Do Predictions of Visual Perception Aid Design?

RUTH ROSENHOLTZ
Massachusetts Institute of Technology

AMAL DORAI
Stanford University

and

ROSALIND FREEMAN
Skidmore College

Understanding and exploiting the abilities of the human visual system is an important part of the design of usable user interfaces and information visualizations. Designers traditionally learn qualitative rules-of-thumb for how to enable quick, easy and veridical perception of their design. More recently, work in human and computer vision has produced more quantitative models of human perception, which take as input arbitrary, complex images of a design. In this paper we ask whether models of perception aid the design process, using our tool DesignEye as a working example of a perceptual tool incorporating such models. Through a series of interactions with designers and design teams, we find that the models can help, but in somewhat unexpected ways. DesignEye was capable of facilitating A/B comparisons between designs, and judgments about the quality of a design. However, overall "goodness" values were not very useful, showed signs of interfering with a natural process of trading off perceptual vs. other design issues, and would likely interfere with acceptance of a perceptual tool by professional designers. Perhaps most surprisingly, DesignEye, by providing in essence a simple visualization of the design, seemed to facilitate communication not only about perceptual aspects of design, but also about design goals and how to achieve those goals. We discuss resulting design principles for making perceptual tools useful in general.

---

1. INTRODUCTION

Visual design is critical to the success of any presentation of information, from a 3-inch mobile phone display to a 100-foot billboard. However, design of user interfaces and information graphics is poorly understood, and somewhat hit-or-miss in terms of its effectiveness. A number of issues influence the success of a design, and these run the gamut of the underlying human behavior. A design must be good cognitively (can the user easily understand the semantic structure of the design?), perceptually (can they effortlessly interpret the visual information present in the design?), and socially (does the design fit into the user's workflow? will they want to use it?).

Following the visual system's "rules" of visual organization makes visual interpretation of designs effortless. However, models for how humans extract meaning from visual displays are incomplete. Designers often "eyeball" it, i.e. use their own visual systems to try to predict the percept.

Researchers have suggested guidelines for good visual design [Ware, 2004; Tufte, 1990; Cleveland, 1980; Mackinlay, 1988]. Such rules-of-thumb, based on percepts of very simple displays, are useful in understanding and guiding design. However, designers may have difficulty applying them to more complex and cluttered displays.

More recently, advances in human and computer vision have produced quantitative models of some aspects of visual perception, including what draws attention in a display [e.g. Itti et al, 1998; Rosenholtz, 2001; Rosenholtz and Jin, 2005], visual clutter [Rosenholtz et al, 2005; Rosenholtz et al, 2007; Bravo and Farid, 2008; Lohrenz and Gendron, 2008, van den Berg et al, 2009], perceptual grouping [see Rosenholtz et al, 2009 for a review], and capabilities of peripheral vision [Balas et al, 2009]. While it seems logical that such tools should aid design, this has not been carefully studied, especially within real work environments.

One objection might be that designers are highly talented, and highly trained both in design techniques and in introspecting on their subjective visual experience. We initially thought that the main utility of models of perception might lie in their use by non-professional designers, or in systems with dynamic displays, in which there can be no human in the loop evaluating every display. However, our study has led us to rethink the utility of perceptual models for professional design teams, to the point that in this paper we now focus on use by such teams.

We have interacted with a variety of potential users of perceptual modeling tools. Our interactions included both brief interviews and in-depth discussions with design teams. As part of some of our interactions, users explored designs with the aid of our perceptual tool, DesignEye, which we describe in its early form in the next section. The purpose of having users interact with DesignEye was not to conduct a user study of DesignEye; our aim was to gain insight about the use of perceptual tools that would generalize to tools other than our own. However, we know of no perceptual tools currently in common use by designers. In order to meaningfully study use of perceptual tools, we needed an instance of a tool so that we could observe its use in more detail. DesignEye served as that placeholder for perceptual tools more generally. In Section 2, we discuss why DesignEye makes a good placeholder. In that section we also discuss DesignEye in more detail.

Overall, we conducted approximately 20 hours of interviews and observations, many of them audio- or videotaped and transcribed, plus email interactions. Our study included a total of 40 participants from 9 organizations. We analyzed our data by first defining concept categories based both on the goals of the study ("problems understanding the

tool," "reasons for desiring the tool," "response to using the tool," "trading off perceptual concerns with other design issues," "user interface issues," "A/B design comparisons," "design materials [sketches on paper, illustrator files, etc.]"), and upon themes emerging from our observations ("issues with the current design process," "communication issues within the design team," "desire for an oracle," "how much do users trust the predictions?," "can users communicate about saliency and clutter?," "interpreting saliency/clutter maps," "interpretation of global clutter values," "discussion of design goals"). We trained a researcher new to our team in these categories, and she used them to sort the transcripts and notes.

We attempted to answer several key questions. First, does anyone want tools like these, and if so, who and why? How do design teams currently work on a design, and what benefits might a perceptual tool bring to the design process? Do users understand and trust the perceptual predictions? What is required for designer acceptance of such a tool? Throughout this paper, we focus predominantly on issues that generalize to other perceptual tools and in some cases to design tools more generally. Based upon what we have learned, we redesigned our own perceptual tool, DesignEye. More generally, we discuss design principles for perceptual tools, in order to enable them to be useful for design teams.

2. THE PERCEPTUAL TOOLS UNDERLYING DESIGNEYE

We work on designing and testing predictive, quantitative models of a variety of visual phenomena. This includes modeling of visual saliency, i.e. what draws one's attention in a display due to its features like color, size, etc., being sufficiently distinct from the features of surrounding regions [Rosenholtz, 2001; Rosenholtz & Jin, 2005]. We also have designed a measure of visual clutter [Rosenholtz et al, 2005; Rosenholtz, Li, & Nakano, 2007]. While our lab also models other perceptual phenomena, here we focus on saliency and clutter for their general applicability, the maturity of the models, and because they are reasonably computationally efficient.

Our saliency model takes as input any image of a display, and can output one or more saliency "maps." These maps represent the saliency of each region of the display such that higher luminance means higher saliency. Higher saliency of a target correlates well with difficulty searching for that target [Rosenholtz, 1999; Rosenholtz, 2001ab; Zhaoping, 2002; Rosenholtz et al, 2004] and is believed to be related to that target's ability to draw attention [Rosenholtz, 2001b]. Saliency maps can make reasonable predictions of the average eye fixations made during free-viewing of an image [Itti et al, 1998; Rosenholtz & Jin, 2005], though other factors such as knowledge of likely locations of the tar-

get are also highly important [Torralba et al, 2006]. Multiple saliency maps may indicate what causes an item to draw attention, e.g. does its color, or size, make it stand out?

DesignEye's second component is a measure of visual clutter. Clutter is known to degrade performance at a wide variety of visual tasks. Clutter is not, however, an unambiguously bad thing. Designs that are too uncluttered can make inefficient use of display space, and require the user to go to another page or view in order to acquire the information needed for their task. Such page switching has its own performance costs [Staggers, 1993]. Given a reliable measure of the visual clutter in a display, designers could perhaps optimize display clutter by trading off giving the user more information against making it more difficult for the user to quickly and veridically extract that information from the display.

Our clutter measure takes as input an arbitrary image. It can output a global value indicating the level of clutter for the display as a whole. We have demonstrated that this global clutter value correlates well with both subjective impressions of visual clutter and with average difficulty in searching within a particular display [Rosenholtz et al, 2005; Rosenholtz, Li, and Nakano, 2007]. Our clutter measure can also output clutter maps. These maps indicate the level of clutter for each region of the display, as well as whether the clutter is due to a particular feature, e.g. due to use of an excessive variety of colors to represent information. Figure 1 shows an example of both a saliency map and several clutter maps.

