Surviving Our Success:
Three Radical Recommendations

Jared Spool

Founding Principal

User Interface Engineering

510 Turnpike St.

Suite 102

North Andover, MA 01845

e: jspool@uie.com

p: +1 978 327 5561

http://uie.com

Blog: http://uie.com/brainsparks

The world of usability practitioners is undergoing massive changes. I know because I read it in the New York Times.

Executive Attention

On July 19, 2007, the *New York Times* ran a story about eBay, reporting how profits are up 50%, yet the number of listings on the auction site fell 6% and the number of active users was stagnant at 83.3 million. According to the paper, eBay Chief Executive, Meg Whitman, reported to shareholders, "that her efforts to improve the user's experience should address these problems."

According to the *Times* article, Ms. Whitman said, "In the next six months, you will see more changes to eBay than you have in the last two or three years, whether that is an improved search experience or fun things that make the site better, like Bid Assistant, which allows you to bid on more than one item without worrying that you will end up buying five iPods by mistake."

There's something exciting about hearing the CEOs of Fortune 2000 companies talk about making the user experience a major priority. Interestingly, user experience is now frequently discussed in boardrooms across the globe. Apple's success with the iPod and

iPhone and Netflix's success beating Blockbuster and Wal-mart in the home-delivered DVD business have shown executives how a well-crafted user experience can be a strategic advantage. Everyone now wants to be the Netflix or Apple of their industry segment.

Be Careful What You Wish For

We should be excited about this new attention to user experience. For years, we've been asking executives to pay attention to what we do. Now, they aren't just paying attention—they are asking for help.

Let's pretend that, in the next year, half of the Fortune 2000 companies each decide user experience is a top strategic priority and want to invest a million dollars into increasing their user research capabilities. ($1 million may seem like a lot of money, but to the average Fortune 2000 company, it's less than one-tenth of a percent of annual revenues.) This investment would create an immediate demand for 10,000 usability professionals. Being there are only 2,373 members of the Usability Professionals Association, many of whom happy with their current jobs, where will these people come from? How would the community satisfy the demand?

Taking Our Cue from CUE

Meanwhile, on the other side of the globe, a researcher named Rolf Molich has been conducting some interesting experiments. Called Comparative Usability Evaluations (or CUE for short), these experiments entail asking multiple usability teams to conduct evaluations of the same interfaces. The results turn common beliefs about usability practice on their head.

In March 2003, Rolf asked 17 professionals to participate in the CUE-4 study. He didn't ask just any professionals. He recruited some of the best in our field—real top-notch folks. Rolf asked each professional to evaluate the reservation-booking system for the Hotel Penn, a Flash-based design that had garnered much attention for its novel single-screen interface.

Like much of Rolf's work, CUE-4 has proven to be a wealth of interesting information about our field. One of the many interesting findings had to do with the severe problems discovered by the teams.

Out of the CUE-4 evaluations, the 17 teams produced a list of 61 problems they rated as catastrophic—serious problems needing immediate attention. Rolf gave the list of problems to the site's designers, who agreed these problems were serious and needed immediate attention.

Good News and Bad News

The good news is the teams found these problems using traditional research techniques, such as usability testing and design inspections. The bad news is only a single team reported each problem. Had the site's designers wanted to hire someone to find the 61 problems, they would've needed to hire all 17 teams.

Rolf's findings in CUE-4 confirm what other CUE studies have told us: when you ask two usability teams to evaluate the same design, you see very little overlap in the issues they discover and report. Evaluators find very different problems and, on the rare occasion when there is an overlap, they report them very differently. Clients get different results, depending on the evaluator they hire.

At first, this may not to appear to be too much of a problem. After all, if you hire different visual designers or information architects, you'll end up with different designs. Based on the skills of those professionals, you'll get different quality results. One would expect a more-skilled professional to produce different results than a less-skilled professional. Why should usability evaluators be any different?

However, evaluation is about identifying problems, much like, say, radiology. If you handed an X-ray to five different radiologists, would you be happy with five completely different diagnoses? Would you expect some overlap in the observations and inferences produced by each doctor?

