkable—almost magical—because it means we can design an interface to get people to react to something *before they are consciously aware that it exists*. It’s the realm of martial arts masters, and may invoke the same fundamental cortical processes¹¹

Shaving two tenths of a second off someone’s performance might help save you in a swordfight, maybe even get you a better price on the Exchange floor, but these pre-attentive distinctions offer value over longer time scales, too. Many can be performed at once, allowing people to choose the most important order, for instance, to address next because it’s a different color than the rest (in the illustration new, unread client orders are black because they’re so important.) Some search tasks can be changed from “linear search” (which requires scanning through every object and takes longer as more objects are added) to “pop-out” tasks (in which the object sought is identified instantly).

And the very name pre-attentive evokes another key benefit, perhaps the most important for complex tasks: conscious thought it not distracted, even for a moment, to acquire whatever information the design encodes in these concrete attributes. Experts can not only be freed from “dealing with the system”; they don’t even have to

¹¹ Some knowledge work decisions need conscious input for a proper resolution; this is the same in martial arts. Picture the usual movie scene in which a martial arts apprentice is startled from behind by his master: he turns, weapon ready, but does not strike. The master is testing him—or serving soup. If we are aware that a decision needs conscious consideration, we can still prime the mind with all the relevant basic data (using pre-attentively readable visual attributes such as shape, color, size, texture, etc.) but stop short of dictating the outcome.

deal with getting their data. Contrast this with how almost every computer system on Wall Street (and other knowledge work domains) tries to feed information to people: in tables of numbers or text. Not only does this require attention, it requires a *lot* of attention; reading in general imposes what cognitive scientists call a “high cognitive load” and reading disjointed numbers may be even worse than reading ordinary text. But to get past this crippling bottleneck, we have to design properly: how can we simultaneously help the eye group similar objects, differentiate between groups, and present variations within groups that can feed information into people’s decision-making processes?

Visual vocabularies

I’ll argue that the best way to feed information to people is to fully and precisely engage their existing visual/cognitive mechanisms: fit the eye and mind as closely as possible and you can get more information in with less uncertainty. And one of the best ways we can understand how to feed those mechanisms is to look at what they evolved to deal with. Our eyes co-evolved with plants and other animals as the “information objects” in scenes supported by rocks and dirt, and lit by the sun and sky. We had to instantly distinguish between food and rock, friend and foe, rival and possible mate using whatever structure we could glean from the light rays bouncing toward our heads. What can we learn from the variations in structure that the natural world presented us?

When we see a snake, for instance, we engage two possible cognitive mechanisms to recognize: we compare it to a prototype or we see if it has the features of snakes. Which mechanism is the true way the brain does the processing is the subject of considerable disagreement among researchers, and may vary depending on the object or task. Fortunately designers need not settle the argument: we can draw objects using teachings from each side; thinking of them as variations on a master object or developed from a list of visual attributes. I’ve found the latter to be very flexible, so I’ve adopted the term “visual vocabulary” from the design community to suggest how we can learn from the natural world to create these visual distinctions.

“Vocabulary” implies a collection of words. What visual “words” (visual attributes) might define a snake—keeping in mind that to define something is to distinguish or hold it in contrast to other things. A snake has sinuous, continuous curves while the plants that might surround it have interrupted ones. A snake may have a different color range. Both snake and leaves may be shiny, but to different degrees; and the snake has a very regular and repetitive pattern to the shiny spots, while the shines on leaves may more linearly follow the veins and leaf structure.¹²

¹² The power of groups of visual attributes like this can be seen in the ability of people to detect threats in the environment more quickly than non-threatening objects.^[6]

Notice how these attributes group together to describe a snake. We'd like to develop similarly cohesive vocabularies of attributes to direct our drawing efforts: make a broker's order types as distinct as a snake and stick. Notice also how they leave a huge amount of room for variation; we can use this variability to express the variability in the data relevant to the given task. This design defines one high-level visual vocabulary for orders in general: well-bounded, outlined, non-cornered rectangles that have roughly a 3:1 aspect ratio. Within that general category two more specific vocabularies distinguish client orders from post orders: client orders are larger, have bright colors, have more text on them, and have rounded corners; post orders are smaller, dark brown or dark blue, and have sharply clipped corners. Just as in the real world, visual features can form hierarchical relationships: curves may define both leaves and snakes as objects distinct from the ground underneath, but adding more features splits those objects into two classes. Finally, we still have a huge amount of variability within the client order vocabulary to express order characteristics like buy vs. sell orders, or states like highlighted (yellow, textured) orders or new (black) orders.¹³