Several features make our saliency and clutter predictions appropriate in the current study as placeholders for perceptual prediction tools more generally. First, predicting where a user will look in a display or scene forms a common subset of computational models of vision; this includes both models like ours that predict what regions will be "salient" due to unusual features [e.g. Itti et al, 1998; Rosenholtz & Jin, 2005; Zhaoping, L., 2002; Bruce & Tsotsos, 2009], and models that incorporate knowledge of the task and type of display [e.g. Torralba et al, 2006] – such as "the user is looking for the fuel gage on a dashboard." Eye tracking heat maps, while a different sort of "model," also make predictions of where users will look. Quantifying the effects of display clutter is a more recent enterprise in the study of human vision, but nonetheless a popular one, with at least 4 competing models [Rosenholtz et al, 2005; Rosenholtz et al, 2007; Bravo and Farid, 2008; Lohrenz and Gendron, 2008, van den Berg et al, 2009].
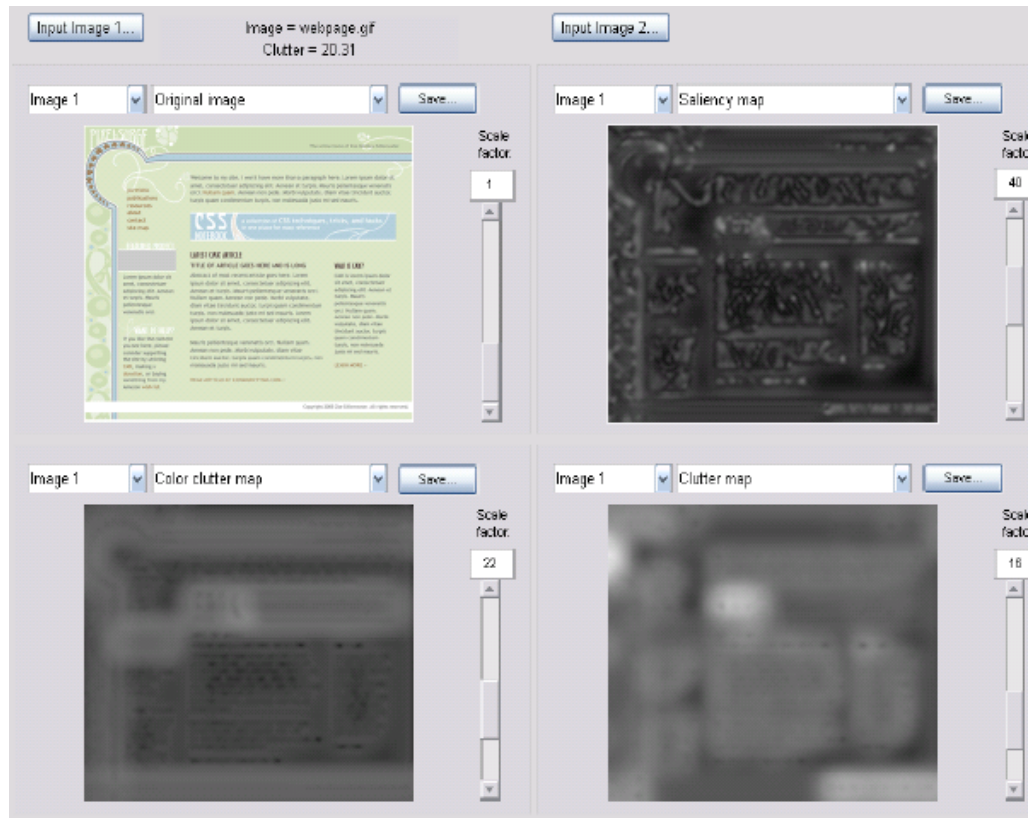
Figure 1: An early version of DesignEye, analyzing a web page (upper left, from http://www.communitymx.com/content/article.cfm?page=3&cid=3D2DA). Shown are the saliency map (upper right), color clutter map (lower left), and clutter combined across all features and scales (lower right).

The second reason why our saliency and clutter tools serve as useful stand-ins for perceptual tools more generally is the form in which our tools present the predictions to the designer. As mentioned above, we present most of our predictions in the form of one or more images. Images form by far the most common way of presenting complex predictions of human perception. In addition to heat map-like saliency and clutter maps, some predictions of where users will likely direct their eyes show a series of predicted fixations, overlaid on the original display [e.g. Itti et al, 1998]. Perceptual grouping predictions are typically shown as one or more images in which either (a) all pixels belonging to a given group are shown in a common color (Figure 2b), (b) boundaries are shown between groups, or (c) only one group is shown per image [e.g. Shi & Malik, 1997; Wattenberg & Fisher, 2003; Rosenholtz et al, 2009; Paris, 2008]. Alternatively, algorithms sometimes depict the hierarchical structure of perceived groups using a tree structures [e.g. Wattenberg & Fisher, 2003, and Figure 2c]. Early models of peripheral vision de-

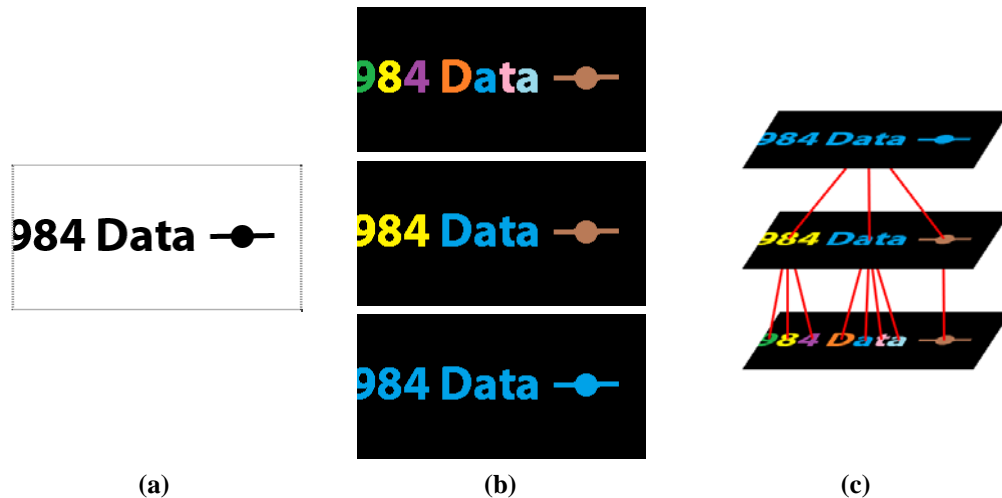| (a) | (b) | (c) |

Figure 2: To get a sense of models of visual perception not currently included in DesignEye, consider a model of perceptual grouping. (a) Original image. Here we show only a fragment, for simplicity. (b) Perceived groups predicted by [Rosenholtz et al, 2009]. Each color represents a predicted group. As grouping scale increases from top to bottom, the algorithm predicts perception of letters, which group into larger structures like words, then form a line of text. (c) An alternative visualization of these results [e.g. Wattenberg & Fisher, 2003] might display fine scale groupings with a graph- or tree-like structure showing how the groups combine to form groups at larger scales.

picted how a display might look in the periphery by increasingly blurring an image of the display with distance from fixation [e.g. Anstis, 1998].

While there are of course differences between a saliency map and, say, a map showing perceptual groupings, they are similar in that they present predictions in "image space," i.e. in an image for which each pixel corresponds to the equivalent pixel in the original display. These prediction images also tend to be simpler than the original display. At the user interface level, interactions with the image-based predictions in DesignEye are likely to be similar to interactions with other perceptual models that present image-based predictions.

Alternatively, some models output, for a given display, a single number like our overall clutter value. This includes not only most of the aforementioned clutter measures, but also such predictions as the reaction time to find a target in a given display, or percent correct measures of ability to perform a given task. The vast majority of perceptual tools in existence, or currently being developed, output image-based visualizations and/or report single, scalar, quantitative predictions, as in DesignEye.