The Feng Shui Connection
According to Wikipedia, Feng Shui is "the ancient Chinese practice of placement and arrangement of space to achieve harmony with the environment." The practice has produced an army of consultants, all touting their ability to make living and working spaces more harmonious, often for big fees.

Penn and Teller, a pair of Vegas magicians, wanted to see if there was something to the modern interpretation of this practice. They hired three renowned Feng Shui consultants to rearrange the furniture in the same apartment. Each consultant produced a completely different arrangement, often contradicting the rationale used by the others. (For example, one consultant claimed the color red had excellent "positive energy" properties while the other two claimed the presence of red objects would negatively "drain" the living experience.)

Because of their little experiment, Penn and Teller decided the modern practice of Feng Shui was conducted by scam artists, looking to extract huge fees out of unsuspecting individuals, all because each consultant produced completely different results from the others. They believed if there was something to this practice, the recommendations of the consultants should have significant overlap.

As demand increases for user research professionals, are the CUE results likely to haunt the profession in a similar way? Is it possible that executives will, upon discovering the small overlap in evaluation results from one team to the next, dismiss the entire practice as a fraud?

Explaining the Successes
The practice of integrating usability into the design process has been around since the early 1980s. In that time, there have been many successful projects, where information about users integrated into the design proved to show dramatic improvements in the resulting design, with increased user satisfaction and better business results. Many in the community can easily cite dozens of projects where everyone was happy with the results of the efforts.

Therefore, we know, from our experiences, that we're not a complete swindle. There is something to what we do. We help teams improve products.

Maybe the CUE results aren't looking at the right component? Maybe there is something else to what we do?

Stone Soup

Stone Soup is a traditional Scandinavian folktale, where a traveler enters a small village with hopes to find someone who will give him food for a meal. As he talks to each villager, he learns the village was hit by a drought and no one has enough food to feed themselves, let alone a stranger.

The traveler suggests they make a pot of "Stone Soup", pulling a stone from his bag. Never having heard of making soup from a stone, the villagers become curious.

"How do you make it?" one villager inquires.

"I'll need a pot and some water," says the traveler. A villager volunteers his stockpot, while another donates some water from her well.

"What will it taste like?" asks another.

"It would be better with some vegetables," suggests the traveler. Quickly, another villager volunteers some carrots; another donates a chicken; others give scraps from their pantry.

Soon a pot of stew has cooked up—enough to feed the entire village. Everyone enjoys the meal and much happiness results. As a gift to the village, the traveler graciously donates the stone to the community, so they can continue to have great meals together and continues on his way.

The Role of the Stone

In the story, the stone didn't make soup. It wasn't even an important component to the recipe.

However, it was critical to getting the villagers to pool their resources and work together. Without it, they would've continued to starve. The stone was a catalyst, getting people to work together.

Just conducting a usability test or producing a list of recommendations won't guarantee change in a design. However, when a team to shifts their attention to the needs of users, we uniformly see improvements. Is it possible the techniques are a stone in the soup and the real benefit comes from the focus on user needs?

Radical Recommendation #1: Stop Making Recommendations

If the real benefit of our work comes from the focus on user needs, maybe we should change our base practice. One place to start would be in making recommendations. Many practitioners see it as their job to produce a document with recommendations for change in the design, often because design teams request it.

In Rolf's CUE studies, teams regularly produce long lists of recommendations for change. In many cases, the recommendations are only lightly supported by the findings of the research. When recommendations don't have the support of the findings, they can appear as opinions of the recommender. Opinions, without data, become sources for conflict, often dividing the team instead of focusing members on a common goal.

Symantec's Manager of User Research and Testing, Meghan Ede, takes a different approach with her team of researchers. She's instituted a policy that each researcher should be willing to bet their personal life savings on every recommendation they make. If they aren't willing to place the bet, the recommendation doesn't go into the results reported to the team.

By forcing a personal investment into the recommendations, Meghan is ensuring the researcher carefully considers the evidence behind it. With this consideration, they can move it from the realm of opinion into one supported by the data uncovered during the research.

However, is Meghan's approach going far enough? If our true goal is to get the team to focus on the users' needs, maybe it warrants a more extreme approach?