Building an ecology of information objects in this way allows interface designers a further luxury: rather than relying on just a few color-coded layers to mix different types of data, dozens or scores of different kinds of objects can be easily distinguished if the right visual vocabularies are developed. We have no feeling of "information overload" in the fruit section of the supermarket; and would have little problem sorting out the fruit if they were mixed—and much less of a problem distinguishing fruits from fruit baskets, signs, carts, people, walls, floor, and dozens of other categories of things visible in the same arena. We have been evolved for exactly this task of distinction and comparison; the task of the next generation of designers is to create a similar perceptual richness—with the goal of making it easy for people to interpret this richness as useful in the knowledge-work tasks they need to accomplish.¹⁴

Vision for meaning

Why do we see at all? Perhaps considering this question can give us clues about how to feed the eyes properly. Eyes evolved for the same reason anything evolves: because they helped the living being survive long enough to reproduce—to be selected by the environment to allow the genes that produced them into the next generation. They helped animals evaluate their surroundings well enough to compete well in the three

¹³ It helps to understand how we process colors, glyphs, and lines at a fine-grained level when designing these visual vocabularies. An exhaustive and scholarly application of Vision and Cognitive Science to the domain of cartography, Alan MacEachran's *How Maps Work*,^[4] can be a great inspiration.

¹⁴ My ecology of information objects shares an approach to the small, focused domains of practice (and therefore well-distinguished and domain-specific characteristics of information flows) studied by Bonnie Nardi^[5] and Edwin Hutchins.^[6]

Fs of evolution: feeding, fighting and, um, reproducing. Since we know that the human animal recognizes things by features, and we generalize sets of lower level features into higher level conceptual categories, what categories might be most useful to confer a selective advantage on an individual? We can look to the top: the organization of concepts in the mind, on the assumption that the mind also conferred a selective advantage and retains some structure that reveals what helped us survive.

George Osgood used language as a window into the structure of the mind, testing how people place concepts (such as LADY, BOULDER, SIN) on dozens of scales (such as young–old, wet–dry, pungent–bland) as probes into how we organize ideas.^[7] The rankings revealed how fundamental each scale was to how we categorize concepts; the three most fundamental factors represent the eigenvectors through meaning space: the scales that best describe how we perceive things. The three factors; *evaluation*, *potential*, and *activity*; described almost 97% of the common variance in the data: 97% of how people made their choices could be described by these three scales. Put in a three Fs evolutionary context these scales make intuitive sense: when we first see an object we want to rank it along a good–bad scale (*evaluation*, describing 68.55% of the covariance), does it serve our goals or not? We also take into consideration its *potential* (strong–weak, 15.46%) and its *activity* (active–passive, 12.66%). The next factor in line (associated with rugged–delicate, beautiful–ugly, and shallow–deep) accounted for only 3.08% of the covariance. If something’s bad, strong, and active does it really matter whether it’s shallow or deep?

I see interface design as the practice of helping people get information relevant to their current task by perceiving things on the screen and acting upon them naturally. We don’t just want to call someone’s attention to something—we want to help them determine whether it’s good or bad for what they’re trying to do, how much it will help them (strong or weak, so they can prioritize among different sub-tasks), and how likely it is to change. Each of these fundamental properties can be inferred from visual cues in the real world. Thinking about it in this way can help a designer decide what to display when creating a specific knowledge-work tool and how to draw it.

Simple-minded consistency

The goal is not just to get someone’s eyes pointed at some place on the screen (which we can think of as being driven by perception) or help them visually categorize and differentiate objects (early cognition), but drive someone’s progress by giving them something meaningful. And meaning in an interface is evaluated with respect to a specific goal. And that goal is *not* to accomplish the chore of some dry “information management,” but to do something meaningful in a living, breathing, perceptually embellished intellectual domain. So why are we reducing every area of human endeavor to data-manipulation tasks that must be done with one combination or another of standardized user interface widgets?

