Editor: Theresa-Marie Rhyne

Understanding Visualization by Understanding Individual Users

Caroline Ziemkiewicz
Brown University

Alvitta Ottley, R. Jordan Crouser, and Krysta Chauncey
Tufts University

Sara L. Su
Google

Remco Chang
Tufts University

Visualization theory research has focused primarily on how to map data to visual forms and how people perceive them. This research has led to identifying fundamental principles regarding how humans perceive colors and visual patterns and to establishing general design guidelines for developing useful visualizations. Perceptual visualization theory attempts to understand and model how users perform fundamental low-level tasks.

However, as visualization gains widespread importance, researchers are studying more complex tasks. Visualizations are now serving as cognitive aids in problem solving, as users come to rely on visualizations to help them solve increasingly difficult problems. Color and perceptual theories remain necessary to make good design decisions but are insufficient by themselves to guide the design of a visualization for a cognitively complex task. These theories don't address how users think or how to apply visualizations as an extension of an individual's cognitive ability.

Clearly, we all think differently. You have aspects that differentiate you from everyone else. Your experiences, personality, and cognitive abilities influence your approach to performing a task and your understanding of a problem domain. Cognitive-psychology research has shown that such differences can significantly impact a user's dexterity with an interface or a tool.

Visualization users differ greatly in experiences, backgrounds, personalities, and cognitive abilities, yet visualizations, like much other software, continue to be designed for a single ideal user. Clearly, designing each visualization for an individual user would be impractical. However, knowledge of broad differences between user groups could help guide design for specific domains and help suggest multiple analysis modes or customization options in a single system. Recently, a promising research area has emerged that takes an opposing approach to traditional "one size fits all" design. This research suggests that the individual users' cognitive style, as much as the visual design, determines the visualization's value.

Although these findings are still at an early stage, they suggest that we shouldn't study visualization in a vacuum but in the context of differences among its users. On the basis of our experiences in studying those differences, we argue that current visualization theory lacks the necessary tools to analyze which design factors lead to differences in user behavior. Developing this understanding would enable researchers to study visualization from the perspective of how an analysis process arises from the interaction between a user and a system. This in turn could lead to a shift in how we evaluate and design visualizations for different user groups, tasks, and domains. For this to happen, we must first understand what individual factors (that is, cognitive and personality factors) affect visualization use.

 Published by the IEEE Computer Society

Cognitive Factors

Cognitive factors such as perceptual abilities, spatial abilities, verbal ability, and working-memory capacity vary substantially between individuals and can affect reasoning in many ways. In particular, perceptual and spatial abilities have been shown to affect how well users perform different tasks in a visualization system. Maria Velez and her colleagues first showed that several of these abilities, including spatial orientation, spatial visualization, visual memory, and perceptual speed (the speed at which a person compares images), affected accuracy and response time on a task involving the comprehension of 3D views similar to those found in scientific visualization applications.[1] Although this is perhaps unsurprising, subsequent research showed that these abilities also affected more abstract 2D visualization tasks.

Perceptual abilities include basic visual proficiencies such as scanning speed and visual-memory capacity. For example, Cristina Conati and Heather Maclaren found that perceptual speed correlated with a user's accuracy at information retrieval tasks in one of two visualization systems: a star graph and a heat-map-like view.[2] In one task, participants compared differences over time between two scenarios at a global level. Users with high perceptual speed performed better with the heat-map-like view; users with low perceptual speed performed better with the star graph. But this wasn't true for the other tasks. This task was perhaps the most complex one the participants had to perform; most of the other tasks involved retrieving or comparing a specific variable value. That the most complex inferential task was the most susceptible to individual differences is notable; our research (which we describe later) revealed similar effects.

We can measure spatial ability by a variety of tests, which might express different aspects of this factor. Generally, however, it refers to the ability to accurately reproduce and manipulate spatial configurations in working memory. Cheryl Cohen and Mary Hegarty found that a user's spatial ability affected the degree to which interacting with an animated visualization helped him or her perform a mental rotation.[3] Participants drew cross-sections of a complex 3D object. They could control two animated rotations of the object to complete the task. Participants with high spatial ability produced more accurate cross-sections and used the visualizations more; those with low spatial ability rarely discovered the best view from which to create the cross-section.

