Figured Foundation
Communicating science through design.

We propose to launch Figured Foundation, a science & arts nonprofit that helps to make science communication more beautiful, lucid, and intelligible by means of thoughtful design.

FIGURED'S APPROACH

Science is communicated visually. The most recent issue of *Nature*, a leading journal of science research, contains 16 articles with a total of 266 graphs and other visualizations, for an average of 17 plots per paper[1]. These are only a few of the 1.5 million research articles published per year,[2] and data visualization is not limited to papers. In 2012, at the world's largest scientific conference, the annual meeting of the *Society for Neuroscience*, there were over 16,000 poster presentations and powerpoint talks[3], each filled with plots and figures meant to convey the latest findings about the brain and other sensory systems. Scientists use visualization as a core method to explore their data, to extract meaning from it, and to share the results publicly. We estimate that there are 10 to 100 million scientific visualizations created per year.

As a skill, visualization is essential to scientific thinking. Leonhard Euler's development of a visualization to help explain the Königsberg Bridge Problem led to rapid advancements in graph theory and our understanding of networks[4]. John Snow's creative use of mapping to track cholera cases in London was essential to our understanding of the spread and prevention of disease[5].

The era of big data has made visualization all the more fundamental to scientific progress. The contemporary explosion of new media has enabled the production of unprecedented amounts of data. Understanding pressing topics like climate science demands that scientists work together with designers to develop new solutions to visualizing data.[6] In the neurosciences, imaging and multi-electrode recording techniques have made it possible to simultaneously monitor the activity of hundreds or even thousands of neurons.[7] The limiting factor in these projects is no longer the biological techniques; instead, it is the quality of the visualization tools used to parse, explore, and understand these data.

---

[1] http://www.nature.com/nature/journal/v493/n7433/index.html

[2] Björk, B. C., Roos, A., & Lauri, M. (2008, December). Global annual volume of peer reviewed scholarly articles and the share available via different Open Access options. In *ELPUB2008*.

[3] http://www.sfn.org/am2012/index.aspx?pagename=press_room

[4] Barabasi, A. L. (2003). Linked: How everything is connected to everything else and what it means.

[5] Tufte, E. R., & Weise Moeller, E. (1997). *Visual explanations: images and quantities, evidence and narrative* (p. 12). Cheshire, CT: Graphics Press.

[6] e.g., see http://ecohacksf.org/

[7] Ahrens, et al., (2012) Brain-wide neuronal dynamics during motor adaptation in zebrafish. *Nature*, 485, 471–477. http://www.nature.com/nature/journal/v485/n7399/abs/nature11057.html

Despite the importance, ubiquity, and pressing need for visualization, there is a lack of dialogue in the sciences about visual thinking. Choices of color, layout, and design are seen as mere aesthetic styling, though in fact they are essential tools for revealing patterns in data. Scientists who develop new methods for data collection have limited time to learn new methods for visualization. Improving the visualization of scientific data provides an opportunity to advance scientific thinking, to make findings more accessible, and to reboot scientific publishing for a new generation steeped in visual media.

This is where *Figured* arrives: we want to make the communication of science more beautiful, lucid, and intelligible by means of thoughtful design. Of course, we are not the first to have considered the problem of imbuing scientific communication with good design, but our approach is unique. By identifying the causes of failure of the two dominant techniques, we were able to remix them to create a new approach that can succeed:

The dominant approach to promoting good design in scientific visualization is to inform through advice and example. The most famous proponent of this approach is statistician and data visualization expert Edward Tufte, whose series of books, beginning with the 1982 classic *The visual display of quantitative information*[8], showed the world what well-designed data looks like. The book has been cited over 6,000 times and is read by millions. Other instances of this approach include blogs that highlight excellent designs (e.g., infosthetics.com, flowingdata.com, visualizing.org, and informationisbeautiful.net), professors and designers who write articles and books (e.g., Kosslyn's book *Graph design for the mind and eye*, Wong's column in *Nature Methods*), and visualization experts who give seminars (e.g., Martin Krzywinski).

The other approach to promoting good design in scientific visualization is to create tools. This is the approach used by software companies (e.g., SigmaPlot, Excel, MATLAB, Mathematica) and the various visualization groups and research communities that invent new techniques, environments, and frameworks for visualizing data (e.g., Ben Fry's *Processing*, Mike Bostock's D3.js, Martin Krzywinski's Circos, and the myriad techniques and tools presented each year at the annual *IEEE InfoVis* conference).

Each of these two dominant approaches — advising and retooling — is deeply flawed when used in isolation.