For the purposes of this study, we developed a fairly crude graphical user interface (GUI) wrapper for our saliency and clutter tools, within the MATLAB graphical user interface design environment (GUIDE). Our initial version of DesignEye allowed users to input up to two images of designs at a time. As shown in Figure 1, users could view

up to 4 images at once. These images could be original images, or saliency or clutter maps for either of the two input images. The saliency maps showed saliency at each location combined across scale and feature (color, orientation, and contrast). Users could view either an overall clutter map, or the contribution to clutter from individual features such as color or orientation. Users could scale the contrast of each map using a slider bar, so as to render differences in saliency or clutter more or less visible. As it is desirable when comparing maps to equate the contrast of those maps, the scale factor appeared next to the slider bars. Finally, the overall clutter values for the two input images appeared at the top of the display.

The ease of developing in GUIDE allowed us to change features within days, and sometimes even online, to iteratively redesign our tool in response to feedback and understanding we gained from studying our prospective users.

3. THE USERS WE STUDIED

We interacted with 9 organizations that were potential users of the DesignEye tool. (Throughout the rest of this paper, we refer to users of a perceptual tool as "users." These users, in turn, produce products of some sort, and we refer to the consumers of those products as "customers.") At each organization, typically 1 potential user contacted us because they were interested in analyzing the saliency or visual clutter in their displays. These users knew about our work through our papers on visual clutter [Rosenholtz et al 2005; Rosenholtz et al 2007], through research talks by Rosenholtz on saliency and clutter, through contact with M.I.T.'s industrial liaison program, and through a press article on our clutter measure put out by M.I.T.'s news office: http://web.mit.edu/newsoffice/2007/clutter-0821.html.

This means of selecting users for our study no doubt introduces bias into our study. Such bias likely affects our results related to interest in perceptual tools, acceptance of them, and perhaps conclusions about the current design environment – users who contacted us may have been more likely to have come from organizations with issues surrounding their design process. However, any effects of such biases are modulated by the fact that overall we studied 40 potential users, the majority of whom were not the ones who contacted us. Furthermore, our most interesting conclusions go beyond mere interest in perceptual tools, to examining interactions with DesignEye. Such details of interacting with the tool likely are less affected by selection bias.

Our users included design teams at software companies, in the transportation industry, individual freelance designers, a print publisher, a company selling advertising space, and a large producer of snack foods. We have also been contacted by interested *researchers*

in user interface (UI) design, human vision, and design of tools for low-vision patients. Here we focus exclusively on professional designers and design teams working on design of products, broadly construed, as their needs may differ from those of researchers. All users and companies were promised anonymity in return for their participation. Companies' designs are often proprietary prior to launch, and neither organizations nor individuals may want to be identified as they talk about their work environment. As a result, we provide only limited, unidentifying details of the participants and our interactions with them, and show no examples of actual designs users provided.

In all cases, we provided users with an explanation of the visual behavior predicted by our saliency model and clutter measure. We explained the meaning of the various saliency and clutter maps displayed by our tool, and for more in-depth interactions explained the functionality of DesignEye.

Interactions with 6 of the organizations consisted of brief conversations (approximately 1 hour, on average) about their desired use of and requirements for DesignEye. Interactions with the other 3 organizations were more in-depth discussions with teams (averaging approximately 5 hours each), about issues in their current work practice, and capabilities and potential use of our tool. For 4 of the organizations, we observed their use of either DesignEye or our pre-GUI perceptual tools to analyze their displays in more depth. For 2 of the more in-depth interactions we recorded audio and/or video, with permission of the companies and individuals involved.

In the real world, most design involves not individual designers, but teams that must collaborate across several disciplines: design, engineering, marketing, and product management [Bucciarelli 1988; Star 1989; Sonnenwald 1996; Hendry 2004; Adamczyk and Twidale 2007]. We refer to this collection of interdisciplinary workers as the "design team." As discussed below, we have found that the interdisciplinary aspect of design is one key to understanding the utility of perceptual design tools. Where possible, we talked to members of the design team from a diverse set of disciplines. In total, we interviewed 6 managers, 7 usability specialists, 1 marketing person, 7 researchers, 12 designers, 6 engineers, and 1 editor. (A number of these people work in two disciplines, e.g. they act as both manager and usability specialist, or usability specialist and human-computer interaction designer. In such cases, we have counted them according to their current dominant role in the organization.)

In the following three sections, we discuss what we learned from these interactions.

4. DOES ANYONE WANT A PERCEPTUAL TOOL? WHY?

Based solely upon the number and diversity of inquiries that our lab has received in response to our predictive models of human vision, clearly users are interested in such tools to aid their design process. In this section, we discuss what potential users seem to *want* from a perceptual tool.

4.1 General Design Guidance

When creating DesignEye, the most obvious use we imagined was to provide guidance during design. Perceptual tools might allow one to decide whether a button should be black or red, how high-contrast to make labels on a plot, and how best to draw attention to an alert. We were uncertain whether such functionality would be helpful for, or welcomed by, trained designers.

Potential users were definitely interested in perceptual tools as providing design guidance. This interest was mentioned by 8 out of 9 of the design teams we studied. We were asked, "Can this tell us where to put components?" "Could it suggest changes?" Can it help make salient the relationship between an old and new version of software during the user's transition to the new version?

Trained designers tend to be more wary of advice from a perceptual tool, a point we return to later. However, we nonetheless received comments from designers like, "It would be hugely valuable to try out the tool on [our] designs." A usability specialist gave one reason why such a tool could augment even skilled subjective judgments of a design. Interaction with a naïve user, he said, had shown that he needed, "to know how a naïve user would look at [the design], because now [he] cannot see it as a naïve person would." Furthermore, many people on a given design team, such as the engineers and managers, may have no training in visual design.

4.2 Assigning an Economic Value to a Good Percept

One desired use of DesignEye that surprised us was the desire to have an objective measure of the goodness of a percept, so as to assign an economic value to it (2 out of 9 study participants). These organizations wanted to charge customers more for better ad space, or were willing to pay more for better perception of their products on, for example, a supermarket shelf.

4.3 Proving Your Design is Better, with Fewer User Studies

A frequent interest in our tools, however, was to in some way prove to someone that the design you had chosen was better than the design another party preferred (5 organizations studied). Furthermore, people wanted to prove their point with fewer slow, often costly user studies (10 users from 3 organizations made comments to this effect). One user

spoke for a number of others, in saying, "[We were] adding lots of stuff [to the design, and needed] to measure the effects. We do a great job with simulators, internal reviews, but a lot of the early analysis is subjective, anyhow. Can we get the same results at a glance, to skip two passes of review and the slow user studies?" Another user added that user testing was fine, but "too slow, for late in a design [process]."

Sometimes these requests for proof were amiable, in which case they resembled requests for general design guidance. Users simply wanted help deciding between two designs.

Other interviewees clearly desired use of the tool in more contentious situations. 5 out of 9 of our participant organizations raised the issue of contention over designs and the desire to prove one design was better than another.[1] "Would [these] tools help in the mediating process?" "We don't have a way to say 'no' to a request to add more [clutter to a design. With this tool] we would not be empty-handed." It would be "good to have an objective measure when responding to a management request... [we] want anything that can help... objectify the data." "Can I just take [the predictions of this tool] back and use it to hit people over the head with?" Such users also hoped that the tool could "be used to prove *why* something should be a certain way."

In the course of our interviews we heard statements like this from numerous people, in diverse occupations, and within companies producing a range of products. As a result, we studied the design environments within 3 of the companies in more depth. Below we largely report issues of relevance to use of a perceptual tool, but also include issues that involve design tools more generally.

5. WHAT IS THE "TYPICAL" DESIGN ENVIRONMENT?

Many of our observations about the design process and design environment echo previous work. Nonetheless, we repeat these observations here, due to their relevance to our current questions about the utility of perceptual tools.

Design rarely involves only one person [Bucciarelli 1988; Star 1989; Sonnenwald 1996; Hendry 2004; Adamczyk and Twidale 2007]. Exceptions include an amateur's design of their own web page, and some research design. At minimum, most design involves a designer and someone who approves the design. More typically, design involves multidisciplinary teams. These teams can include designers, their manager, engineers, their manager, usability experts, and their manager. In addition, there is often in-

---

[1] This ratio likely underestimates the extent of the issue. In our early interactions, before contention over designs became a clear theme, we did not ask about difficulties coming to agreement if the participants did not raise the issue themselves.

put and perhaps veto power from customers and marketing. All of these personnel are potential users of a perceptual tool.