Similarly, Chaomei Chen and Mary Czerwinski found a relationship between spatial ability and the visual search strategies users employed in a network visualization.[4] Participants viewed an interactive node-link visualization of a paper citation network and had to find papers on specific topics. Spatial ability generally correlated positively with search task performance and predicted the use of a better navigation strategy. Low-spatial-ability participants were more likely to click through every node in a cluster even after determining that the cluster was irrelevant to the target topic. High-spatial-ability participants pursued a more hierarchical strategy in which they jumped from cluster to cluster until they found a likely neighborhood.

We argue that current visualization theory lacks the necessary tools to analyze which design factors lead to differences in user behavior.

Notably, two of these findings showed that differences in spatial ability affected not just overall performance but also how users approached a task. In the cross-section study, high-spatial-ability participants were more likely to seek out an optimal view for cross-sectioning; in the network visualization study, they employed a more hierarchical search strategy. The use of different strategies by users with different cognitive profiles suggests that, when user characteristics vary, there's no single way for a visualization to best support a given task. If people with varying cognitive abilities employ different strategies for the same task, a visualization designed for that task must take this into account to be effective.

Further research is needed to elaborate on these studies' implications. However, they suggest that at least some factors of cognitive ability affect the strategies people use in visualization tasks. What remains is to fully characterize these strategies and to be able to predict when differences will arise.

Spatial ability is perhaps the natural first individual difference to study in visualization, but it's not the only one that has an effect. As Ji Soo Yi recently argued, a better understanding of individual factors beyond basic spatial ability might be necessary to understand the variability in visualization evaluation.[5] Yi suggests further factors for study, including visual literacy and personality factors such as field independence and openness to experience. Although this area needs more research, several experimental results already show that users' personality differences can significantly influence visualization use.

Personality Factors

Personality psychology is a well-established research area, making it a useful lens through which to better understand how different users approach visualization tasks. A common model in personality psychology, the *five-factor model*, categorizes personality traits on five dimensions: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. Studies show that a person's traits remain consistent throughout adulthood.[6]

Some human-computer interaction (HCI) research has shown that these personality factors correlate significantly with a user's preference for visual-interface designs. For example, Batul Saati and colleagues compared preferences for five "skins" (visual themes) for a music player interface.[7] The skins varied only in the dominant color. Introverted users preferred blue, more conscientious users preferred yellow, and more imaginative users preferred black. Although user preference might affect adoption rates for a visualization system, performance differences are more valuable in understanding how people use visualization.

The examination of how personality affects performance remains relatively sparse in HCI research. This situation reflects a commonly held, if implicit, assumption that personality is more superficial than cognitive ability and therefore unlikely to affect much beyond surface reactions to design. Nonetheless, such studies, including those summarized in an early review by Kym Pocius, have suggested that personality can affect interface use at a deeper level.[8] For example, Pocius's meta-analysis indicated that introversion consistently correlates positively with both programming ability and performance in computer-assisted-instruction tasks.

In visualization specifically, recent research has shown that personality traits can significantly affect complex-task performance. For example, Caroline Ziemkiewicz and Robert Kosara investigated how conflicting metaphors affect tree visualization evaluations.[9] By varying verbal and visual metaphors in the evaluation conditions, they studied the extent to which users slowed down in response to metaphor conflicts. Users who scored highly on the openness dimension were unaffected by conflicting verbal and visual metaphors. This study also found a similar effect for spatial ability, but the two factors were independent of one another. This study didn't directly compare performance on different types of visual design. However, it does suggest that participants with high openness and spatial ability might more easily switch between design metaphors, such as those in a multiview system.

Tera Green and Brian Fisher[10] studied the use of visual-analytics interfaces by users with varying scores on the five-factor model and a personality dimension called *locus of control*.[11] Locus of control measures how much a person sees himself or herself being in control of events (internal locus of control), as opposed to being controlled by outside factors (external locus of control). The study compared two complex, dissimilar information retrieval systems—a visual-analytics system and a Web interface with a more list-like view. Users with a more external locus of control performed better at complex inferential tasks when using the visual-analytics interface. The authors also discovered additional correlations between neuroticism and task performance.

Building on Green and Fisher's research, we've conducted studies to identify visual elements that appear to be stronger classifiers of users.[12] We aimed to identify the design factors responsible for the reported results. We hypothesized that the layout's underlying metaphor was the most significant factor. So, we studied performance on four simple visualizations (see Figure 1) that were similar in all aspects except for their overall layout style.