What would happen if researchers stopped delivering recommendations altogether? Instead, the researchers' deliverable would elaborate the evidence, allowing the team to come to their own conclusions. When teams create their own recommendations, they end up with more ownership of the results. The process of negotiating the recommendations increases their familiarity with the findings of the research.

We tried this with our clients a few years ago as an experiment. We stopped delivering recommendations, instead only presenting the observational findings from the study. Each project contained a workshop where we collaborated with the team to produce the recommendations from the findings. We found the teams became more engaged with the research and their recommendations were often better than any we

might have suggested ourselves. This is now our preferred way of engaging with teams.

Radical Recommendation #2: Stop Conducting Evaluations

Over at The Mathworks, Mary Beth Rettger, who manages the user research team, has instituted an interesting policy: Every member of the Mathworks' engineering group goes through a one-day workshop on paper prototyping and usability testing. This includes the engineers, quality assurance teams, and product managers.

Something interesting happens when everyone knows how to conduct basic usability techniques: It becomes part of the culture. It is no longer a special activity that only a few understand. Anyone who needs a study can put one together and learn about their users.
The next step on this evolutionary path is to do the unthinkable: usability professionals should stop conducting usability research altogether. Push all the research onto the team themselves.

If a usability professional is conducting a study, it should be a red flag. Something has gone wrong. They should put down the clipboard and back away slowly.

Instead, the job of the professional should be to coach the team to do the testing themselves. Sure, they can help by making it easy to recruit participants and execute the study, but all the groundwork should be done directly by the team.

In our shop, we've adopted the following mantra: "Outsourcing your usability research is like outsourcing your own vacation. It gets the job done, but it misses

the point." When teams do their own studies, they are more likely to absorb the subtle information that surfaces about their users and the users' needs.

Over the last few years, we've changed our focus to be training teams to do their own testing instead of hiring us (or anyone else). These teams produce substantially better designs in a much shorter time. The immersive influx of user data really changes their approach to creating great products and services.

Radical Recommendation #3: Seek Out New Techniques

The practice of using usability tests to assess software designs started in the late 1970s. Inspections and heuristic evaluations were initially proposed in the 1980s. Field studies and contextual inquiry also had its origins in the mid-1980s. None of the techniques we commonly use today have changed much in more than 20 years.

Unfortunately, we can't say the same for the technology we're evaluating. It's changed dramatically in that period. Similarly, the user population is now substantially broader. In addition, the development teams are radically different. Everything has changed, except for the tools we favor.

If we're no longer responsible for putting together recommendations or for conducting the research ourselves, we can devote some of that 'found' time to innovating new techniques. Not necessarily new techniques for conducting user research. It should be new techniques for helping teams focus on user needs.

After all, focusing on user needs is what we're really about. The user research techniques are just the stone in the soup.

Surviving Our Success
Oscar Wilde once said, "In this world there are only two tragedies. One is not getting what one wants and the other is getting it."

It's possible the Fortune 2000 executives won't discover what a valuable strategic advantage a well-crafted user experience is to the organization. If that's the case, we can keep going down the road we've always gone down, continuing to do the things we've always done.

However, that's not what we want. We want them to realize that a great customer experience increases brand engagement, which, in turn, increases customer loyalty and word-of-mouth advertising. Then, like the Apple iPod and the Netflix movie service, the customers become a major part of making the product or service a success.

Yet, if the executives do what we want, are we ready? Can we produce 10,000 more practitioners in a year? Or, do we need to take a more radical approach—something that changes the fundamental ways we approach our work? (Would we still call ourselves "practitioners?")

The world of usability practitioners is undergoing tremendous change. I'll keep reading about it in the *New York Times*.

Jared Spool has built User Interface Engineering into the largest research organization of its kind in the world. He's been working in the field of usability and design since 1978, before the term "usability" was ever associated with computers.

Jared spends his time working with the research teams at the company, helps clients understand how to solve their design problems, explains to reporters and industry analysts what the current state of design is all about, and is a top-rated speaker at more than 20 conferences every year. He is also the conference chair and keynote speaker at the annual User Interface Conference, is on the faculty of the Tufts University Gordon Institute, and manages to squeeze in a fair amount of writing time.