How Prototyping Practices Affect Design Results

Steven Dow
Stanford University | spdow@stanford.edu

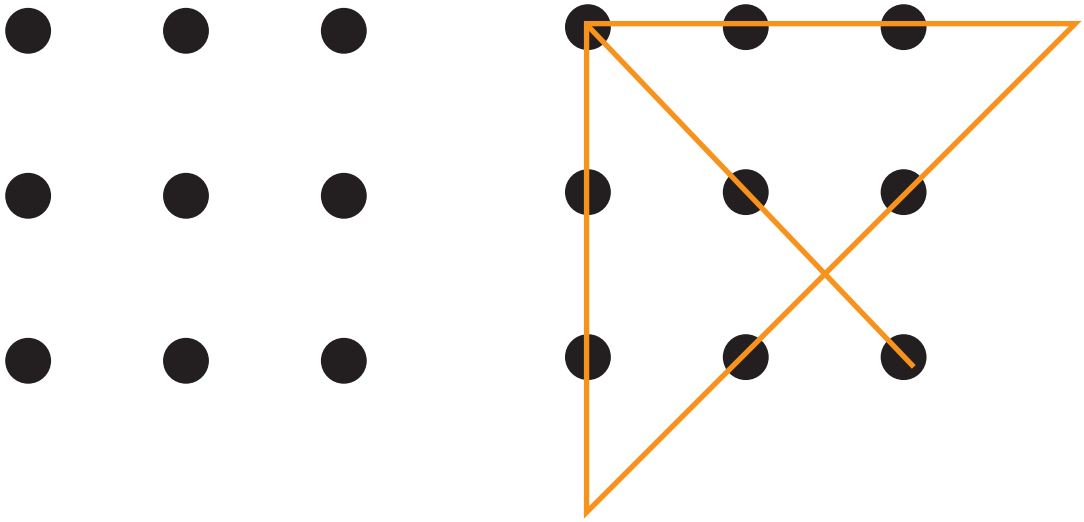
Iterate rapidly. Explore broadly. Gather feedback from multiple sources. Don't conflate ego with object. These pearls of wisdom state principles and values with which few designers disagree. Behind these mantras lie decades of human science research that can enrich our understanding of design. For the past few years, my Stanford colleagues and I have studied how and when design practices affect results. Our experiments begin to clarify why these designerly rules of thumb matter and where breakdowns can occur. By examining the cognition of comparison and the social psychology of sunk-cost reasoning, practitioners and educators can more fully realize the value of creating multiple alternatives throughout a design process.

There's a story about a ceramics teacher who divides his class into two groups. He tells one group they will be graded on quantity: Produce as many ceramics as possible. He tells the other half to focus on quality. He would base their grade on one good ceramic. It was reported "while the quantity group was busily churning out piles of work—and learning from their mistakes—the quality group had sat theorizing about perfection" [1]. Iterative deliberate practice led to better results. While some people resonate with this story, others point out that production schedules often discourage iteration in favor of realization. As Michael Schrage says, "It is hard to persuade companies that one more iteration costs less than a flawed product" [2]. It raises an empirical question about design practices: Is iterative prototyping valuable when time is highly constrained?

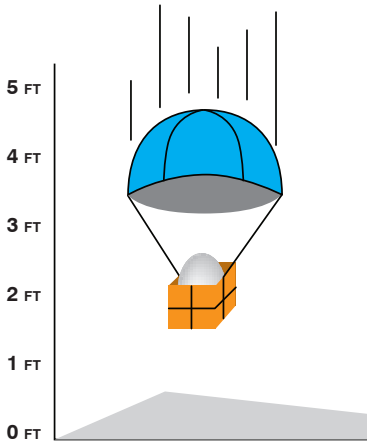
In 2005 Stanford opened the Hasso Plattner Institute of Design, also known as the d.school, to begin teaching a creative problem-solving process known as "design thinking." Imagine the ultimate cross-disciplinary studio space, full of configurable furniture, sketches, electronics, and ongoing student projects. On the wall, a sign reads "Believe in Process" (see Figure 1). This commitment to a particular strategy rests largely on faith. The goal for our research—sponsored by the Hasso Plattner design-thinking research program—has been to study the principles behind practices and to articulate how and why process affects creative results.