The four views progressed from a list metaphor to a containment metaphor. We first measured the participants for locus of control and other personality factors, and then performed a series of search and inferential tasks similar to those that Green and Fisher used. For the inferential tasks, participants with an internal locus of control showed increased performance as the views became more list-like. Their accuracy increased by up to 70 percent (from 40 to 68 percent correct), and their response time improved by 34 percent (from 338 to 222 seconds), compared to using the containment view. Participants with an external locus of control showed less difference in performance overall but were slightly more adept with the most container-like view. Like Green and Fisher, we found this effect in complex tasks but not simple search tasks.

These studies suggest that personality differences might account for some of the observed individual variability in visualization use. However, this relationship isn't straightforward. Performance differences based on personality factors appear to manifest for tasks requiring inference and metaphorical reasoning. It's under these cognitively demanding situations that visualization will likely be the most valuable.

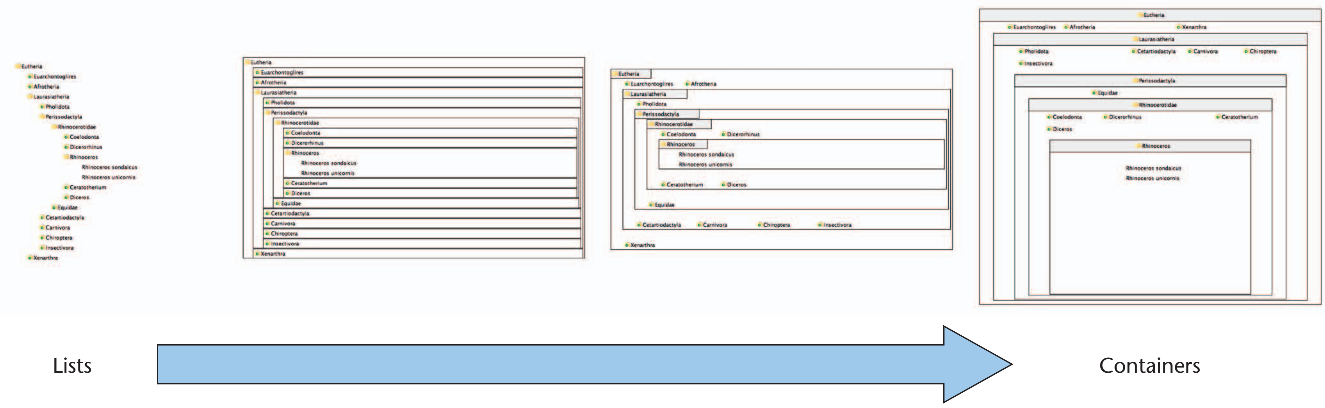

Figure 1. Our study on locus of control used these four visualizations, which progressed from a list metaphor to a container metaphor.[12] Participants with an internal locus of control performed progressively worse (in both accuracy and speed in solving complex analytical tasks) as the metaphor shifted from list to container. Those with an external locus of control were adept with all views, but especially with the most container-like view.

Relating Individual Factors to Design

To generalize from the findings we've described, we must isolate the visualization factors and evaluate which ones are helpful or harmful to a user with a given cognitive profile. In turn, we must identify relationships between two primary sets of factors:

- the individual factors describing the user and
- the design and structural factors describing the visualization.

Regarding individual factors, a large body of established research in psychology exists, but little agreement on which factors are most relevant to visualization. Regarding design factors, we have no real standard language to use when decomposing a visualization. In both cases, we must identify a set of reliable, measurable factors to identify useful correlations.

Individual Factors

Research has already given some indications as to which personality factors are most relevant to visualization use: spatial ability, extraversion,[8] and locus of control. This work is just the beginning, however. To identify the relevant individual factors, we need studies that both confirm the factors already found and investigate new factors. For example, although many HCI studies have found that extraversion is significant, it hasn't shown an effect in any of the visualization studies in which it has appeared. Is this due to inherent differences between visualizations and other interfaces, differences in the tasks being studied, or just differences in methodology or study population?