Advice, though it may be principled, thoughtful, and worthy of following, is hard to apply in a new setting once it has been learned[9]. Changing one's workflow to incorporate a new habit is a slow, costly, and effortful process[10]. This difficulty of incorporating advice into one's workflow limits the utility of that advice. For example, scientists may know the principle that good

---

[8] Tufte, Edward R., and P. R. Graves-Morris. *The visual display of quantitative information*. Vol. 31. Cheshire, CT: Graphics press, 1983.

[9] Argote, L., & Ingram, P. (2000). Knowledge transfer: A basis for competitive advantage in firms. *Organizational behavior and human decision processes*,*82*(1), 150-169.

[10] http://captology.stanford.edu/projects/behaviordesign.html

visualizations have a high "data to ink ratio" (less clutter, more data) and may have even seen plots that exemplify the principle, but may still not know how to recognize clutter in their own graphs or how to use their tools to reduce the clutter once it is recognized. The result is visualizations that don't follow the known principles.

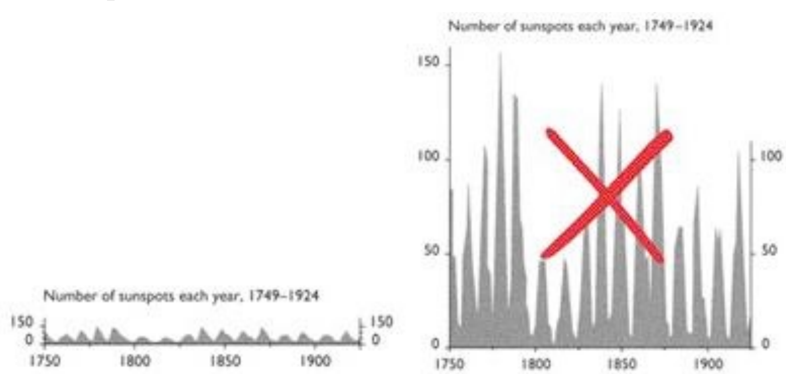
Tools suffer from a different problem. The most popular tools, such as MATLAB and Excel, have default behavior that produces notoriously bad design, requiring an expert to wrangle a serviceable plot (e.g., see excelcharts.com/blog/examples/). On the other hand, powerful tools that encourage good design and enable new forms of exploration and interactivity (e.g., Circos or D3) require the user to learn an unfamiliar programming language and put the onus on the user to use the tools correctly to produce a successful visualization.

We have developed a method that circumvents these problems by remixing the two dominant approaches. It's resulted in *Figured*'s guiding principle: *smart defaults*. We think the best way to make science communication more beautiful, lucid, and intelligible is first to explain the principles, and then immediately provide users with a way to modify their existing workflow such that it automatically follows the new principle, by default. This technique has the same benefits of advising and retooling, but does not suffer the problems of costly and effortful adoption. Successful adoption of smart defaults can serve as a critical stepping stone by introducing scientists to cutting-edge techniques in the context of a tool they already know.

In the next section, we describe a case study showing how our method can be applied to a recently-developed design technique, leading to its widespread adoption in a situation where the two dominant approaches failed.

CASE STUDY: INCORPORATING *MULTI-SCALE BANKING TO 45°* INTO A VISUALIZER'S WORKFLOW

In 2006, Jeffrey Heer and Maneesh Agrawala, visualization researchers at UC Berkeley, introduced a new technique[11] based on William Cleveland's 1993 observation that trends in a time series are most clearly seen when the average tilt of the lines on the page is 45°. This technique was described in Edward Tufte's *Beautiful Evidence*, where he shows two variants of the same plot with different aspect ratios:



[11] Heer, J., & Agrawala, M. (2006). Multi-scale banking to 45 degrees. *Visualization and Computer Graphics, IEEE Transactions on*, *12*(5), 701-708.

On the right, the plot is square and the average tilt is very steep. On the left, the plot is improved by squashing the vertical axis such that the average tilt is 45° — this is called *banking*. Banking makes it easier to see the ups and downs of each sunspot cycle, revealing patterns that were hidden in the original plot. Heer & Agrawala's technique builds on banking by directly analyzing the raw data, determining the dominant trends in the data, and then building plots that are tuned to optimally visualize each of these dominant trends.

The technique is simple, powerful, and beautiful, but it has not seen widespread adoption despite its 6 years of existence since publication. We attribute this to the vast gap between understanding the rule and incorporating it into the scientist's existing workflow.