But how can we experiment on design practices? Scientists have

▸ Figure 1. Signs of faith adorn the walls of Stanford's d.school.



▸ Figure 2. The nine-dot problem—the origin of "outside the box."



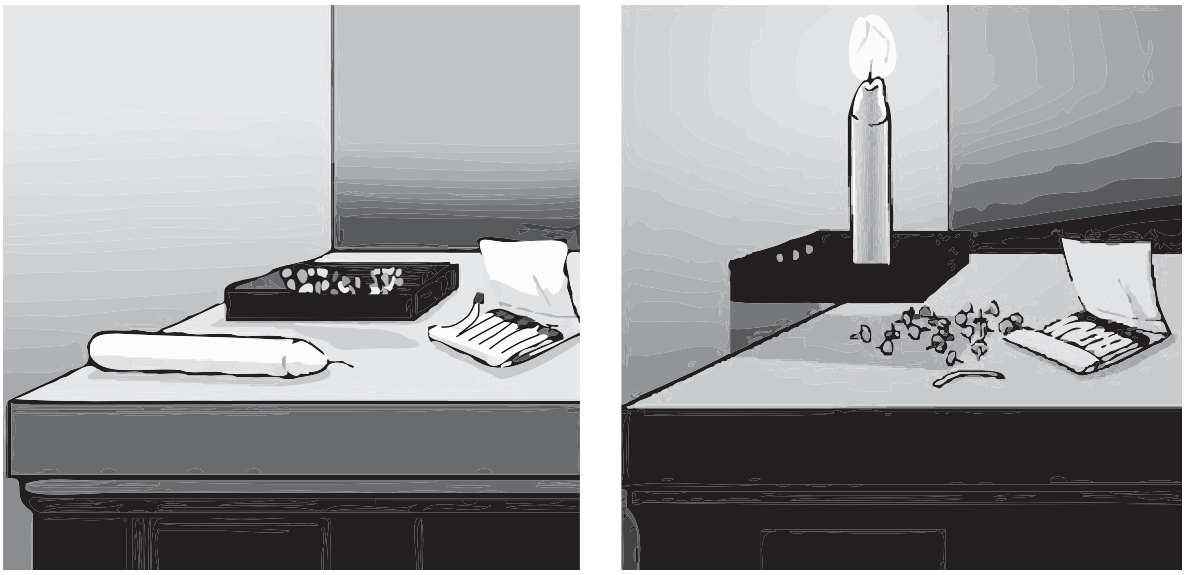
▸ Figure 3. Using the egg-drop problem to study design process.

long been interested in creativity. One classic creative insight experiment asks participants to connect nine dots with four straight lines without lifting the pen (see Figure 2a). The oft-missed insight is that lines must extend "outside the box" (see Figure 2b). As a proxy for creativity, scientists measure how long it takes people to solve the problem [3]. Other researchers ask people to invent alternative uses for objects. For example, a brick can be a paperweight, a boat anchor, a blunt weapon, and so on. As a creativity measure, scientists count up the number of valid and unique ideas. In Finke, Ward, and Smith's experiments on creative cognition, they ask people to sketch "creatures from another planet" [4]. Experts can then judge each idea on various criteria.

These approaches all contributed to an understanding of creativity. However, as my colleagues and I reflected on design and how practices affect the real world, we realized we needed a different Petri dish. Unlike the nine-dot problem, we wanted to give participants a problem in which outcomes cannot be defined by success/failure/right/wrong, but by what concept best fits the design context. More important, we wanted to measure creative impact using more objective criteria. How could we objectively contrast creative solutions? We found inspiration for our Petri dish in a classic middle-school activity: We had people design and construct vessels from raw materials to protect a raw egg's plummet (see Figure 3).