Answering these questions will require both experiments that examine a broader array of individual factors and experiments that study known factors more deeply. To focus this research agenda, a first step might be a formal meta-analysis of the existing findings. This would demonstrate which factors have the most consistent effects. It would also be a useful way to find connections between the research on individual differences in visualization and broader HCI research on this subject. Also important is establishing benchmark tasks and datasets so that we can compare different studies' findings more directly. Progress in this area will not only help focus the set of factors we study in any given experiment but also produce information about which aspects of the individual user are significant for visualization. A more difficult question is how to uncover information about which aspects of a visualization are significant to the individual user.

Design Factors

Much research in individual differences in visualization has been case based and domain specific. Although the previously cited studies found individual differences in performance, it's unclear whether those differences are inherent to all visualization use or reliant on aspects of the specific systems the researchers used. It would be helpful to analyze these systems to discover how they vary. Unfortunately, there's no standardized set of dimensions on which to analyze, let alone synthesize, visual designs.

This proved a challenge when we designed our locus-of-control study.[12] Much previous research has aimed to find individual differences in the use of a single visualization system or in one or more real-world systems that differ in many respects. In contrast, we aimed to take a known individual difference in performance and isolate the design factors that influence it. So, we needed to create a set of visualizations that systematically varied in a particular visual quality: in our case, whether participants used the list or container metaphor.

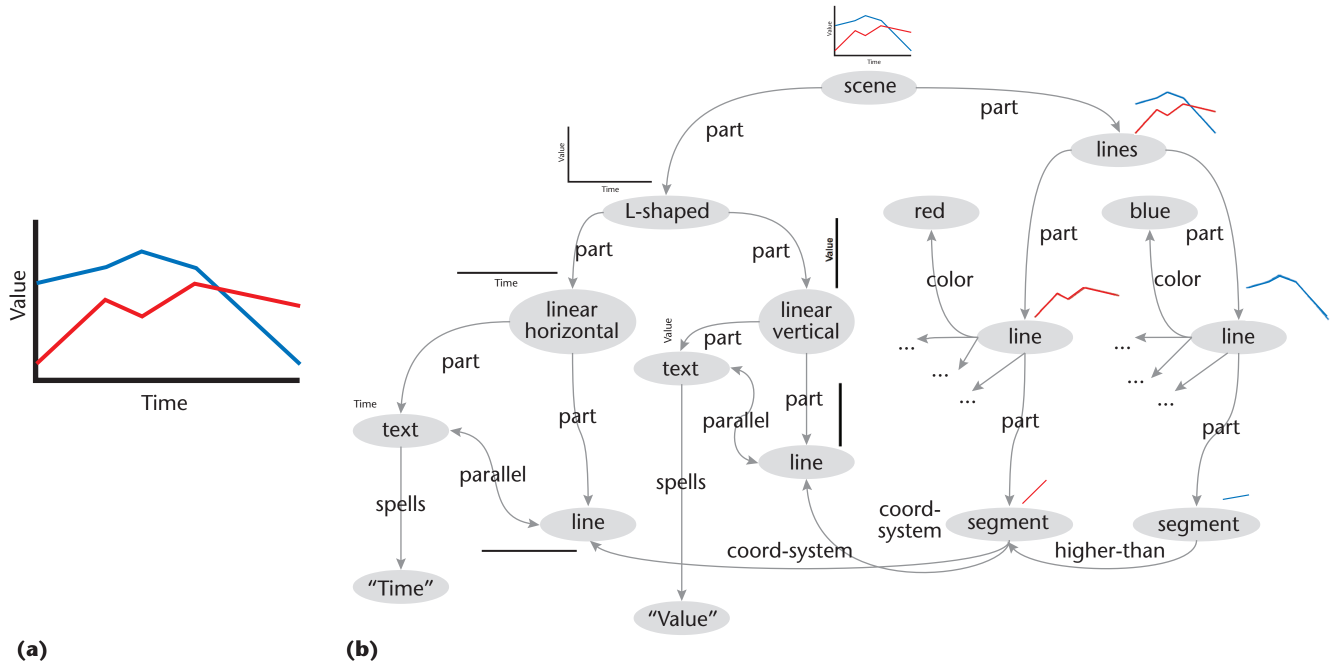

Figure 2. An illustration derived from Steven Pinker's decomposition theory.[14] (a) A simple line chart. (b) A decomposition of its parts. The decomposition includes visual mappings, perceptual qualities, and structural elements such as axes and labels.

We arrived at the solution in Figure 1 through trial and error. Although it was sufficient to isolate the effect we were testing, a more controlled approach and common language would more quickly advance this field. This requires some way to classify a visualization on the basis of its design factors.