In our studies, users frequently complained about disconnects in communication between these disciplines (4 of the organizations studied overall, 14 comments in longer, in-depth interviews) [as in Bucciarelli 1988; Sonnenwald 1996; Hendry 2004]. These disconnects in communication are reflected in both work practice and in the different tools and representations used by the different groups. [Bucciarelli 1988; Hendry 2004; Henderson 1995] The process we describe here is to some extent a worst-case scenario created from a composite of several companies.

Designs may begin with market researchers, customers, or management, who make preliminary suggestions for design. Designers then often produce static graphic designs. The design team must then somehow transition from static sketches, through a dynamic design, and ultimately produce a working UI.

The next stage typically involves a round of user testing. User testing at this early stage is often used to facilitate communication about usability issues and conflicting goals for the design, as much as to actually test users (8 comments from in-depth interviews; mentioned in 10 comments by 3 organizations overall). One group, for instance, might have good reason to believe that a design would lead to usability issues, or be distracting, but would have trouble convincing other groups without a user study. One participant said, "We are asked to demonstrate things that are known and obvious." A typical manager commented, "Communication is a huge issue. We need engineers, managers, and designers all to talk to each other. What's the common language? …Costly user studies." User studies are of course quite important for uncovering usability issues in a design, but a number of participants complained about the use of them merely as a communication device.

Since a main purpose of these user studies is to communicate across disciplines and departments, the studies need to consume quite a number of man-hours in order to be effective. Not everyone on the team is typically present for the studies, and it can take a lot of time and effort to convince those who were absent to accept the results. One higher-level manager explained that it is "too time-consuming for designers to watch the user studies, but they could learn a lot from that about requirements, good and bad points of the design." In meetings, those who conducted the user study attempt to explain why they made changes to the design, but complain that the conversation is "lossy." Even once designers, engineers, and usability experts agree on a design, they need to effectively communicate their results to managers and marketing personnel, who may complain

that the design has "no pizzazz" or deviates from (possibly ill-conceived) requirements from the customer.

At one company, researchers (who had acted as engineers on a recent project at one company studied) complained primarily of difficulties communicating with designers, and with designers not being sufficiently comfortable with the tools the engineers wanted to use to produce a design (9 comments). Engineers told us, "Better to start with a designer/computer engineer in one person." "The designer's education is usually more art than technology."

Designers, on the other hand, complained about being overly constrained, and having their talents underutilized (7 comments). "[There is a] tension between what to display, aesthetics of designers, requirements of users, and management or marketing requirements for selling the product." "98% of our design is rules we're obliged to follow. But new technology outpaces the rules, and it takes years to eliminate unnecessary rules." Designers also complain about communication issues. They say they work with their "gut sense," and extensive knowledge derived from training. It's hard to explain one's gut sense, but the designers need to justify many decisions to their managers and others (6 comments). "You hired me because of my intuitions, and now I need to justify each one." "We are butting heads against what people will let you do."

Of course, sometimes the process proceeds more smoothly. At one of the three organizations studied in depth, personnel with different training reported that they work well together. We were told that interdisciplinary teams work better than collaboration across teams consisting of only one discipline (3 comments). However, a number of interviewees from a number of different companies resonated with these difficulties.

Other researchers have suggested tools to aid the transition between multiple representations of a design – e.g. that make it easier to transition from sketches to other design representations [Landay and Myers 2001; Sinha and Landay 2001; Newman et al 2003; Adler and Davis 2004; Hammond and Davis 2005]. Other tools have focused on collaboration and personal interactions in design teams; process innovations driven by the need to create a shared vocabulary and agree on common goals.

What role might a perceptual tool have to play within this environment, and how likely is such a tool to be accepted and used? Clearly some of the issues raised above are beyond the scope of a perceptual tool. However, in observing user interaction with DesignEye, we have found it to help in sometimes surprising ways. In the following section, we describe our observations in watching users interact with DesignEye.

6. HOW USERS INTERACT WITH A PERCEPTUAL TOOL

Most of the in-depth interactions with DesignEye that we describe come from observing users at 2 companies: one manager responsible for designing physical devices within the transportation industry, and 7 HCI/usability experts designing and testing software tools. We also had more lightweight interactions with an individual web designer and with an individual at another organization interested in the use of the overall clutter values for a marketing-related purpose. All interactions were with the explicit purpose of exploring whether DesignEye might aid their design process. In order to test our tool, users provided images of candidate designs. Only two users (the manager in the transportation industry, and the web designer) were in the midst of a design at the time we studied them. Other users nonetheless often brought images about which there was ongoing debate about the best design. Overall, during our study, users tested DesignEye on a total of 37 images they provided, including 11 information visualizations, 3 images of a software interface, 2 physical devices with LCD displays, 2 web pages, and 19 marketing-related images.

Overall, response was positive, and users found the tools potentially useful at 3 of the 4 organizations that tested the tool. The 4th organization, which found the tool less obviously useful, was predominantly looking only at overall clutter values for a particular marketing application. We discuss issues around overall clutter values further, below. The positive remarks about DesignEye included 8 comments along the lines of: "There's a lot of interesting stuff I could do just with this [early version]." "[Someone] who is redesigning the desktop should see this." "This is really exciting."

Below, we have divided our discussion into subsections corresponding to key lessons learned.

6.1 How Accurate Do People Think the Tool Is?

Users, by default, seemed to believe that the tool was accurate. In the two more in-depth interactions with the tool, several users (4 comments) made explicit statements to the effect of, "I am assuming that the algorithm works," and in general users seemed basically to accept as a given that the tool gave the right answers. Nine user comments implicitly assumed the tool had given the right answers, e.g. because the users suggested ways to change the design given the predictions of the tool. In comparison, only 4 user comments implicitly or explicitly questioned the specific predictions of DesignEye.

Trained designers were generally more skeptical and wary of the tool. The wariness seemed more due to fears of their intuitions being replaced or constrained by an algorithm than actual performance of that algorithm. Designers did comment that the tool was, e.g., "effective in showing things that are salient." One designer expressed a com-

bination of confidence in and wariness of the tool when he said he would want the tool, but "wouldn't want other people to know [it] existed."

In at least 4 examples, confidence in the tool was supported when the tool confirmed what users already knew about a design. For example, one user was pleased when the tool showed that the help content for his software package was highly cluttered. He said that he had always believed that the non-help text in other portions of the software was easier to manage than help text, and that this was shown in our computed clutter values. "This text I know is a problem… It would be nice to definitively say why."

Users were in fact so confident about the tool's predictions that in 2 instances it was difficult to shake that confidence even when we intervened to tell them about known situations in which the tool might give the wrong answers. When we once tried to discourage over-reliance on the difference between overall clutter values between two displays, the user responded that he believed the difference was correct. One user told us that he of course had his own visual impressions, but he didn't know whether or not to trust them unless they were backed by numbers. One user seemed initially to question the predictions of the tools, saying, "The boundaries look so big on this [clutter] map, and yet I have no trouble ignoring them [sic]." But even when the authors agreed that the tool might over-emphasize the saliency of high-contrast boundaries, he then told us that he trusts the tool, saying his design "spends too much energy having the borders stand out."

As designers of perceptual tools, we were disturbed by users' tendency to overestimate tool accuracy even when it disagreed with their own perceptions or with our estimates of the accuracy. This tendency interacted negatively with a strong desire among users for a tool that could tell them the absolute goodness of a design relative to other designs. Regardless of their reasons for wanting a perceptual tool, users at 6 of the 9 organizations studied (10 users in total) desired an "oracle." (Notably absent in the desire for an oracle were the designers, who already feel their design process is over-constrained.) This oracle would preferably report a single number that would say definitively whether one design was better than another, help make decisions, and prove to others either that a decision is the correct one or that a design has a given economic value. In later interactions, we felt the need to actively discourage users from thinking of the tool as an oracle. As discussed below, we believe perceptual tools have great utility even without being 100% accurate at all times, and that one may increase designer acceptance of such tools by emphasizing non-oracle affordances.