The rallying cry of “consistency above all” is often manifested as a slavish adherence to written interface guidelines and xenophobic dismissal of non-standard widgets. This was an important “training wheels” stage in the development of computer interfaces; admittedly much better than having to learn a different set of keystrokes for the same function in a dozen different programs. And it still serves a useful purpose for a certain class of interfaces: if a general-purpose interface is needed, or if the interface deals with new tasks where people have absolutely no prior experience to draw upon.

I’m not suggesting that we re-invent interfaces for specific computer-related tasks, like CTRL-S for *save* and CTRL-Z for *undo* on a PC. But rules that work on one level don’t necessarily work on a higher conceptual level, and the real world provides an example that we are well equipped to deal with a great variety of actions that might map to the same concept in a UI tool kit. Imagine living in a world where to open anything—a car door, a wrapped sandwich, a book, a water faucet, or someone’s abdomen for surgery—you simply and consistently tapped it twice. We use hundreds or thousands of opening strategies in everyday life, and though training for this kind of life takes a decade or so we seem to be built for it.

I *will* suggest that not only are we built for it in the real world, we can learn why it works there and apply the same principles to interaction design. I’ll also suggest that *open* means something different in each of these contexts and it’s actually useful to have different behaviors to implement the different things we want to do. The old saw “variety is the spice of life” might even indicate that we find some kind of delight in varying approaches to similar tasks. I think it also helps up to avoid what Cognitive Science researcher Gary Marcus calls “interference” in human memory: we don’t remember where we parked the car because we have semi-automatically done the same thing hundreds of times and they all blur together in memory; the actions share the same features. At least we don’t try to park the car in the same way or place we park a bicycle; they have distinct enough visual/behavioral signatures that we distinguish them effortlessly.

Facilitated interpretation

We do need consistency, but a consistency of a higher order: the interface should be consistent with our internal experience of the task itself. If we can tap directly into how someone thinks about their problem domain we sidestep an entire inefficient and effortful two-sided translation. We no longer have to desiccate a culturally rich organically structured domain of practice into some amateur ontologist’s Procrustean hierarchy. We can just draw what people are thinking.

And they'll know what to do—not because they've mastered a widgetized codification of a business analyst's guess at what an expert is thinking, but because they know their field.¹⁵

While many of the ideas we've considered in perception and early cognition could be applied to making general-purpose widgets more effortless (how about making the highlight color for a desktop the brightest or most saturated color on the screen?) what we'll look at now is best considered in a more restricted scope on interaction design: designs for experts who already know how to do something in a specific domain of practice. I use the word "expert" in a generous way: we're all experts in the domain of practice of getting into a car. It's not automatic: conscious decisions sometimes need to be made as we do it. But many of the actions are automatized: we grab for a half-dozen different handle types without thinking about it, and we settle back into seats of widely varying height & tilt.

The process of becoming an expert in any domain is partially the process of developing a flexible array of automatized "knee-jerk" reactions to common situations. But human (animal) memory is better at differentiating situations if they're accompanied by perceptual information that has a consistent relationship to the nature of the situation.

To facilitate the interpretation of things in an interface we can draw them to show some of the characteristics of the things they represent. This is not "chart junk" when restraint is used. I use an upside-down version of the principle of Ockham's razor: introduce only the features that will help someone familiar with the field recognize what they're looking at.¹⁶ Not gratuitous graphics; simple reference to (and respect for!) the unique cultural history that accompanies any intellectual domain. If we have the luxury of designing for a well-developed domain of practice we can draw from it to directly support the expertise, rather than wasting development time translating the domain into widgets, and worse: wasting the cognitive effort of (possibly millions) of people trying to figure out how to do a task they already know how to do.