Many high-level classifications of visualization types exist, but they don't offer the level of detail needed to isolate design variables. Classic visualization taxonomies are often based on data variables, such as dimensionality and data type (for example, categorical, ordinal, or numerical). Other taxonomies, such as the one Stuart Card and Jock Mackinlay developed in their research on the visualization design space,[13] rely on describing the mapping between data variables and visual variables such as color or position.

Although such research is useful in describing the kinds of data a visualization can depict, it's limited regarding describing factors of the visual design itself. This makes it difficult to isolate the factors that cause performance differences for varying user types. For example, visualization theory has no language to easily describe the differences between the visual designs we studied in relation to locus of control. The visualizations are the same in terms of basic visual mapping; the significant design differences are at a structural or metaphorical level. Card and Mackinlay's descriptions of visual mappings generally denote these properties with an asterisk indicating a special case. No systematic way exists to fit structural design differences into a visual mapping schema.

What's lacking is a usable decomposition of visualization design. To correlate individual factors and design factors, we need to know what those design factors are and be able to manipulate them in a controlled way. This means being able to take a single visualization and reliably analyze its components and how they relate to one another. This approach is closer to the one that Steven Pinker proposed.[14] He aimed to represent charts in a way that would make it possible to computationally model chart comprehension. The result is a decomposition that includes visual mappings, perceptual qualities, and structural elements such as axes and labels in a single graph (see Figure 2).

Pinker based his decomposition theory on simple, static charts; it's unclear how to extend it for more complex situations such as interactive visualizations or multiple linked views. As research shows, it's in these more complex situations that individual differences in visualization users will most likely arise. Nonetheless, Pinker does offer a model for a more comprehensive analysis of visualization designs. An abstract representation of a visual design produced by decomposition analysis could be measured and analyzed more quantitatively, producing metrics researchers can usefully correlate to individual personality factors.

Putting It All Together

After building a rich taxonomy of design factors that interact with various personality traits and a good understanding of which traits are significant for visualization use, we'll be poised to run valuable experiments to determine how individual differences affect visualization use. Such experiments could examine how other factors such as data and task complexity play into individual performance differences. This experimental toolkit could also form the basis for deeper questions about how people make sense of visual information under varied circumstances. This is an ambitious long-term research agenda, but its results could transform our understanding of visualization.

Toward Adaptive Interfaces

Our formal model of visualization must incorporate models of individual users, their personality profiles, and their situational strategies. Many rich areas for exploration accompany this ideological shift. For example, because the possible combinations of personality traits are functionally limitless, designing for the user as an individual inherently demands developing and adopting adaptive interfaces. That is, the interfaces we develop should learn about the user as an individual. They should then adjust to best support the unique combination of personality factors expressed in the user at that time. Such personalized interfaces aim to enhance an individual user's strengths and address individual weaknesses and have been extensively studied in HCI. Combining this existing research with the knowledge of users we gather in visualization studies will enable us to tune a visualization interface in accordance with the principles uncovered by this research.

Adapting visualizations to broad classes of users is a valuable design strategy. However, it's impractical to subject every user of a real-world system to the kind of multiple-choice personality inventories or tests of cognitive ability used in the experiments we described. In lieu of laborious tests, we propose building a model of a user's personality and cognitive ability by analyzing his or her interaction history.

Although this area has seen little research, Iftikhar Khan and his colleagues demonstrated significant correlations between interaction measures in a programming task and several personality measures, including those in the five-factor model.[15] For example, they found a negative correlation between openness and the number of times a participant switched between windows. Because this dimension also correlated positively with the length of time between interaction events, this result suggests that more open participants spent more time in each window. Findings such as these, extended to visualization-specific tasks, could form the basis for a model of user personality based on tracked interactions. Such a model of usage patterns will let us extrapolate a user's cognitive profile and adapt the visual design accordingly.

The challenges to designing for individuals are great, but the potential benefits make this a challenge worth pursuing. At the individual level, we each stand to benefit from systems that improve our efficiency and accuracy. On the more global scale, many marginalized and traditionally underserved user groups stand to benefit from increased access to visualization systems tailored to them, rather than those designed for only the average user. Finally, this research will result in a much deeper understanding of how users make sense of visual information.