We tested the ceramics teacher's hypothesis about rapid iteration. Half of the 28 participants were encouraged to rapidly iterate; the other half focused on perfecting one design. As a dependent measure, we dropped the vessels from one foot up, then two feet, and so on, until the egg eventually cracked. Everyone came up with a different idea, with varying degrees of success. Our results showed quantitatively that—even under tight time constraints, when people have the tendency to focus on realization—rapid iteration led to better results [5].

What really surprised us was, independent of condition, participants tended to pick one idea and stick with it. The time constraints certainly contributed to participants' limited exploration, but people felt they had fully explored the concepts. Many talked about how "they could not see any other alternatives for the materials." Participants exhibited a psychological effect known as functional fixation, first studied by Karl Duncker back in the 1940s [6]. He did a series of experiments where



▸ Figure 4. Duncker's candle problem.

he presented a candle, a book of matches, and a box of tacks (see Figure 4). He told participants to affix the candle to the wall so that the wax did not drip down. The hidden insight is that the box of tacks can be used to support the candle. People often exhibit functional fixation in viewing the box's primary function as a container for tacks. It turns out that if the exact same materials are provided, but the tacks are left outside the box on the table, people are much more likely to solve the puzzle.

Could iteration, in some cases, increase fixation around a particular design? Prototypes elicit feedback, whether it's from the physical world, through simulations, or from colleagues and potential users. Feedback often frames subsequent actions around the existing solution; it provides a road map for how to improve designs but doesn't explicitly encourage exploration. We wondered if we could combat this fixation through a simple change in process.

Instead of just iterating solutions to a problem, what if people created and tested different designs in parallel? To answer this empirical question, we recruited people to participate in a design task where the solutions are creatively diverse and objectively measurable. This time, instead of egg-drop vessels, participants designed Web advertisements. Participants all created ads for the same client, *Ambidextrous*, a student-run magazine at Stanford. Online advertising presents an unprecedented opportunity to study the creative process. People of all skill levels can design simple Web graphics, and then the ads can be placed online to gather a host of performance metrics, such as click-through rates.

Study participants created an equal number of ad designs in the same time frame, but the process differed across conditions. Serial participants received a descriptive expert critique directly after each prototype. Parallel participants created multiple prototypes before receiving any feedback (see Figure 5).

The study found a parallel prototyping approach led participants to create better ad designs [7]. Web users clicked more parallel ads per appearance than serial ads. Not only did parallel ads generate more visitors to the *Ambidextrous* website, but those visitors also spent more time on the client site; the parallel ads did better at reaching the target audience. Moreover, independent expert raters—both ad professionals and the magazine editors—judged the parallel ads to be better than serial ads.

Why did a parallel approach lead to better results? One reason has to do with our fundamental human ability to draw contrasts. Dedre Gentner and colleagues' many experiments on comparison show people are much more likely to transfer a principle to a new context when explicitly prompted to draw contrasts between cases [8]. People do a better job of capturing knowledge when they compare. So perhaps viewing and thinking about two ads side-by-side helped people to understand and apply graphic design principles to subsequent designs.

Our study also revealed parallel participants created more diverse ad designs. Using a crowdsourcing platform, we asked independent judges to rate the similarity within participants' set of designs. The judges deemed serial ads to be very similar and the parallel ads, more diverse. The timing of feedback affected how broadly people generated ideas. By simply waiting for a critique, parallel participants had time to explore.

Moreover, more than half of the serial participants reacted negatively to the expert critique; none of the parallel participants felt this way. One serial participant complained, "[The expert is] telling me I am completely doing something wrong here…there was a period where the emotional response overwhelmed any positive logical impact that this ended up having." The critiques were not any more negative for serial participants, but they were perceived that way. Parallel participants showed several ideas at once, so they were less invested in any particular idea. By spreading investments, parallel participants were more open to diverse feedback.