¹⁵ A somewhat out-of-fashion mode of literary analysis known as Hermeneutics developed the concept of "projected introspection," useful here. If I'm reading something you wrote and I don't completely understand it—but I recognize you as a kindred thinker—I can think about what I would have meant had I written those words. If a system exhibits some cultural, even idiosyncratic, parts of its field it helps people identify with the interface's author and do this sense-making trick. (Hermeneutics is out of fashion for exactly the reason it's useful to us: it assumes the writer intends to communicate and the reader intends to extract what the writer wanted to communicate; *très* pre-Postmodern.)

¹⁶ If anyone can connect the real entities in their field with the undifferentiated rows in the ubiquitous scrolling table I'm very impressed. It does happen, but because they develop an association between low-level visual attributes (number and widths of columns, unusual features in a row in a column) and what they care about. I suggest a stronger and more effortless connection can be made if we design higher-level features into the entity as it is drawn on the screen.

Two approaches to connecting screen objects to thought objects

There are two ways to make interpretation easier: copy the objects used in the domain of practice to the screen, or invent new visual representations.

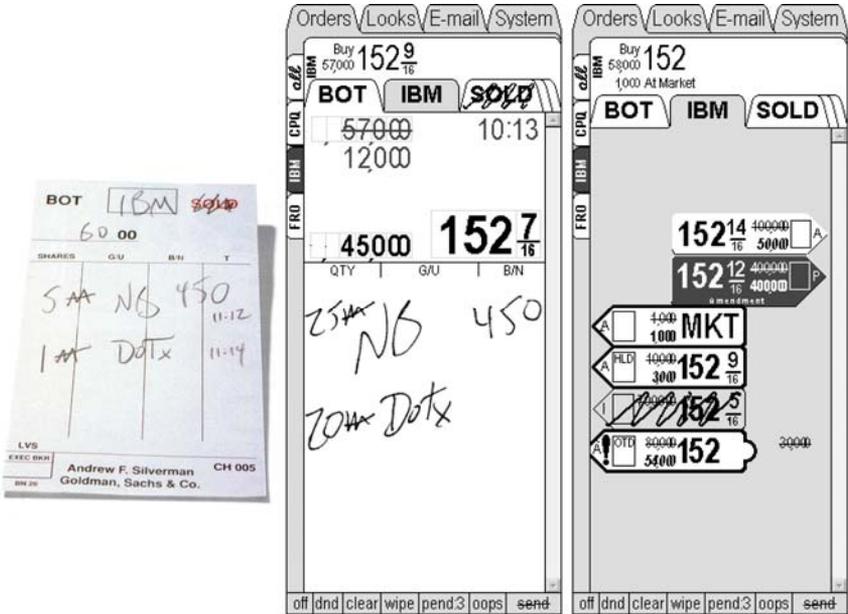
The founders of the financial software company Intuit paid careful attention not just to the ideas involved in bookkeeping, but the *embodied* ideas: the ideas as they looked and worked in real life. Scott Cook did a simple, brilliant thing for Kwik-Chek when he designed it in 1982. He copied both the look of a checkbook and the structural layout of the printed check register directly to the screen; not an easy thing to do on the text-based monochromatic green monitor of IBM's first PC. Today's Quicken is the flagship of a fleet of programs that in my mind deserved to take over that industry; they did careful research and clever design, but more important they respected the deeper roots of the task and made it simpler for people by (ironically) putting *more* on the screen. Kwik-Chek wasn't an abstract number-manipulation task; it had the additional graphics to look like a checkbook. Computer games are another very successful example of a design approach that works because it copies interaction and display modes from the natural world.

The second way of easing interpretation is slightly trickier, but more satisfying to me and potentially more valuable in a world turning increasingly toward a knowledge work service-based economy. We can invent new representations that *align with how people think of their specific tasks, in their specific domain*. Sometimes simple copying doesn't work—but having visual distinctions can still help people develop domain- and task-specific automatic reactions to situations. I emphasize *aligning with specifics* to distinguish this kind of invention from inventing new general-purpose representations. We now have a deeper way of understanding the only partial success of information visualization; it may cost too much effort to attach the real-world problems to the abstract images. And even where the new representations support useful insights, if we use the same tool for different tasks there may be too high a cost for assigning the insights back to the right task. If we make handles and visuals that are identical for cars, bikes, planes, and boats, we shouldn't be surprised if they all blur together in memory and people occasionally park their cars just past the end of the dock.