Ultimately, many users do have a healthy desire for testing of perceptual tools within their domain, to validate for themselves how well they work (7 comments, total). They

often have creative ideas of how to do this. They have old designs they know worked or did not work. In some cases customers have sent in pictures of the state of the display when it was confusing. Users looked forward to testing those designs to see if the tool agreed with previous conclusions. Furthermore, users wanted to compare predictions with eye tracking or comprehension data on their designs. It may be that users' overconfidence in a perceptual tool is only a short-term problem.

6.2 Do Users Understand the Perceptual Predictions?

Certain aspects of the perceptual predictions were quite clear to users. The meaning of saliency and clutter seemed clear to all users studied. The saliency (and clutter) maps used a scale in which higher luminance mean higher saliency (or clutter), and this required no explanation. One knowledgeable user, who had read a paper about the clutter algorithm, understood without instruction that a blurrier region in the saliency map meant that the item was salient at a coarser scale, i.e. when viewed from a greater distance.

While the saliency *predictions* seemed clear and natural, some users, particularly a couple of higher-level managers less intimately part of the design process, were uncertain how to think about the *desired* saliency. The desired saliency of an item depends upon its expected use. We told such users that: "Salient items should be things that the customer may have to search for. This includes items that are rarely used, or not always in the same place. Alerts also need to be salient; customers do not actively look for them, so alerts need to draw attention. Items that always appear in the same place, and that a customer will look for often enough to learn their location, can be less salient. Of course, items that a customer need never search for can also be less salient, unless those items are useful for spatial orienting or some other purpose." This tutorial in how to think about what saliency might be desired seemed quite valuable in interactions between the user and the tool.

We also wanted to see whether users could naturally trade off perceptual predictions with other design concerns. A tool based upon a few perceptual algorithms cannot make the complex design tradeoffs of information vs. simplicity, or accessibility vs. discoverability. Users at times needed to trade off perceptual issues, such as high clutter, with more cognitive issues. They had no apparent difficulty doing so. "Well, sure, that design is more cluttered, but it has important information in it," said one user. Additional cognitive modeling might well help design teams, for some of the same reasons that perceptual tools may be useful. But cognitive modeling seems not to be *required* in order for perceptual tools to be useful.

There were several confusions specific to our algorithms, which we mention only briefly here. At least 3 users had difficulty understanding the difference between saliency maps and clutter maps. They seemed to prefer the clutter maps and (incorrectly) used them as indicators of what would draw attention in a display. Furthermore, comparing values for the overall clutter in a display was less useful than expected, in all 4 interactions with DesignEye. When users made side-by-side comparisons of two designs, the two typically differed only slightly. As a result, the overall clutter values were often quite similar, and most likely not significant. However, small design differences did show up clearly in the saliency and clutter maps. As one user said, "The overall clutter values don't seem very useful, but some of the [maps] are pretty explicit." A last confusion specific to our saliency predictions is that at least 2 users did not find it intuitive that a visible, isolated item against a uniform background was typically fairly salient. These users thought of salient items as being, e.g., a bright color, not items that were distant from other items. This might be one place where more transparency of the algorithm would help, e.g. by optionally indicating what makes an item salient.

Other confusions about the algorithms likely hold for perceptual tools more generally. For many of the problems that the visual system solves (what to attend to, how to group or segment the image, what objects are present, etc.), it first analyzes a number of basic features at a number of different scales. Such features include color, orientation, and motion. Analysis likely occurs at many scales because one does not know, *a priori*, how big relevant items appear on the retina. This analysis is related to viewing the design at a number of different distances. Most effective models of human vision mimic such multiscale processing in order to best predict visual perception. Figure 2bc, for instance, depicts predicted perceptual groupings at 3 different scales.

Various parts of this process are confusing to users. At least 2 users had trouble with the concept of different scales of processing. For example, it was not obvious to one user that lines of text, at a coarse scale, have a strong horizontal orientation (see Figure 3). Similarly, early versions of our tool displayed the original image at a small screen size of only about 2" × 3." For designs that in real life often take up at least an entire computer screen, this is a very small view. Many of the interesting percepts predicted by the algorithms were not visible at this small scale. Three separate comments indicated that this was confusing to users. DesignEye allows entry of the size that the display will appear to the customer, so as to make valid predictions for a typical viewing distance. However, perceptual tools, in addition to getting the computations right for a particular viewing

natural images, and grouping in more complex diagrams.

PREVIOUS WORK

A great body of computer vision work exists on the topic of perceptual grouping. However, for user interface (UI)

(a)   (b)

Figure 3: A snippet of text, shown at two different scales. A user viewing the coarse scale, (a), might wonder why an algorithm predicted that the header was salient, whereas a user viewing the fine scale, (b), might find a prediction of horizontal structure (the lines of text) unintuitive. A perceptual tool needs to facilitate viewing of input images at varying scales, as predictions will often make use of multiple scales of processing.

distance, should facilitate display of the original images at different effective "distances," e.g. at different levels of zoom, to aid in understanding the perceptual predictions.

A final confusion occurred that is likely common to other models that return their predictions in the form of an image (for us, the saliency and clutter maps). If these maps are high resolution and contain sharp boundaries between regions, users take those sharp boundaries quite literally. At least 2 users, for instance, thought that high clutter within a region, and low clutter at a neighboring pixel, meant that that neighboring pixel was clutter-free, in spite of its immediate proximity to the high clutter regions. This misunderstanding was severe enough that early on we began blurring the clutter maps before displaying them for the user.

6.3 Does a Perceptual Tool Give Design Teams a Common Vocabulary?

In general, there seemed to be hope that a perceptual tool could provide a common language for discussion of the design. Users across the spectrum (designers, engineers, researchers, managers) found it natural to talk about saliency and clutter. With occasional instruction (for 2 managers, mentioned above) on how these relate to design goals, they also found it natural to talk about the *desired* saliency and clutter.

Members of the design team, whether designers, engineers, researchers, or managers, expressed a desire for better communication across disciplines and groups. Perceptual tools like DesignEye, which create a natural and relevant language for talking about the design, may aid in communication between diverse groups. Members of design teams from all disciplines expressed hope that this might be true (5 comments). As one participant put it, "We could [use this tool] as a way to communicate common concerns." Designers seemed to worry that a perceptual tool would only be used to put additional constraints on their already over-constrained designs. Focus on the communication benefits might help to alleviate their concerns. Designers desired a means to explain their often non-verbalized design intuitions and choices to those without design training (6 comments). The perceptual language associated with a perceptual tool, as it seems natural for

users from the range of disciplines present on a typical design team, may be of particular help to them in explaining their "gut sense" about a design to non-designers. This possibility deserves further study.

6.4 What a Typical Interaction Looks Like

Below we describe key features of typical interactions with DesignEye. We focus on two aspects that were clear, often repeated, and largely unexpected by the authors.

6.4.1 By Giving a Simplified Visualization of the Design, a Perceptual Tool Enables a "Guided Tour" of the Design

While users did sometimes bring specific questions to DesignEye ("which choice of colors is better?"), in many cases they more generally wanted to evaluate the "goodness" of a design. What we saw in virtually all interactions with DesignEye was that when looking at the design itself, users were unsure where to begin in evaluating it. (The 2 exceptions were A/B comparisons where the user wanted to test whether the tool would prove they were right about which design was better.) All of the designs had a lot going on, both visually and in terms of ways in which customers might interact with the UI. (Figure 4a gives a hypothetical example image, which one might imagine a designer using to design the UI for Microsoft Powerpoint.) Many different aspects of a design need to be perceptually and cognitively right, and users found it difficult to know where to start their evaluation. Even if (perhaps with training), design team members might be able to use their own visual systems to say what the likely percept will be, it is difficult, when looking at the whole of a complicated design, to know how to break it apart and ask the necessary specific questions: Does A perceptually group with B? Does C stand out?