Here's what Figured would do. We solve this problem by releasing a tutorial that describes multi-scaling banking, and at the end of the tutorial, provide scripts that change the default behavior of the major plotting tools such as MATLAB, Excel, R, and Mathematica such that they automatically perform multi-scaling baking every time a plot is created — smart defaults. Our method provides a vast improvement in outcome with minimal adoption costs. We bundle our scripts together into toolboxes such that a newcomer receives the benefits of all previous tutorials. Critically, our method allows quick adoption of the latest techniques, requiring only one motivated implementer.

This case study, where we applied our principle of smart defaults to multi-scale banking to 45°, is an example of our approach and can be scaled to a wealth of other design principles and new techniques: using sensible color schemes, eliminating clutter and chart junk, using small multiples, incorporating simple forms of interactivity[12].

THE TEAM

Our team has wide-ranging professional experience in science, data visualization, perception, and design:

Bang Wong, MS (bang.clearscience.info, @bangwong) is the creative director of the Broad Institute of MIT and Harvard and an adjunct assistant professor in the Department of Art as Applied to Medicine at the Johns Hopkins University School of Medicine. His work focuses on developing visual strategies to meet the analytical challenges posed by the unprecedented scale, resolution, and variety of data in biomedical research. Bang also writes a monthly column for *Nature Methods* that deals with the fundamental aspects of visual presentation.

Jordan Suchow (jordansuchow.com, @suchow) is a graduate student of cognitive science in the Department of Psychology at Harvard, where he studies minds, brains, machines and how they process visual information. His research involves building computational models of human memory and perception and using the results to inform the design of information displays.

---

[12] Tufte, E. R., & Graves-Morris, P. R. (1983). *The visual display of quantitative information* (Vol. 31). Cheshire, CT: Graphics press.

Benjamin Golder (benjamingolder.com, @bgolder) is a graduate student at MIT in the departments of Urban Studies and Planning, and Architecture. Ben has taught creative applications of technology in design for the past five years, most recently at Architecture Association Visiting Schools in Tel Aviv, Mexico City, and San Francisco, and as a lecturer in the Architecture Department at UC Berkeley and in the Industrial Design Department at California College of the Arts.

Chana Haouzi (linkedin profile) is a graduate student in the Graduate School of Design at Harvard, where she studies architecture and design. Chana graduated from McGill University with a masters degree in architecture. She taught digital representation and freehand sketching at McGill and teaches CNC milling at Harvard.

Jeremy Freeman, PhD (jeremyfreeman.net, @istudyvision) is a neuroscientist at NYU and HHMI's Janelia Farm, studying how animals use their brains to perceive and interact with the world around them. Jeremy spent his graduate work at NYU studying the primate visual system. He loves wrangling complex data sets into clean, effective, and interactive visualizations.

Our team has a flat organizational structure. Everyone on the team has existing professional and long-term interests in the mission of the organization. We are designers, scientists, and visual communicators.

CORE PRODUCTS, PROGRAMS, AND SERVICES

The following sections provide an overview of the implementation details of Figured, describing the means by which we will attack our goal of making science communication more beautiful, lucid, and intelligible by means of thoughtful design. Figured's initial content will appear on our website. Critically, we have identified the following key areas where our technique of smart defaults can be leveraged to produce the biggest possible impact on the quality of design in scientific communication:

Toolboxes that enhance standard tools with smart defaults

Our first project will be to develop toolboxes that provide smart defaults for MATLAB, R, Mathematica, and Excel. These are the most common tools used to create scientific plots and for which changing the defaults would yield the largest impact. Our team has the design expertise (Bang, Chana, Ben) and technical competence (Jordan, Jeremy, Ben) needed to produce these toolboxes. We have created an alpha version of the MATLAB toolbox. We estimate that its completion will require approximately three months, and that once complete porting the code to R, Mathematica, and Excel will take an additional 3–6 months. Full funding would allow us to commission advanced features. For example, after building the basic toolboxes, we can leverage our userbase's comfort with our tools to expose them to more advanced visualization techniques. We plan to add to the MATLAB toolbox software that makes it easy for current MATLAB users to export and visualize their data using web-based interactive tools (like D3.js). For many scientists, this kind of push could be the jumping off point for launching an entirely new exploratory data analysis project. When we release any new feature for our toolbox, we will write a blog post on the design principles that informed the feature, while also pointing the user to other resources that they can use to learn more.

Rewards and potential impact: These toolboxes have the potential to transform the quality of visualization in science by nearly eliminating poor design, and encourage a generation of scientists to incorporate new techniques and interactivity into their designs. Together, the toolboxes could saturate the target audience of approximately 5 million users. As a digital asset, the toolboxes scale easily, with essentially no added cost per user. Impact can be measured by the number of downloads and by the number of published or presented visualizations relying on our toolboxes.