An earlier stock trading example

An earlier design for a Broker's handheld provides a good example of each method of tapping into the very specific domain of trading on the floor of the New York Stock Exchange. There were two primary categories of entities the brokers dealt with: orders

and executions. At the time executions were mostly done one at a time,¹⁷ so copying the existing execution ticket to the screen let people instantly recognize what they needed to do. Importantly, it also allowed the design to piggyback on decades of ad-hoc design by the practitioners themselves: every time the forms were reprinted during the two hundred years they were in use there was an opportunity for expert brokers to add a field, move a line, change a dimension (the pad fit perfectly into people's palms); or modify spelling, font size, or color. This physical artifact was the end result of an evolution of intentional design tweaks that were selected to remain in the design by their fitness for the niche on the Exchange floor.¹⁸



But orders needed to be dealt with all at once, to sort and prioritize, and a broker might be carrying a dozen or more tickets about the same size as the full-screen execution ticket. This is where the idea of information objects was developed: I determined the defining characteristics of an order by interviewing, observing, and mimicking the actions of the brokers. Then I drew objects that encoded the key task-specific issues as shape (buy and sell orders point in opposite directions) and position (orders with higher limit prices are higher on the page); attributes our visual system acquires for us pre-attentively. The order objects were then decorated with text and glyphic details that were important to the business task, but not necessary for the

¹⁷ This 1999 design pre-dated the hundred-fold increase in trading volume due to electronic trade entry and algorithmic trading. More details can be found in the June 2000 ID Magazine feature article on the completely deployed project.

¹⁸ A compromise between random-mutation-driven Darwinian Evolution and assume-perfect-foresight Intelligent Design that I'm sure satisfies neither side...

initial critical task of instantly deciding which order to act upon. All other information was hidden a stylus-tap away; though useful parts of that information was revealed in the three-letter “detail page” thumbnail that could display up to three three-letter codes. Notice that there was no need to translate the textual information into graphical form since the decisions it supported were consciously analytical and numerical in nature.

A happy accident strongly enhanced the design, fixing one of the worst problems caused by the initial introduction of computers. When orders were paper tickets they originated in the home booths around the edge of the Exchange floor and were handed to brokers when they passed. This physical token-passing was ideally suited to the sensitive financial nature of the information being passed: Both the originating clerk and the broker knew who had responsibility for it: whoever was holding it. And it clearly alerted the broker that he had a new order to deal with. Wireless handhelds broke this interlock: when an order magically appeared in the broker’s handheld he had to go looking for it in the list of orders; not something people are naturally good at. So they sorted orders by time in the tabular list of orders: most recent at the top.

I learned that the second-most important feature of an order was whether it was a new one (for which the broker accepted responsibility), so new orders were colored with the most salient visual attributes this eight-level grayscale device could offer: a very dark gray pseudo-3D object on a light field that was usually populated with only white/gray (accepted/fully-executed order) objects. It was so hard to miss new orders that we could sort them vertically by the third most important attribute, price. That’s where the happy accident occurred: when we sorted them by price we recapitulated on the screen the way they were arranged in the broker’s mind. And because the buys and sells were so clearly distinguishable the boundary between them also became pre-attentively available—telling the brokers exactly which orders were closest to the current market price (“in the market”: act on those) and which were farthest from the market (and so could be ignored)!

Since then this kind of “accident” has occurred often enough that I’ve realized it isn’t an accident at all. If as designers we are respectful enough of the way experts mentally represent their domains, and accurate enough in transcribing that to the screen, the design shows not only the things we intentionally transcribed—but also meaning in the relationships among the parts. Both the representation’s parts and the domain’s parts. Draw the objects well, and you illustrate the domain’s relationships too.

We get these gifts for free, in a sense, by trusting that the experts have an internal representation that helps them do their job well. But how do we extract these internal representations to guide our designs? We just listen. Literally listen.