The "parallel process" led to a number of learning and motivational benefits for individuals, but we wondered how such strategies could affect design interactions in groups. Many designers live by the principle "never go to a client meeting without a prototype." However, the presence of a concrete prototype may (for better or worse) focus the discussion on refining that idea rather than thinking more broadly. Moreover, people tend to polish prototypes to look good in front of colleagues. What kinds of dynamics occur when group members share multiple concepts as opposed to sharing only their best idea? We hypothesize that sharing multiple designs leads to better results because people will be more open to adopting and merging new ideas.

We recruited pairs of participants to work together on an ad-design task. Participants worked individually to create either multiple designs or a single design. Then they shared their design(s) with a partner and critiqued each other's ideas. Each person created a final ad design, which we launched in a Web ad campaign. The results show when participants create and share multiple prototypes—rather than devoting their time to polishing one concept—they produce better results [9]. Moreover, participants who shared multiple designs borrowed more specific features and provided higher reports of group rapport. By a number of indicators, the collaboration was more productive.

Iteration helps designers integrate feedback into their designs but may have some limitations. With only one idea on the table, designers may take feedback and use it to concentrate on improving design without considering other options. Creating multiple alternatives and getting feedback on them in parallel encourages designers to enumerate more diverse solutions, helps reduce fixation, discourages emotional investment in any one idea, and gives group members license to be more candid and critical of their own and others' ideas.

What do the results mean for the interaction design community? While parallel strategies may be common practice for seasoned designers, the rationale behind these practices often eludes people. Empirical evidence may help persuade disbelievers to adopt a culture of prototyping. More important, understanding the cognitive and social underpinnings of prototyping practices can challenge designers and non-designers alike to reflect on why and how their actions affect results. What follows are practical implications for how practitioners and educators can structure creative group work.

The "enlightened trial and error" of prototyping offers a way to explore the opportunities and constraints of new design contexts. As the egg-drop experiment illustrates, iteration helps people discover unknown variables and their interrelationships. This notion of "design as discovery" is particularly important when addressing wicked problems. Often, trying solutions helps uncover the right problems to solve. Iterative prototyping initiates a conversation with the space of design possibilities.

Perhaps overlooked is how *parallel* design provides value throughout a process, not just in early stages. In many domains, sketches can be produced quickly, but creating complete designs is costly and time consuming. When creating multiple comprehensive designs is impractical, designers can still prototype and share alternatives to subproblems. In Web design, for example, it may be infeasible to produce three very different functional sites, but invaluable to create and test strategically selected elements.

Indeed, many organizations practice alternative generation beyond the early brainstorming stages. When IDEO redesigned the shopping cart in the infamous *Nightline* report, they created four physical mockups around different user needs [10]. By putting multiple functional prototypes in front of customers, they could gather

the kind of comparative feedback needed to make effective design decisions.

Multiple alternatives help refine a design. When Dan Siroker served as director of analytics for the Obama campaign's website, data played a central role [11]. With millions of visitors arriving each day, they could easily evaluate different combinations of image media, color, and button phrasing and measure the impact on sign-up rates and donations. The results often surprised the campaign staff. For example, while many members of the team assumed Obama's stirring videos would lead to improvements, the data told a different story. Page versions with images rather than videos helped the campaign collect the most email addresses.

Further, our results provide an opportunity to reassess the dynamics of client interactions. Clients may not want to hire a design firm that presents several half-baked ideas. However, finely polishing a concept in anticipation of a big client presentation can lead to fixation and overinvestment. David Kelley, founder of the design firm IDEO, claims that part of his company's mission is to "train" clients about their approach. Effective design practice is not a straight march to a particular solution, but a process of trying out alternatives and tolerating shifts in direction.