Risks: Our potential for impact is capped by the number of users who install the toolbox and find it worthwhile. There are two potential pitfalls that we may encounter in acquiring new users. First is the problem of compatibility: our toolboxes must function as drop-in replacements for the existing functionality, without breaking the user's workflow; this would lead to frustration and distrust, which is particularly important to avoid when dealing with data — trust is everything. The best way to mitigate this risk is through extensive beta testing with our target users and by collaborating with people who are expert developers with each of the tools (e.g., a MATLAB developer). Full funding would allow us to commission targeted testing and code review by these expert developers. The second risk relates to customer support. As a small team, it is beyond our capabilities to provide email and phone support for each user. This may lead to dissatisfaction when something goes awry. To mitigate this risk, we will borrow a technique from successful software companies: providing FAQs, online forums, and user-groups where users of the toolboxes can ask questions and answer those of other users. The Meteor group (meteor.com) has successfully used Stack Overflow for this purpose. We hypothesize that these risk-mitigation techniques are sufficient, and we will test this through targeted follow-up emails with randomly-selected users; we will ask whether they are satisfied with the product and whether they have ever run into problems for which they could not find help. We will also monitor the user forums and measure the frequency of questions. Once the number of users exceeds ~1000, we will consider hiring part-time support staff.

Poster templates for conferences

Our second project will be to develop a web application that generates custom poster templates for Powerpoint, Keynote, and Illustrator. These are the most common tools used to create posters presented at scientific conferences. Our web app will produce templates tailored to the users' needs (e.g., meeting the requirements of a particular conference or venue) and that follow the best-practices and customs of their field. Our team has the design expertise, technical competence, and experience with scientific presentation standards to produce these templates in house. Funding would allow us to commission additional design work and to hire programmers to expand the web interface.

Once our web application template service is well established, we will scale the service along multiple dimensions to earn revenue. The base service will be free, but using the FedEx Kinko's

API, we will integrate printing and delivery, charging the user a small fee to have the poster printed and delivered to their conference venue or hotel. We will also integrate a design editing service, described in more detail below.

Reward: The potential user base is large (~1M posters per year) and there is only one major competitor in this space: PosterGenius, a company that provides a Powerpoint-like tool for creating posters. Their templates are not informed by the principles of good design.

Risk: Many conferences have specific size and layout requirements. It is necessary that our templates meet these requirements. There is a risk that our templates will at some point recommend an incorrect size. We can mitigate this risk by making it clear to the user that it is their responsibility to check the requirements of their conference and venue. Another mitigation technique is to allow conference-organizers to update our recommendations when the requirements change.

Design editing services

Our third project is even more ambitious: we will try to integrate design services into the workflow of postermakers and journals. This will begin small, as a series of educational blog posts that document the redesign of certain figures. It will grow through a series of documented collaborations between high-profile scientists and designers (see section below, "Collaborations between scientists and designers"). Finally, we will integrate design services into our poster template app and website, through which users will be able to purchase poster and figure editing services, analogous to copyediting (https://wordy.com/). This will provide us with income to sustain and grow the company, while furthering our mission. Our team has organizational experience in hiring and managing designers.

Reward: These design services have the potential to scale beyond our web application to the standard workflow of journals, who currently hire or commission work for similar services such as copyediting. Design services would provide sustainable revenue.

Risk: Design editing services scale poorly because each new figure or poster requires additional time to review and edit. There is risk because hiring a full time designer is costly. To mitigate this risk, we will begin by hiring freelance designers on a per-project basis. Slowly, as the service grows, we can expand to hiring part-time or full-time design staff.

AUDIENCE-BUILDING PRODUCTS, PROGRAMS, AND SERVICES

We also have a number of ideas for projects that will build the organization's visibility among scientists and the general public:

Collaborations between scientists and designers

We plan to pair up high-profile scientists and information designers to work together to create visualizations for ongoing research projects or manuscripts. We will then write about the process and outcome. For the kinds of complex data sets increasingly common in the biological sciences, the scientists doing the biology do not have the resources to learn — or create — sophisticated visualization techniques, opening an exciting window for collaboration. Success is a collaboration where, together, a scientist and designer manage to create a visualization that reveals a new scientific truth, leading the pair to co-author the resulting paper. Our team has a deep network of both designers and scientists who we can recruit to participate in this effort.