Metaphors revealed in jargon

For my next clients I mentioned that I might (maybe, possibly) be able to apply some of those basic findings to the project, even though the design was incredibly specific; and I was able to use information objects again. This happened once or twice more and though I was pleased and it seemed like more than luck, I couldn't completely count on it working every time. Then I found Psycholinguist George Lakoff's research into metaphors as a basis for thought.¹⁹ I interpreted his ideas to understand why I might theoretically expect to be able to apply these idea-physicalization techniques to any knowledge domain—even (*especially!*) the most abstract ones. And I simultaneously realized they gave my students and me a workable tool we could use to extract these metaphors.

Lakoff argues that far from being just a literary device enhancing prose and poetry, metaphor is a fundamental basis for our conceptual structures. In the domain of knowledge work this makes a lot of sense: we learn abstract concepts by associating them with concrete things: atoms had electron orbits, then probability clouds; DNA is a spiral staircase. It seems reasonable that these associations would continue since they help tie the abstract ideas more firmly into human memory. I'd also suggest that it's another example of evolution's parsimony: where new needs (the ability to manipulate abstract ideas) are mapped onto structures developed for other purposes; as fins became legs. When we try to extend beyond the very limited cognitive resources we can apply to abstract ideas, we fall back on the much better developed parts of mind that animals have had for tens of millions of years. The ability to understand space and social relationships, e.g., become grounds to which metaphors map abstract thoughts. This supports my optimism that every endeavor of abstract thought will have at least some level of entities, states, and relationships that can be illustrated in an interface as information objects.

These mappings may continue whether they're accurate or not (as in the case of electron orbitals, or the false "Great Chain of Being" in evolution) until experts discuss faults with the metaphor and propose new ones (evolution's a bush, not a chain). This metaphor-pruning role of discussion in a culture of experts is a critical one for us: the metaphors are more than mappings inside one person's head; they are exposed as language in discussion. Language, fortunately, is freely available to a designer who listens carefully to the experts they're trying to support. And the way a metaphor works—in mind and in language—gives us the linguistic tool that we can use to start to expose the experts' metaphors that should form the basis for accurate design. Simply listen for any concrete modifier to an abstract thought: "high price" (there's no a priori reason a large number should be up—in grade school we count *down* the blackboard), or "buy side/sell side" buys are on the left throughout the world in finance.

¹⁹ The excellent early book *Metaphors We Live By*.^[3]

With this in mind we can better define the proper scope for a domain-specific information-object kind of interface. We look for tasks that are shared by a culture of experts within a domain of practice. And the domain of practice is simply the set of people sharing the metaphors: the people who all understand the same jargon.

If different groups of people *seem* to do the same job but don't understand one another's jargon, we should be designing different interfaces for each group.

Or we can fall back to teaching the experts we're trying to serve our new way of thinking about their domains: we can have them throw away their (perhaps centuries-old) language, metaphor, education and practice that brings life to their ideas, and replace it all by learning how the widgets in our general-purpose interface tool kits work. Please consider the former.

References

[1] J. Abrams "Stock Options," *I. D. The International Design Magazine*, June 2000, Volume 47, Number 4, pp. 48-51

[2] E. Hutchins, *Cognition in the Wild*, MIT Press, 1996

[3] G. Lakoff, M. Johnson, *Metaphors We Live By*, The University of Chicago Press, 1980

[4] A.M. MacEachren, *How Maps Work*, Guilford Press, 2004

[5] B.A. Nardi, V.L. O'day, *Information Ecologies*, MIT Press, 1999

[6] A. Öhlman, Anders Flykt, F. Esteve "Detecting the Snake in the Grass," *Journal of Experimental Psychology: General*, September 2001. Vol. 130(3), pp. 466-478

[7] C.E. Osgood, G.J. Suci, P.H. Tannenbaum, *The Measurement of Meaning*, University of Illinois Press, 1957

[8] S. E. Palmer, *Vision Science: Photons to Phenomenology*, MIT Press, 1999

[9] M.K. Singley and J.R. Anderson. *The Transfer of Cognitive Skill*, Harvard University Press, 1989 (also see the excellent J.R. Anderson *Cognitive Psychology and its Implications*, Worth Publishers, 2000)