This puzzlement about where to begin often evaporated when users looked at, e.g., a design's saliency map (all 3 organizations interacting with DesignEye's saliency and clutter maps, involving analysis of 16 designs). (Figure 4b.) It was as if the saliency and clutter maps gave users a natural walk-through of the design. Typically users would look first at the most salient or cluttered regions of the maps, look back to the design to see what they corresponded to, and discuss whether those regions should be so salient or cluttered (as described in more detail in the next subsection). Users would then check the saliency of items they felt *should* be salient, and puzzle over why they were not.

It is worth comparing this guided tour of the design with Kosslyn's systematic method of analyzing charts and graphs [Kosslyn 1989], or with task analysis and cognitive walk-throughs [Wharton et al 1994]. Both types of systems are built upon the same observation that often, in analyzing whether a design has flaws, one does not know how to proceed, and both consist of specific programs to help analyze a design. Kosslyn's solu-
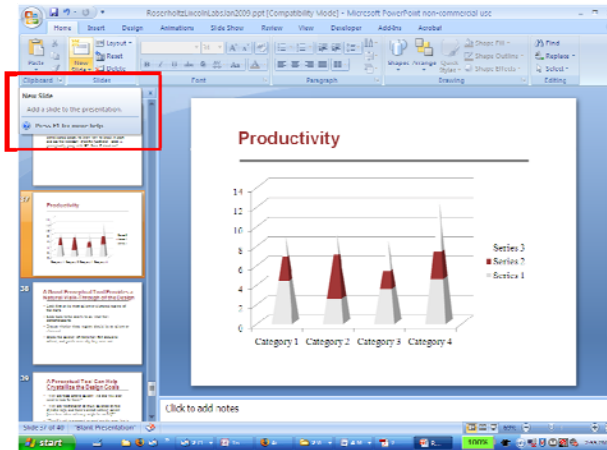
(a)

(b)

Figure 4: An example design (a) and its saliency map (b). Users faced with (a) have difficulty knowing where to begin analysis of the quality of the design. When shown perceptual predictions, such as (b), as well, users naturally alternate between the visualization of the prediction and the original image. Users typically first pick out more regions of higher saliency or clutter in (b), then look for what they correspond to in (a). They asked themselves whether those regions should be so salient or cluttered. Then users typically looked for items that "should" be salient. For example, the help dialog box, (boxed in red on both the original image and saliency map), should be salient, but the saliency model predicts that it is not.

tion, for example, is a step-by-step process of "isolating four types of constituents in a display, and specifying their structure and interrelations at a syntactic, semantic, and pragmatic level of analysis." First look at the background – does it violate any design rules? Then examine the "framework" of a graph, and so on.

Kosslyn's approach and cognitive walk-throughs explicitly capture many more aspects of the design than a perceptual tool. However, the guided tour provided by a perceptual tool appears far more natural than following a list of steps. Based on user facial

expressions and tone of voice during this guided tour, discovery of the source of high saliency seems to combine just enough challenge and reward to make the process fun.

Perceptual predictions like saliency and clutter maps essentially serve as low-information visualizations of a complicated design. Such visualizations enable easier design evaluation. Other low-information visualizations of the design no doubt have similar utility. Such visualizations might come from other visual predictions, such as perceptual grouping, or from another source. Heatmaps, either from eye tracking or clickstream data, for instance, provide visualizations similar to saliency maps, based on eye tracking or mouse click data. Knowing how to simplify a design requires knowing what is important, and perceptual models afford knowing what is important to the visual system.

6.4.2 Perceptual Tools Help Crystallize the Goals of the Design

Prior to our study, we expected interactions with DesignEye to center around visual aspects of the design. However, an immediate consequence of the "guided tour" of a design was often discussion of the goals of the design, and increased clarity of those goals.

We frequently heard from users such comments as, "Why are these arrows salient? *No one will ever need to look for them!*" "Why are we spending so much salience on this stylistic logo, and there's almost nothing salient [in this location where saliency might be useful]?" "This file tab *is supposed to stand out the most*, but it is actually the least salient." (Comments from all 3 organizations interacting with DesignEye's saliency and clutter maps, including analysis of 7 designs or design pairs.) Interestingly, several user realizations mirrored Tufte's exhortations to minimize non-data "ink" [Tufte 1990]. One user was upset that borders were just as salient as text that a user "might be trying to read." Another said there was "no reason for the meaningless title" on a window to be so salient, when his application should be aiming for content inside the window to have the most saliency.

6.5 What is the Role of a Perceptual Tool in Guiding Design and Participating in Design Conflicts?

Users want from a perceptual tool both design guidance and help resolving conflicts. There is hope for a perceptual tool to play both of these roles. Whether a perceptual tool helps is, of course, difficult to definitively prove, without much more extensive study. An individual web designer has already received positive feedback from customers due to a change he made to his web site based on DesignEye. Clearly DesignEye helps people to have interesting conversations and insights about a design and its goals. Users found the tool valuable for making A/B comparisons between competing designs (users performed 9 A/B comparisons during the course of our study). It frequently validated in a

lightweight manner insights the design team had already gathered through more extensive interactions with the design.

Our saliency and clutter maps seemed the most useful parts of DesignEye for improving design and resolving conflicts. These maps enable in-depth A/B comparisons of designs. They enable a "guided tour" of the design and the resulting discussions of design goals. The fact that the perceptual tools are separate from the tools used to create a design (e.g. Adobe Illustrator) may be a benefit, as users can interact with them without fear of messing up the actual design [Henderson 1995], and can use them regardless of what design tools they prefer to use. Furthermore, by not making a rigid judgment in favor of one design over another, the saliency and clutter maps allow designers to easily trade off what the maps predict with other design concerns, and do not overly constrain the playful nature of the design process [Adamczyk and Twidale 2007; Bødker, Nielsen, and Petersen 2000].

7. PRINCIPLES FOR DESIGNING A PERCEPTUAL TOOL

In this section we discuss generally applicable principles for perceptual tools for aiding design teams. We have made changes to DesignEye corresponding to those principles.

7.1 Underlying Perceptual Models Need to Be Good, Not Perfect, But Manage Expectations

Though users often want a perfect oracle they can consult, they clearly can derive benefit from algorithms that can reasonably inform A/B comparisons and provide a sufficient visualization of the design to guide the design team through its analysis in a way that is fun rather than frustrating. The predictions, if good enough to be meaningful, can provide a common language for discussing the design in what one participant called "a less emotion-driven, opinion-driven" way. The human in the loop is quite capable of modulating the predictions when they seem to be incorrect or are in opposition to other design constraints.

Nonetheless, one might need to manage expectations, particularly through removing or downplaying more global "this design is better" predictions in favor of more qualitative representations of the design. Our global clutter values, for instance, overly encourage "oracle" thinking.

7.2 If Possible, Models Should Operate on Images (or Videos) of the Design

Many designs consist of a mix of LCD and physical components. Ideally, a perceptual tool should analyze a variety of the many artifacts created during the design process, from sketches, through story boards, to completed designs. Furthermore, many applications require an understanding of perception of a design under real-world lighting condi-

tions, e.g. glare, and surrounded by non-design surrounds like other products on a shelf, the view out a windshield, etc. In all of these cases, it is difficult to adequately capture the percept of a design without an image of how it will appear in the real world, as opposed to a mere design description. When possible, perceptual models should operate on images of the designs, as does DesignEye. Though perceptual models certainly exist which operate on descriptions of displays rather than on images [e.g. Wolfe, 2007] in general models that generalize to more complex and realistic displays operate on images.

If at all possible, tools should make predictions for dynamic displays. The current version of DesignEye operates only on static images. However, dynamic versions of a number of perceptual models – saliency, perceptual grouping, etc. – already exist [e.g. Paris, 2008] Since descriptions of such displays can be misleading, operating on actual video of the design in action is optimal.