Rewards: If successful, this is the sort of collaboration that could see widespread interest in both the design and science communities, receiving attention and press from both communities.

Risk: It is hard to anticipate the outcome of a project like this, one that depends so much on the particular research group, design firm, and their ability to collaborate. However, one of our team members, Bang Wong, in his position as creative director of the Broad Institute, has expertise in facilitating communication between artists and scientists (e.g., in creating Pathline[13]).

Book: 100 Scientists' Favorite Figures

Better visualization of scientific data also benefits the public understanding of science. We plan to create a coffee table book featuring 100 scientist's favorite figures, with each figure appearing alongside an ELI5 explanation (http://reddit.com/r/explainlikeimfive) that walks the reader through the science. At first glance, the figure will look beautiful and enigmatic. After reading further, the figure will take on meaning. We have contacted a number of well-known scientists who have told us about their favorite plots. The demo post was very popular, seeing 7,000 views over the first two days, for a site that had not otherwise built an audience.

Many people are scared of math and science and think that it's too hard for them to understand. A book like this can give people the chance to succeed at understanding challenging material, and will boost their confidence and comfort in approaching technical subjects. A recent book "Portraits of the Mind"[14] featured neuroscientific images alongside explanatory text and was enormously successful among scientists and the general public[15].

Our book will be funded by a Kickstarter campaign. We expect to raise $10–50k[16]. The initial design work for a 1-page spread will be commissioned. Full funding may allow us to expand our scope, finding even higher profile science journalists and scientists to provide figures and author explanations. If the first book is successful, we plan to follow it up with books featuring the nicest figures published in the previous year ("Best Figures of 201*x*"). We have also been in contact with editors from science magazines such as Wired and Scientific American, who have expressed interest in running a story on the best figures published in the previous year.

---

[13] Meyer, M., Wong, B., Styczynski, M., Munzner, T., & Pfister, H. (2010, June). Pathline: A tool for comparative functional genomics. In *Computer Graphics Forum* (Vol. 29, No. 3, pp. 1043-1052). Blackwell Publishing Ltd.

[14] Schoonover, C. E. (2010). *Portraits of the Mind: Visualizing the Brain from Antiquity to the 21st Century*. Abrams.

[15] http://www.carlschoonover.com/, section "Pixels, ink & waves" for press coverage

[16] Projections based on www.kickstarter.com/year/2012 and sales figures of similar projects such as http://www.amazon.com/Portraits-Mind-Visualizing-Antiquity-Century/dp/0810990334

Rewards: If successful, our book could bring both education and enjoyment to large swaths of the scientific community and the general public. It would also bring attention to our website and organization, increasing the number of scientists using our tools.

Risks: None of us have experience designing or producing a full book. We can do much of the curation and writing and layout in house, and we will use our initial funding for a sample treatment, but for the full process we will seek additional funding and hire experienced design professionals.

Interviews with people who have created great figures, revealing their design process

We have contacted a number of plotmakers who are interested in being interviewed. Our interviews will give plotmakers the space to describe why their figure is effective, and also what previous versions of the figure looked like, educating the reader on how a figure evolves. Success for each interview will be measured by the number of pageviews and the response in social media.

ACQUIRING RESOURCES AND BUILDING SUSTAINABILITY

Figured is an early stage startup that's on the hunt for its first taste of seed funding. We're hard at work producing some of the initial content. What follows is a plan for how we will acquire initial funding and use it to build a sustainable business model.

The initial $5000

We'll use the initial $5000 to launch the organization. We'll incorporate as a nonprofit ($1000, with help from iLab and Harvard's Transactional Law Clinic), purchase domain names ($100) and hosting ($200), do a bit of targeted advertising ($1000), commission design work for poster templates ($500) and a sample book spread for use in a Kickstarter campaign ($1500), and purchase printing ($500) to help us launch outreach projects like the *100 Scientists' Favorite Figures* book, which will help us to build a large audience. We have also applied to the Awesome Foundation and are looking for other sources of seed funding.

ROADMAP FOR BUILDING SUSTAINABLE INCOME

Year 1.

- Launch organization
- Acquire seed funding (President's Challenge, Awesome Foundation, etc.)
- Produce regular examples of primary content (interviews, tutorials, redesigns, etc.)
- Release toolboxes
- Launch Kickstarter campaign to fund the book

Year 2.

- Release book (Revenue source #1)
- Release poster template application
- Begin curating "Best of" series in journals and conferences
- Recruit data visualization evangelists to give demos to their labs

Integrate print and delivery service into poster application (Revenue source #2)

Year 3.

Launch editing service (Revenue source #3)