Visualizations are tools for thinking, and we can't understand visualization until we understand what people do with those tools. Understanding that there's no one answer to that question is an important step toward truly understanding visualization.

References

1. M.C. Velez, D. Silver, and M. Tremaine, "Understanding Visualization through Spatial Ability Differences," *Proc. 2005 IEEE Conf. Visualization* (VIS 05), IEEE, 2005, pp. 511–518.
2. C. Conati and H. Maclaren, "Exploring the Role of Individual Differences in Information Visualization," *Proc. Working Conf. Advanced Visual Interfaces* (AVI 08), ACM, 2008, pp. 199–206.
3. C.A. Cohen and M. Hegarty, "Individual Differences in Use of External Visualisations to Perform an Internal Visualisation Task," *Applied Cognitive Psychology*, vol. 21, no. 6, 2007, pp. 701–711.
4. C. Chen and M. Czerwinski, "Spatial Ability and Visual Navigation: An Empirical Study," *New Rev. of Hypermedia and Multimedia*, vol. 3, no. 1, 1997, pp. 67–89.
5. J.S. Yi, "Implications of Individual Differences on Evaluating Information Visualization Techniques," *Proc. BELIV '10 Workshop*, ACM, 2010.
6. B.W. Roberts and W.F. DelVecchio, "The Rank-Order Consistency of Personality Traits from Childhood to Old Age: A Quantitative Review of Longitudinal Studies," *Psychological Bull.*, vol. 126, no. 1, 2000, pp. 3–25.
7. B. Saati, M. Salem, and W.-P. Brinkman, "Towards Customized User Interface Skins: Investigating User Personality and Skin Colour," *Proc. 11th Int'l Conference Human-Computer Interaction* (HCI 05), Lawrence Erlbaum Associates, 2005, pp. 89–93.

8. K.E. Pocius," "Personality Factors in Human-Computer Interaction: A Review of the Literature," *Computers in Human Behavior*, vol. 7, no. 3, 1991, pp. 103–135.
9. C. Ziemkiewicz and R. Kosara, "Preconceptions and Individual Differences in Understanding Visual Metaphors," *Computer Graphics Forum*, vol. 28, no. 3, 2009, pp. 911–918.
10. T. Green and B. Fisher, "Towards the Personal Equation of Interaction: The Impact of Personality Factors on Visual Analytics Interface Interaction," *Proc. 2010 IEEE Symp. Visual Analytics Science and Technology* (VAST 10), IEEE CS, 2010, pp. 203–210.
11. J.B. Rotter, "Generalized Expectancies for Internal versus External Control of Reinforcement," *Psychological Monographs*, vol. 80, no. 609, 1966, pp. 1–28.
12. C. Ziemkiewicz et al., "How Locus of Control Influences Compatibility with Visualization Style," *Proc. 2011 IEEE Conf. Visual Analytics Science and Technology* (VAST 11), IEEE, 2011, pp. 81–90.
13. S.K. Card and J. Mackinlay, "The Structure of the Information Visualization Design Space," *Proc. IEEE Symp. Information Visualization*, IEEE, 1997, pp. 92–99.
14. S. Pinker, "A Theory of Graph Comprehension," *Artificial Intelligence and the Future of Testing*, R.O. Freedle, ed., Lawrence Erlbaum Associates, 1990, pp. 73–126.
15. I.A. Khan et al., "Measuring Personality from Keyboard and Mouse Use," *Proc. 15th European Conf. Cognitive Ergonomics* (ECCE 08), ACM, 2008, article 38.

***Caroline Ziemkiewicz** is a research engineer at Aptima. This work was performed while she was a postdoctoral researcher at Brown University. Contact her at cziemkiewicz@aptima.com.*

***Alvitta Ottley** is a PhD student in computer science at Tufts University. Contact her at alvittao@cs.tufts.edu.*

***R. Jordan Crouser** is a doctoral candidate in Tufts University's Computer Science Department. Contact him at rcrous01@cs.tufts.edu.*

***Krysta Chauncey** is a postdoc in computer science at Tufts University. Contact her at krysta.chauncey@tufts.edu.*

***Sara L. Su** is a product manager at Google. Contact her at sarasu@google.com.*

***Remco Chang** is an assistant professor in Tufts University's Computer Science Department. Contact him at rchang@cs.tufts.edu.*

Contact department editor Theresa-Marie Rhyne at theresamarierhyne@gmail.com.