7.3 Support the "Guided Tour" Use of Perceptual Tools

When possible, perceptual predictions should be in the form of a clear visualization of the design. Implicitly, one need is to facilitate understanding of the mapping between the original image and representations of the perceptual predictions. It is this mapping that allows one to read out the predictions about a particular region of the display: What is the most salient item in the display? Is the help-box too cluttered? (A related issue is discussed below, in the section "Support A/B Comparisons of Designs.") When transitioning back and forth between the predictions and the original design becomes difficult, this may be a sign that the visualization is too complex (or that the design itself is). For example, one could easily imagine from the simple example in Figure 2bc that for even a modestly complex design, perceptual grouping predictions might involve many groups and thus be quite complicated. Given that our study suggests that a low-information visualization of the design enables a guided tour of that design, a too-complicated visualization might thwart the benefits of perceptual modeling. Each new perceptual model requires careful thought about how best to present the predictions.

7.4 Support Crystallization of Goals

While our users typically understood the perceptual predictions, some users required training in thinking about the *desired* percept. This interfered with their ability to use DesignEye to clarify their design goals. In our experiments, simple instruction from us was sufficient, but ideally the tools should be useful without hands-on training. Documentation might help, if users read it. The tool itself might facilitate thinking about the desired percept by asking the user questions about the likely use of individual compo-

nents. In some cases it might be possible to generate a visualization of the desired percept based upon the answers to these questions.

7.5 Support A/B Comparisons of Designs

Another common use of DesignEye was to make direct comparisons between similar designs. Such comparisons aid in suggesting design improvements as well as participating in discussions about the optimal design.

The tool should enable easy comparison of a collection of similar designs and their associated perceptual predictions. Users need to see original designs alongside the predictions. We have added batch processing to DesignEye, which produces a document showing a number of designs alongside perceptual predictions. We have also enhanced on-screen comparison of designs, with added zoom functionality.

Furthermore, users often wanted to compare in more detail the predictions for particular regions of the design. For example, a user might want to know amount of clutter surrounding a particular component in two versions of the design. One user went to great lengths to compare individual regions by clipping them out by hand and pasting them against a black background. Isolating these regions against a clutter-free background greatly changes the predicted percept of those regions, invalidating predictions of their percept within the original design. If perceptual tools do not provide region-based computations, users will likely implement such computations themselves, and to do so correctly requires too much understanding of the underlying models. We added region-selection tools to DesignEye to enable more perceptually valid *in situ* region comparisons. Within a region, DesignEye computes average saliency and clutter. Selection regions from one design can easily be copied to other designs, to aid in comparing equivalent regions. For other kinds of perceptual models, tools for enabling A/B comparisons might be quite different. For instance, for perceptual grouping models, one might want to select a small region of the display, and show an entire grouping hierarchy (e.g. Figure 2c) for that region; a grouping hierarchy for the entire display would typically be intractably cluttered, but such a hierarchy might facilitate careful comparisons of specific regions of the display.

Enabling easy comparisons to some extent involves protecting users from misuse of the tool. Support for A/B comparisons should enable *trusted* comparisons, since such comparisons are often used in disagreements about the design. Furthermore, tools should make it difficult to inadvertently make unfair comparisons between two designs. For example, tools need to protect against users inadvertently comparing designs that are of different quality merely because they have been scanned in at different resolutions. Tools

need to discourage users from displaying perceptual predictions in a misleading way, whether intentionally or not. Our new version of DesignEye, for instance, by default autoscales the contrast of all open clutter and saliency maps for display, enabling fair comparisons across design options.

7.6 Support Cross-Disciplinary Design Discussions

In addition to the principles already mentioned, one can support discussion about designs by facilitating a user showing prediction results to a colleague. DesignEye now allows users to save a working session, so that another user can reopen it and see with their own eyes where the previous user left off, and can thus validate the perceptual predictions.

8. CONCLUSIONS

Perceptual aspects of design are important to the ultimate usability of the design. In the past, design teams have relied upon design guidelines, the expertise of the designer, or user studies, to ensure the perceptual "goodness" of the design. More recently, work in human and computer vision has made it possible to automatically make reasonably accurate perceptual predictions on arbitrary designs. But are such predictions actually useful to design teams?

We conducted field studies of a number of design teams, both examining their usual work practise and studying their use of our perceptual tool, DesignEye. We found hope that such tools will be useful, but not always for the reasons we expected.

We went into this study expecting that perceptual tools could help by giving guidance on perceptual aspects of design. This does seem one likely role for such tools. Design Eye did facilitate observations about the perceptual goodness of designs, useful A/B comparisons between candidate designs, and successful redesign.

One caveat, however, is that users may want these tools to provide absolute answers about what is the best design. However, this is unlikely to be the most useful role for these tools; "oracle" mode not only relies much more on tools' accuracy, but also reduces creativity, and may interfere with natural process of design team members balancing perceptual concerns with other design issues. Furthermore, such "oracle" use may render a perceptual tool less acceptable to professional designers. Where such issues are a concern, designers of perceptual tools may need to take steps to discourage such oracle thinking, such as eliminating overall "goodness" values, providing some indication of how certain the model is about a given prediction, or perhaps through setting expectations in marketing such tools and in training design teams to use them.

Perhaps the most powerful potential use of perceptual tools, though, is as a facilitator of conversations among multidisciplinary team members about design intuitions and de-

sign goals. By providing a common "perceptual" language, these tools may function as "boundary objects" [Star 1989], mediating communication across disciplinary boundaries. By providing a simple visualization of the design, thus providing a "guided tour" of the design, and facilitating conversations about goals, perceptual tools are like the "conscription devices" of [Henderson 1995], or "persuasive artifacts" of [Wagner 2000]; they enlist participation, elicit comments, and "recover problem-solving context."

One of course must always ask whether our the apparent positive effects of a guided tour and conversations about goals are a real effect of the tool, as opposed to effects of studying our users (a Hawthorne effect), or due to the novelty of the tool. A Hawthorne effect seems unlikely, as we studied our users examining their designs for several minutes without the presence of DesignEye, yet introduction of the tool led to more thoughtful analysis of their designs.

An effect of the novelty of the tool is more difficult to rule out without a controlled study of DesignEye vs. other tools, or vs. the common technique of "squinting" at a design to gain insight about it. In general, the ultimate test of perceptual tools like DesignEye is of course more extensive study of its use during the actual design process, over the course of months. We are working to establish relationships with beta testers in order to carry out such user studies.

Related to this question are potential issues arising from the fact that we were contacted by one member of each design team. This raises concerns that our sample may have been biased. At the very least, one might expect this to have affected interest or acceptance of our tool in the study. However, of the 35 people involved in our study of existing work practises, only 5 were individuals who contacted us. Of the 10 people involved in use of DesignEye, only 3 contacted us with interest in DesignEye. Furthermore, the fact that people thought use of a perceptual tool might help them with the design process is among the least interesting of our conclusions, and it is more difficult to see how details of interacting with such a tool would be biased by our selection process.

Based on what we have learned thus far, we have modified DesignEye. We have tried as much as possible to make DesignEye extensible to other perceptual models, so that effectively with this tool designers can keep their designs up to date with theoretical results from perception research, even if they are not focused on these results themselves. At present, DesignEye allows one, with a small modification to the menus, to essentially plug in any modules that take a single image as an input and output predictions in the form of one or more images. As mentioned in Section 2, many existing perceptual models have these characteristics. A complete abstraction between interface and algorithm is

not possible. Different perceptual "modules" may require different user input, different techniques for visualizing the predicted percepts, and different techniques for aiding A/B comparisons. Nonetheless, we have suggested principles important for any perceptual tool: Focus on qualitative rather than quantitative predictions; if possible, take images as input; provide a simple visualization of the predictions; make sure the users understand how to think about the *desired* percepts of their design; and support A/B comparisons and easily shared, trusted output from the model.

ACKNOWLEDGMENTS

REFERENCES

ADAMCZYK, P. D. & TWIDALE, M. B. 2007. Supporting multidisciplinary collaboration: requirements from novel HCI education. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA, April 28 – May 03, 2007). CHI '07. ACM, New York, NY, 1073-1076.