Educators may look for ways to improve project-based design courses by teaching parallel practices. Scott Klemmer and his teaching assistants have largely revamped the curriculum for Stanford's course on HCI design around generating alternatives (See cs147.stanford.edu). In the first assignment, students brainstorm at least 20 ideas for how to redesign the "waiting in line" experience. In week three, students create storyboards for two points of view. In week eight, teams create multiple redesigns of functional prototypes and then gather data on these alternatives. When students form teams, they each bring multiple project ideas to help avoid imposing preconceived notions of their project's focus.

Parallel design is a strategy for coping with unpredictability. It's about avoiding commitment and signaling to others that the process could go in a number of directions. By enabling comparison, parallel design helps problem solvers reason about the implications of possible futures. While the variance of design alternatives necessarily narrows as deadlines approach, the parallel mind-set provides designers rational and emotional support throughout a design process.

In terms of research methods, our approach opportunistically leverages the modern Web. Using banner ad design and data analytics, we bring a fresh perspective to questions about human creativity, motivation, and teamwork. Our future experiments will examine how novices transition to experts, how reflective techniques affect fixation, how value-centered strategies reflect stakeholder perspectives, and how the dynamics of feedback affect client relations.

Design excellence goes beyond learning to sketch and prototype. It's not only a craft skill but also a way of thinking. How can the community harness the most value from these practices? With a deeper understanding of why prototyping practices matter, perhaps new pearls of wisdom emerge. Engage in conversation with the design space. Create prototypes that examine big unknowns. Discover problems in addition to solutions. Focus on interpreting and integrating feedback. Keep multiple possibilities in play as long as possible.

Acknowledgments

I want to thank Scott R. Klemmer and Daniel L. Schwartz for valuable research contributions.

ENDNOTES

[1] Bayles, D. and Orland, T. *Art & Fear: Observations on the Perils (and Rewards) of Artmaking.* Image Continuum Press, Eugene, OR, 2001.

[2] Schrage, M. *Serious Play: How the World's Best Companies Simulate to Innovate.* Harvard Business School Press, Cambridge, MA, 1999.

[3] Maier, N.R.R. Reasoning and learning. *Psychological Review* 38, 4 (1931), 332-346.

[4] Finke, R.A., Ward, T.B., and Smith, S.M. *Creative Cognition: Theory, Research, and Applications.* The MIT Press, Cambridge, MA, 1996.

[5] Dow, S.P., Heddleston, K., and Klemmer, S.R. The efficacy of prototyping under time constraints. *Proc. of the 7th ACM Conference on Creativity and Cognition.* (Berkeley, CA, Oct. 27-30) ACM, New York, 2009, 165-174.

[6] Duncker, K. *On Problem-solving.* Greenwood Press Reprint, Westport, CT, 1945.

[7] Dow, S., Glassco, A., Kass, J., Schwarz, M., Schwartz, D.L., and Klemmer, S.R. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. *Transactions on Computer-Human Interaction* 17, 4 (2010).

[8] Thompson, L., Gentner, D., and Loewenstein, J. Avoiding missed opportunities in managerial life: analogical training more powerful than individual case training. *Organizational Behavior and Human Decision Processes* 82, 1 (2000), 60-75.

[9] Dow, S.P., Fortuna, J., Schwartz, D., Altringer, B., Schwartz, D.L., and Klemmer, S.R. Prototyping dynamics: sharing multiple designs improves exploration, group rapport, and results. *CHI 2011: Conf. on Human Factors in Computing Systems* (Vancouver, Canada, May 1-12).

[10] The Deep Dive: Five Days at IDEO. Nightline with Ted Koppel, 1999.

[11] Siroker, D. How Obama raised $60 million by running a simple experiment. 2010; http://blog.optimizely.com/how-obama-raised-60-million-by-running-an-exp/.

ABOUT THE AUTHOR

Steven Dow is a postdoctoral scholar in the HCI Group at Stanford, where he researches HCI, creative problem-solving, prototyping practices, and crowdsourcing methods. He is recipient of Stanford's Postdoctoral Research Award and co-recipient of a Hasso Plattner Design Thinking Research Grant.

DOI: 10.1145/1962438.1962451