ADLER, A. & DAVIS, R. 2004. Speech and sketching for multimodal design. In *Proceedings of the 9th International Conference on Intelligent User Interfaces,* 214-216.

ANSTIS, S. 1998. Picturing peripheral acuity. *Perception, 27,* 817-825.

BALAS, B. J., NAKANO, L. M., ROSENHOLTZ, R. A statistical model of peripheral vision predicts visual crowding. *Journal of Vision*. *9*(12):13, 1-18, http://journalofvision.org/9/12/13/, doi:10.1167/9.12.13.

BØDKER S., NIELSEN, C., PETERSEN, M.G. 2000. Creativity, cooperation and interactive design. In *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practises, Methods, and Techniques* (New York City, New York, USA, August 17 – 19, 2000). D. Boyarski and W. A. Kellogg, Eds. DIS '00. ACM, New York, NY, 252-261.

BRAVO, M. J. & FARID, H. 2008. A scale invariant measure of clutter. *Journal of Vision 8*(1):23, 1-9. http://journalofvision.org/8/1/23/, doi:10.1167/8.1.23.

BUCCIARELLI, L. L. 1988. An ethnographic perspective on engineering design. *Design Studies 9*, 159–168.

BRUCE, N. D. B., & TSOTSOS, J. K. 2009. Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision*, *9*(3):5, 1-24, http://journalofvision.org/9/3/5/, doi:10.1167/9.3.5.

CLEVELAND, W. 1980. *The Elements of Graphing Data.* Wadsworth, Monterey, CA.

HAMMOND, T. & DAVIS, R. 2005. LADDER, a sketching language for user interface developers. *Computers and Graphics 28,* 518-532.

HENDERSON, K. 1995. The visual culture of engineers. In *The Cultures of Computing*, STAR, S. L., Ed. Wiley-Blackwell, NY.

HENDRY, D. G. 2004. Communication functions and the adaptation of design representations in interdisciplinary teams. In *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practises, Methods, and Techniques* (Cambridge, MA, USA, August 01-04, 2004). DIS '04. ACM, New York, NY, 123-132.

ITTI, L., KOCH, C., & NIEBUR, E. 1998. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence 20*(11), 1254-1259.

KOSSLYN, S. M. 1989. Understanding charts and graphs. *Applied Cognitive Psychology 3*, 185-225.

LANDAY, J. A. & MYERS, B. A. 2001. Sketching interfaces: Toward more human interface design. *IEEE Computer 34*(3), 56-64.

LOHRENZ, M.C. AND GENDRON, M. L. 2008. A 3D clustering algorithm to model clutter in electronic geospatial displays. *Journal of Management and Engineering Integration. 1,* 83-88.

MACKINLAY, J. D. 1988. Applying a theory of graphical presentation to the graphic design of user interfaces. In *Proceedings of the 1st Annual ACM SIGGRAPH Symposium on User Interface Software* (Alberta, Canada, October 17 – 19, 1988). UIST '88. ACM, New York, NY, 179-189.

NEWMAN, M. W., LIN, J., HONG, J. I. & LANDAY, J. A. 2003. DENIM: An informal web site design tool inspired by observations of practice. In *Human-Computer Interaction, 2003. 18*(3), 259–324.

PARIS, S. 2008. Edge-preserving smoothing and mean-shift segmentation of video streams. In *Proceedings of the 10th European Conference on Computer Vision: Part II* (Marseille, France, October 12 - 18, 2008). D. Forsyth, P. Torr, and A. Zisserman, Eds. Lecture Notes In Computer Science, vol. 5303. Springer-Verlag, Berlin, Heidelberg, 460-473.

TORRALBA, A., OLIVA, A., CASTELHANO, M., & HENDERSON, J. M. 2006. Contextual guidance of attention in natural scenes: The role of global features on object search. *Psychological Review. 113*(4), 766-786.

ROSENHOLTZ, R. 1999. A simple saliency model predicts a number of motion popout phenomena. *Vision Research, 39,* 3157-3163.

ROSENHOLTZ, R, 2001a. Visual search for orientation among heterogeneous distractors: Experimental results and implications for signal detection theory models of search. *J. Experimental Psychology, 27*(4), 985-999.

ROSENHOLTZ, R. 2001b. Search asymmetries? What search asymmetries? *Perception & Psychophysics 63*(3): 476-489.

ROSENHOLTZ, R., NAGY, A.L. & BELL, N.R. 2004. The effect of background color on asymmetries in color search. *Journal of Vision, 4*(3), 224-240.

ROSENHOLTZ, R., & JIN, Z. 2005. A computational form of the statistical saliency model for visual search [Abstract]. *Journal of Vision 5*(8), 777a.

ROSENHOLTZ, R., LI, Y., MANSFIELD, J., & JIN, Z. 2005. Feature congestion, a measure of display clutter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Portland, OR, USA, April 02 – 07, 2005). CHI '05. ACM, New York, NY, 761–770.

ROSENHOLTZ, R., LI, Y., & NAKANO, L. 2007. Measuring visual clutter. *Journal of Vision 7*(2):17, 1-22. http://journalofvision.org/7/2/17/, doi:10.1167/7.2.17.

ROSENHOLTZ, R., TWAROG, N. R., SCHINKEL-BIELEFELD, N., WATTENBERG, M. 2009. An intuitive model of perceptual grouping for HCI design. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems* (Boston, Massachusetts, USA, April 04 – 09, 2009). CHI '09. ACM, New York, NY, 1331-1340.

SHI, J. & MALIK, J. 1997. Normalized cuts and image segmentations. In *Proceedings of the 1997 IEEE Conference on Computer Vision and Pattern Recognition* (San Juan, Puerto Rico, June 17 – 19, 1997). CVPR '97. IEEE Computer Society, Washington, DC, 731-737.

SONNENWALD, D. H. 1996. Communication roles that support collaboration during the design process. *Design Studies 17*(3), 277-301.

SINHA, A. K. & LANDAY, 2001. J. A. Visually prototyping perceptual user interfaces through multimodal storyboarding. In *Proceedings of the 2001 Workshop on Perceptive User Interfaces* (Orlando, Florida, USA, November 15 – 16, 2001). PUI '01, vol. 15. ACM, New York, NY, 1-4.

STAGGERS, N. 1993. Impact of screen density on clinical nurses' computer task performance and subjective screen satisfaction. *International Journal of Man-Machine Studies, 39,* 775-792.

STAR, S.L. 1989. The structure of ill-structured solutions: Boundary objects and heterogeneous distributed problem solving. In *Readings in Distributed Artificial Intelligence 3*, HUBS, M. AND GASSER, L. Eds. Morgan Kaufmann, Menlo Park, CA.

TUFTE, E. R. 1990. *Envisioning Information.* Graphics Press, Cheshire, CT.

VAN DEN BERG, R., CORNELISSEN, F. W., & ROERDINK, J. B. T. M. 2009. A crowding model of visual clutter. *Journal of Vision*, *9*(4):24, 1-11,

WAGNER, I. 2000. Persuasive artifacts in architectural design and planning. In *Proceedings of CoDesigning 2000*. (Nottingham, UK, September 11 – 13, 2000). Springer-Verlag, New York, NY, 379-390.

WARE, C. 2004. *Information visualization: Perception for Design.* Elsevier, San Francisco.

WATTENBERG, M. & FISHER, D. 2004. Analyzing perceptual organization in information graphics. *Information Visualization, 3*, (June 2004), 123-133.

WHARTON, C., RIEMAN, J., LEWIS, C., & POLSON, P. 1994. The cognitive walkthrough method: a practitioner's guide. In *Usability Inspection Methods*, J. NIELSEN & R. MACK, Eds., 105-140.

WOLFE, J. M. 2007. Guided Search 4.0: Current Progress with a model of visual search. In W. Gray (Ed.), *Integrated Models of Cognitive Systems* (pp. 99-119). New York: Oxford.

ZHAOPING, L. 2002. A saliency map in primary visual cortex. *Trends in Cognitive Sciences, 6*(1), 9-16.